

**RODRIGO DE VASCONCELLOS VIANA MEDEIROS**

**SAÚDE E SANEAMENTO: EFEITOS DA EFICIÊNCIA E DA DESIGUALDADE DE  
ACESSO AOS SERVIÇOS SOBRE A SAÚDE DE CRIANÇAS BRASILEIRAS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Economia Aplicada, para obtenção do título de *Doctor Scientiae*.

Orientadora: Maria Micheliana da Costa Silva

Coorientadora: Raquel Pereira Pontes

**VIÇOSA – MINAS GERAIS  
2022**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade  
Federal de Viçosa - Campus Viçosa**

T

M488s  
2022  
Medeiros, Rodrigo de Vasconcellos Viana, 1992-  
Saúde e saneamento: efeitos da eficiência e da desigualdade  
de acesso aos serviços sobre a saúde de crianças brasileiras /  
Rodrigo de Vasconcellos Viana Medeiros. – Viçosa, MG, 2022.  
1 tese eletrônica (105 f.): il (algumas color.).

Inclui apêndices.

Orientador: Maria Micheliana da Costa Silva.

Tese (doutorado) - Universidade Federal de Viçosa,  
Departamento de Economia Rural, 2022.

Referências bibliográficas: f. 93-98

DOI: <https://doi.org/10.47328/ufvbbt.2022.530>

Modo de acesso: World Wide Web.

1. Crianças - Saúde e higiene - Aspectos econômicos.  
2. Saneamento. 3. Hospitais - Eliminação de resíduos - Custos.  
4. Hospitais - Eliminação de resíduos - Eficiência. 5. Análise de  
regressão. 6. Econometria Espacial. I. Silva, Maria Micheliana  
da Costa, 1986-. II. Universidade Federal de Viçosa.  
Departamento de Economia Rural. Programa de Pós-Graduação  
em Economia Aplicada. III. Título.

CDD 22. ed. 338.433621

**RODRIGO DE VASCONCELLOS VIANA MEDEIROS**

**SAÚDE E SANEAMENTO: EFEITOS DA EFICIÊNCIA E DA DESIGUALDADE DE  
ACESSO AOS SERVIÇOS SOBRE A SAÚDE DE CRIANÇAS BRASILEIRAS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Economia Aplicada, para obtenção do título de *Doctor Scientiae*.

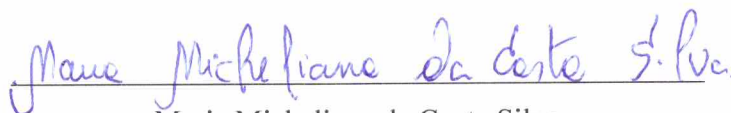
APROVADA: 05 de agosto de 2022.

Assentimento:



---

Rodrigo de Vasconcellos Viana Medeiros  
Autor



---

Maria Micheliana da Costa Silva  
Orientadora

## AGRADECIMENTOS

Inicialmente, quero agradecer aos meus familiares que sempre me apoiaram nas decisões que eu tomei. Deixo aqui o meu muito obrigado a minha mãe (Vilma), meu pai (Nei), minha avó (Zelita), minha madrinha (Sônia) e meu padrinho (Viana). Todos vocês foram importantes em algum momento, com um gesto, atitude ou palavras nessa caminhada que se iniciou lá em 2010, quando entrei na graduação.

Quero agradecer também aos professores que passaram pela minha trajetória acadêmica, principalmente na graduação, quando estava com incertezas sobre qual caminho seguir após terminar a faculdade. Aos professores Ralph, Vanessa, Abend e Eliane o meu muito obrigado pelas conversas e por abrir minha mente para várias possibilidades ainda na graduação.

Não posso deixar de agradecer as minhas orientadoras Micheliana e Raquel. Sempre muito dedicadas e com contribuições que foram valiosíssimas para que este trabalho fosse se aperfeiçoando. Sempre ouviram minhas ideias e apoiavam aquilo que deveriam apoiar e criticavam aquilo precisava ser melhorado, ou seja, desempenharam o papel de orientadoras com muito rigor, e o melhor, sem perder a amizade. Quando eu crescer quero ser assim também com meus alunos orientados.

Alguns amigos também merecem o meu agradecimento. Um grande abraço para meu amigo João Gabriel, que sempre estávamos discutindo questões acerca da academia e também sobre nosso Flamengo que nos deu tanta alegria nesses últimos anos. Muito obrigado pelos momentos de estudo e descontração nesse período. O meu muito obrigado também para o Thiago. Desde a graduação sempre conversando sobre economia, modelos matemáticos, estatística, naruto e outros animes. Agradeço pelos bons momentos durante todo esse tempo e mesmo distante sei que posso contar com você.

Não posso terminar meus agradecimentos sem agradecer a minha esposa Marina que, não sei por qual motivo, sempre disse para eu ficar tranquilo, pois eu ia conseguir alcançar as minhas metas. Muito obrigado por acreditar em mim durante todo esse tempo, as vezes eu não tinha certeza se eu ia conseguir, mas você estava sempre ali dizendo que eu ia conseguir e que eu ia encontrar uma solução para fazer o modelo dar certo.

Por fim, agradeço a CAPES, pois o presente trabalho foi realizado com o apoio da Coordenação de Pessoal de Nível Superior – Brasil – Código de Financiamento 001.

## RESUMO

MEDEIROS, Rodrigo de Vasconcellos Viana, D.Sc., Universidade Federal de Viçosa, agosto 2022. **Saúde e Saneamento: efeitos da eficiência e da desigualdade de acesso aos serviços sobre a saúde de crianças brasileiras.** Orientadora: Maria Micheliana da Costa Silva. Coorientadora: Raquel Pereira Pontes.

A presente tese teve como objetivo investigar a relação entre saúde e saneamento sob duas óticas distintas. Em um primeiro momento, foi analisado os efeitos de transbordamentos da eficiência de custos das operadoras dos serviços de água e esgoto sobre as hospitalizações de crianças com até 5 anos de vida nos municípios brasileiros. Utilizando uma base de dados que compreende o período de 2007 a 2018, a estratégia empírica adotada foi construir uma fronteira de eficiência estocástica que desagrega a eficiência em fatores variáveis e persistentes e, logo em seguida, relacionar tais indicadores diretamente com a saúde infantil através de modelos econométricos espaciais. Os resultados encontrados evidenciaram que as firmas possuem melhores níveis de eficiência variável e que instrumentos regulatórios e jurídicos, além do desempenho financeiro, são os responsáveis por aumentarem a eficiência variável das operadoras. Ao relacionar os níveis de eficiência com as internações infantis, verificou-se de maneira inédita que tais variáveis possuem efeitos positivos na saúde infantil, isto é, reduzem a taxa de internações por mil habitantes nas cidades em que há melhorias na eficiência além de transbordar seus efeitos em municípios vizinhos. Isso significa que quando uma firma busca otimizar seus recursos, além de alcançar melhores níveis de eficiência, ela promove externalidades positivas que ultrapassam a sua fronteira de competência. Nos modelos espaciais também foi verificado que choques exógenos ao município também podem transbordar e afetar cidades em seu entorno. Na segunda parte da tese, com uma outra base de dados, o objetivo foi investigar o efeito da desigualdade de acesso ao saneamento, entre os meios urbano e rural, por meio das diferenças de dotação e retorno, sobre as internações em crianças com idade inferior a 5 anos, no âmbito municipal. Para alcançar o objetivo proposto, foi utilizada de forma inédita a decomposição quantílica para identificar quais características detinham influência sobre as diferenças nos casos de doenças ligadas ao saneamento inadequado (DRSI). Com os resultados obtidos foi possível verificar que todas as variáveis explicativas tinham efeitos mais expressivos nos quantis superiores da distribuição de casos de DRSI, tanto para a zona urbana quanto para a rural. Logo, quanto pior a situação do município, em relação ao número de casos por mil habitantes, mais relevante se torna o ambiente no qual as crianças vivem. Em relação aos resultados da decomposição quantílica,

ficou constatado que existem diferenças significativas de casos entre as zonas urbana e rural. Características como o percentual de domicílios com água insegura, esgoto a céu aberto e o baixo nível de escolaridade foram fundamentais para determinar os maiores casos por mil habitantes no meio rural. Entretanto, se as dotações referentes ao esgotamento por fossa e se o número de famílias com plano de saúde aumentarem na zona rural, essas diferenças podem ser amenizadas. Considerando essas características, o governo tem maior facilidade de impactar positivamente o status de saúde infantil se promover e buscar soluções alternativas que visem melhorar as condições de esgotamento no meio rural. Projetos de fossas biodigestoras podem ser uma alternativa desde que sejam bem planejadas para utilizar a mão de obra local para a sua produção, pois, além de gerar renda para a população, pode aumentar a sua conscientização em relação aos efeitos nocivos do saneamento inadequado nas condições saúde das crianças.

**Palavras-chave:** Saúde Infantil. Saneamento Básico. Ineficiência. Regressão Quantílica Incondicional. Econometria Espacial.

## ABSTRACT

MEDEIROS, Rodrigo de Vasconcellos Viana, D.Sc., Universidade Federal de Viçosa, August 2022. **Health and Sanitation: effects of efficiency and inequality in access to services on the health of Brazilian children.** Advisor: Maria Micheliana da Costa Silva. Co-advisor: Raquel Pereira Pontes.

This thesis aimed to investigate the relationship between health and sanitation from two different perspectives. At first, the effects of cost efficiency spillovers of water and sewage service providers on hospitalizations of children up to 5 years of age in Brazilian municipalities were analyzed. Using a database that covers the period from 2007 to 2018, the empirical strategy adopted was to build a stochastic efficiency frontier that disaggregates efficiency into variable and persistent factors and, from that, to relate such indicators directly to children health through spatial econometric models. The results found showed that the firms have better levels of variable efficiency and that regulatory and legal instruments, in addition to financial performance, are responsible for increasing the variable efficiency of service providers. When relating efficiency levels with children hospitalizations, it was found in a brand new way that such variables have positive effects on children health, that is, they reduce the rate of hospitalizations per thousand inhabitants in cities where there are improvements in efficiency in addition to overflowing their effects in neighboring counties. This means that when a firm seeks to optimize its resources, in addition to achieving better levels of efficiency, it promotes positive externalities that go beyond its competence frontier. In the spatial models it was also verified that exogenous shocks to the municipality can also overflow and affect cities in its surroundings. In the second part of the thesis, with another database, the objective was to investigate the effect of inequality in access to sanitation, between urban and rural environments, through differences in endowment and return, on hospitalizations in children under the age of to 5 years, at the municipal level. To achieve the proposed objective, the quantile decomposition technique was used in brand new way to identify which characteristics had an influence on the differences in cases of diseases linked to inadequate sanitation (DRSI). With the obtained results, it was possible to verify that all explanatory variables had a more expressive effect in the upper quantiles of the distribution of DRSI cases, both for urban and rural areas. Therefore, the worse the municipality situation, in relation to the number of cases per thousand inhabitants, the more relevant the environment in which children live becomes. Regarding the results of the quantile decomposition, it was found that there are significant differences in cases between urban and rural areas.

Characteristics such as the percentage of households with unsafe water, open sewage and low level of education were fundamental to determine the highest cases per thousand inhabitants in rural areas. However, if endowments related to sewage disposal and the number of families with health insurance increase in rural areas, these differences can be mitigated. Considering these characteristics, the government is more likely to positively impact children's health status if it promotes and seeks alternative solutions that aim to improve sewage conditions in rural areas. Biodigester cesspool projects can be an alternative as long as they are well planned to use local labor for their production, because, in addition to generating income for the local population, it can increase their awareness of the harmful effects of inadequate sanitation in children's health conditions.

**Keywords:** Child Health. Sanitation. Inefficiency. Unconditional Quantile Regression. Spatial Econometrics.

## SUMÁRIO

1. INTRODUÇÃO .....	9
1.1 Considerações Iniciais.....	9
1.2 Problema e sua importância .....	11
1.3 Contribuição para o tema .....	14
2 Objetivos.....	16
2.1 Objetivo geral.....	16
2.1.2 Objetivos específicos .....	16
3. REFERENCIAL TEÓRICO .....	16
3.1 Conceito de Eficiência e Função de Custo.....	16
3.2 Abordagem Econômica para Mensurar a Saúde Infantil .....	20
4. METODOLOGIA.....	24
4.1 Fronteira de Custo Estocástica para Dados em Painel e Modelagem da Ineficiência.....	24
4.2 Mensurando o impacto da ineficiência nas condições de saúde infantil.....	33
4.3 Medindo o efeito da desigualdade urbano-rural no acesso ao saneamento na saúde infantil.....	41
5. ANÁLISE DOS RESULTADOS .....	47
5.1 Eficiência de Custo no Saneamento e Saúde Infantil.....	47
5.1.1 Análise Descritiva.....	47
5.1.2 Estimção da fronteira estocástica e da ineficiência do serviço de saneamento básico no Brasil entre 2007 e 2018.....	49
5.1.3 Análise Exploratória de Dados Espaciais (AEDE).....	57
5.1.4 Efeitos da eficiência do serviço de saneamento básico sobre a saúde infantil .....	63
5.2 Desigualdade urbano-rural no saneamento e seus efeitos na saúde infantil .....	71
5.2.1 Análise Descritiva.....	71
5.2.2 Resultados Econométricos.....	78
6. CONSIDERAÇÕES FINAIS .....	90
REFERÊNCIAS.....	93
APÊNDICE A.....	99
APÊNDICE B.....	103

## 1. INTRODUÇÃO

### 1.1 Considerações Iniciais

A universalização do acesso à água potável, coleta e tratamento de esgoto é um desafio que incita a comunidade internacional de países desde os anos 1980, quando foi declarado que aquela seria a década do saneamento e da água. Naquele período, diversas metas foram traçadas para que toda população global obtivesse serviços básicos de água e esgoto (WHITTINGTON *et al.*, 2007).

Apesar dos avanços consideráveis desses serviços, a cobertura dos serviços de água e esgoto em países menos desenvolvidos ainda são significativamente menores se comparados com outras nações desenvolvidas e com isso milhões de pessoas sofrem com algumas Doenças Relacionadas ao Saneamento Inadequado (DRSI), que podem ser evitáveis. Cerca de 10% de todas as doenças no mundo estão ligadas a essa baixa cobertura de saneamento e água inadequadas. Infecções diarreicas e doenças provocadas por parasitas (Teníase, Esquistossomose, Amebíase, entre outras) causam, todos os anos, cerca de 1,7 milhões de mortes evitáveis. Aproximadamente 90% dessas mortes são em crianças menores que 5 anos de vida e localizam-se em países em desenvolvimento (FEWTRELL *et al.*, 2007).

A consolidação da temática do saneamento no mundo veio em 2010 quando a Organização das Nações Unidas (ONU) reconheceu que o acesso a água tratada e o saneamento básico é um direito fundamental dos seres humanos e os incluiu nos objetivos do milênio. Com isso ficou reconhecido que o acesso ao saneamento básico contribui para que outros Objetivos do Desenvolvimento do Milênio (ODM) sejam alcançados tais como a redução da mortalidade infantil e a desnutrição, maior número de matrículas nas escolas e a melhora da qualidade de vida em regiões de periferia (HUTTON *et al.*, 2014).

Em países de crescimento rápido, como o Brasil e China, a taxa de mortalidade por doenças infecciosas relacionadas ao saneamento inadequado reduziu ao longo dos anos 1990. Porém, a taxa de infecção por essas doenças, nesse mesmo período, não reduziu na mesma magnitude, revelando que o acesso a esses serviços ainda está incompleto, sugerindo que governantes e organizações civis deem maior atenção para a criação de soluções e melhorias em água encanada e coleta de esgoto, itens basilares para qualquer sistema de saneamento básico (JEULAND *et al.*, 2013).

Especificamente para o caso brasileiro, o primeiro grande esforço para a criação de serviços de água e esgoto em ampla escala foi em meados de 1970 com a criação do Planasa (Plano Nacional de Saneamento). O principal objetivo desse plano era ampliar

significativamente o acesso ao abastecimento de água para 80% da população e levar serviços de esgoto para a metade dos domicílios. Contudo, a situação de instabilidade econômica e política nas décadas subjacentes dificultaram a universalização das ideias do Planasa e, até o final do século XX, o avanço nesses serviços ficou praticamente estagnado (ROCHA, ROSSONI e FARIA, 2018).

Assim, com a descontinuidade do Planasa, o Brasil passou toda a década de 1990 e início dos anos 2000 sem uma política nacional de saneamento o que deixou a infraestrutura instalada insuficiente para a demanda da população, gerando um déficit em saneamento básico. Em termos legislativos, esse quadro só começou a se modificar quando, em 2007, a Lei nº 11.445 foi regulamentada, estabelecendo as diretrizes nacionais para o saneamento e criando o Plano Nacional de Saneamento Básico (Plansab). Nessa Lei, ficou estabelecido que o saneamento é o conjunto de serviços, infraestruturas e instalações operacionais de abastecimento de água potável, esgotamento sanitário, limpeza urbana e manejo de resíduos sólidos, drenagem e manejo de águas pluviais (BRASIL, 2007).

Com o marco legal do saneamento estabelecido através da Lei nº 11.445/2007, um novo ciclo da gestão dos recursos públicos voltados ao esgotamento sanitário, água encanada e rede de infraestrutura foi inaugurado e programas nacionais como o Programa de Aceleração do Crescimento (PAC) passaram a destinar boa parte dos investimentos para a ampliação desses serviços, resultando em grandes avanços nas atividades de saneamento. Em 2018, segundo o Instituto Trata Brasil (ITB, 2018), o número de pessoas atendidas com abastecimento de água alcançou cerca 166,6 milhões o que significa um aumento aproximado de 2% ao ano comparado com o ano 2004 (data base para a pesquisa).

Contudo, mesmo com os avanços recentes, ainda existem indicadores que necessitam de atenção. De acordo com o ITB (2018), o Brasil é pior país da América Latina em quesitos básicos como esgotamento sanitário e disponibilidade de banheiro. Além disso, atualmente ainda existem cerca de 33 milhões de pessoas (aproximadamente 17%) que vivem sem água encanada em suas residências e mais da metade das residências (cerca de 55%) continuam sem esgotamento sanitário. Com isso, ainda persistem muitos casos de DRSI que acabam prejudicando não apenas a saúde infantil, mas também o desenvolvimento da própria criança em seu ambiente escolar e social. Esses dados ressaltam o quão atual e relevante é a temática do saneamento básico, incitando e justificando o surgimento de trabalhos que busquem novas evidências do impacto do saneamento inadequado nas condições de saúde da população nos municípios brasileiros.

## 1.2 Problema e sua importância

O acesso inadequado ou, em muitos casos, a falta de saneamento é um potencial causador de impactos negativos na saúde da população como um todo, ocasionando, sob a ótica econômica, externalidades negativas e constituindo-se em um complexo problema de saúde pública. Do ponto de vista macroeconômico, os impactos da falta de saneamento adequado podem reduzir, em média, 2% o PIB per capita dos países. Isso, em partes, pode ser explicado pela redução do desempenho produtivo dos trabalhadores com DRSI ou daqueles com filhos acometidos por essa doença têm seu desempenho produtivo afetado onde, em muitas situações, optam por faltar ou até mesmo são demitidos por não cumprirem as tarefas estabelecidas. Além disso, parte das despesas direcionadas a custear os tratamentos das DRSI poderiam ser aplicadas em outros setores da economia. (HUTTON *et al.*, 2014).

Pelo enfoque da saúde pública, Fonseca e Vasconcelos (2017) denotam o elevado número de internações por infecções causadas por Tracoma, Esquistossomose, Hepatite A, doenças de pele, Teníase, infecções diarreicas, entre outras, típicas de ambientes com falta de saneamento. Segundo Teixeira *et al.* (2014), até 2013, o gasto médio anual dos hospitais públicos com internações devido a essas doenças foi de, aproximadamente, R\$ 2,11 bilhões de reais. Somando as internações com as consultas médicas, a despesa total com DRSI compromete cerca de 3% dos gastos do Sistema Único de Saúde (SUS). Em um contexto onde cada vez mais se tem um orçamento estrangulado, qualquer redução de despesas com doenças evitáveis pode gerar benefícios para uma melhor alocação de recursos.

Os impactos da falta desses serviços na saúde da população se tornam mais evidentes quando analisado o número de mortes por mil habitantes de crianças menores que 5 anos de vida. Segundo Montgomery e Elimelech (2007), aproximadamente 60% da mortalidade infantil mundial é causada pelas doenças ligadas ao saneamento inadequado. Com uma base de dados referente as estatísticas de mortalidade infantil em 2004, os autores mostraram que em países com baixo desenvolvimento (África sub-saariana) a morte de crianças chega a 24 por mil habitantes. Dados da ONU de 2011 apresentaram alguma melhora para esses países, entretanto, ainda permanece uma média de mortes de crianças superior a 20 por mil habitantes. As estatísticas de países em desenvolvimento, incluindo o Brasil, obtiveram melhorias ao longo dos anos (taxa de mortalidade próxima de 10 por mil habitantes), porém ainda estão muito longe de atingir os padrões obtidos em países desenvolvidos. Nestes países, a mortalidade de crianças menores que 5 anos de vida não ultrapassa 3 por mil habitantes (ONU, 2021).

Dentre todas as formas capazes de veiculação dessas doenças (parasitoses, infecções diarreicas, doenças de pele, entre outras), o contágio pela via hídrica é a que possui maior conexão com o saneamento. Segundo Uhr, Schemel e Uhr (2016), entre 2007 e 2010, o número de internações de crianças com até 5 anos no Brasil por infecções (amebíase, giardíase, esquistossomose, entre outras) diarreicas causadas por parasitas ligados a falta de saneamento chegou a 60 mil casos por ano. Muitas dessas hospitalizações e, conseqüentemente os óbitos, poderiam ser evitados com o aumento do acesso aos serviços de água encanada e esgoto. É neste ponto que, segundo Ferro, Romero e Castiglioni (2011) a eficiência pode se tornar uma variável chave, uma vez que para uma empresa (privada ou pública) ser capaz de aumentar a produção de tais serviços é necessário que sua gestão de custos (mão de obra, equipamentos, entre outros) seja eficiente, caso contrário a redução de custos causaria uma redução na oferta dos serviços. Com isso, segundo os autores, uma eficiência nesses serviços significaria ser eficiente em salvar vidas por mortes vinculadas às DRSI.

No Brasil, ainda é um grande desafio garantir que todos tenham acesso ao saneamento básico, pois, assim como em outros serviços públicos essenciais, o déficit existente se espalha por todo território brasileiro, deflagrando a dificuldade do poder público em garantir a igualdade e a qualidade desses serviços para a população. O déficit torna-se ainda mais visível quando se destaca as diferenças entre o meio urbano e rural. Cerca de 75% da população moradora da zona rural vive sem nenhum tipo de tratamento ou coleta de esgoto em suas residências. Apesar da população rural no Brasil flutuar em torno dos 15% do total populacional, esse valor, em número absolutos, representa quase 30 milhões de pessoas que não possuem qualquer tratamento de seu esgoto doméstico. As conseqüências desse fato são graves principalmente para as crianças que podem ter até 5 vezes mais chances de adoecerem por infecções diarreicas se comparadas com aquelas que habitam em regiões urbanas com saneamento (COSTA e GUILHOTO, 2012).

Outro fator essencial para o saneamento básico rural é a água tratada. Segundo Silveira (2013), em 92% das áreas urbanas o acesso a água é realizado através da rede geral de fornecimento, seja público ou privado. No meio rural, apenas 28% da população possui esse serviço o que gera insegurança no consumo desse bem, pois a obtenção da água se dará em situações impróprias. Esse baixo acesso no meio rural por tais serviços essenciais é considerado por Lopes e Moraes (2019) resultado da inadequação das matrizes tecnológicas para o meio rural que se baseiam no princípio da tarifação dos serviços sem considerar propostas alternativas para o meio rural.

Nesse contexto, fica evidente o modo como o déficit de saneamento acarreta em externalidades negativas para as condições de saúde de toda a população. Além disso, ainda existe a problemática quanto à alocação dos recursos destinados ao saneamento básico. Segundo Saiani e Azevedo (2018), a expansão desses serviços está sujeita ao aumento dos investimentos no setor que, se bem destinados, aumentam a eficiência na entrega. Porém, em muitos casos o aumento dos custos de expansão dos serviços não reflete uma melhora na qualidade, pois de acordo com Da Hora *et al.* (2015), durante a década de 1990, o setor sofreu com a dificuldade em coordenar investimentos públicos e regulação de contratos em razão da ausência de um plano nacional com diretrizes e metas bem definidas. Isso resultou em uma prestação de serviços ineficientes com baixa qualidade na água que chega até as residências, interrupções nos serviços além de perdas por vazamentos na rede de distribuição. A ineficiência destes serviços provoca o crescimento dos gastos com saúde pública que objetivam combater os sintomas e agravos de DRSI, além de um gerenciamento falho dos recursos públicos (FERREIRA *et al.*, 2016).

Diante dessa realidade, compreender a forma como que o saneamento inadequado afeta as condições de saúde das crianças se torna importante no contexto atual por, pelo menos, três razões. Em primeiro lugar, apesar de se mostrar fundamental para as condições de vida humana, o saneamento parece não ser prioridade para governantes brasileiros, pois dados recentes do ITB (2020) mostram que apenas 41,5% dos municípios possuem plano de saneamento básico regulamentado ou não.

Em segundo lugar está a falta de confiança na disciplina jurídica no setor de saneamento brasileiro por parte das empresas privadas e outras organizações, provocando insegurança quanto aos benefícios econômicos que o investimento no setor pode trazer. Faltam sólidas garantias sobre o prazo para exploração, influência sobre os preços, entre outros aspectos. Pelo lado da iniciativa pública, segundo Hutton *et al.* (2014), o tema saneamento é menos atraente se comparado com pautas da saúde e educação para a mídia, provocando nos formuladores de política pública certa desmotivação em tornar tal tema um dos principais nos debates eleitorais o que justifica, em partes, o gasto entre 2% e 5% do PIB com saúde e educação, enquanto o gasto com saneamento, em média, alcança 0,7%.

Em terceiro lugar, além da falta de apoio político, o setor sofre com a falta de financiamento de projetos. No contexto econômico-fiscal atual não é novidade afirmar a dificuldade financeira que os municípios e estados brasileiros enfrentam, sendo muitas vezes “socorridos” pela União. Hutton *et al.* (2014) apontam que a falta de recursos para o financiamento do saneamento é frequentemente associada a falta de planos que detalhe os

custos e os benefícios que as ações irão promover. Além disso, pode-se apontar que a baixa capacidade técnica para a formulação dos planos também contribui para a ausência dos mesmos, ou ainda, transformando-os apenas em documentos administrativos sem nenhuma funcionalidade prática. Desse modo, o planejamento para entender previamente os impactos do saneamento pode estimular novos financiamentos além de atrair investidores interessados em ações que proporcionem benefício social.

Nesse sentido, torna-se relevante responder às questões de pesquisa: Quais são os efeitos da eficiência de custo dos serviços de saneamento básico na saúde infantil? Existem efeitos de desigualdades de acesso, entre as áreas urbanas e rurais, e nos indicadores de saneamento dos municípios brasileiros sobre a saúde de crianças entre 0 e 5 anos?

As informações incompletas sobre os planos municipais somado ao cenário econômico, falta de planejamento e contratos mal redigidos tornam o avanço do setor de saneamento lento e duvidoso. Logo, a relevância desse estudo emerge não somente para expor os benefícios da melhoria dos serviços de saneamento, mas também para avaliar a eficiência de custo atual para que não apenas os projetos de financiamento tenham uma base sólida de informações, mas também para mostrar as potenciais externalidades positivas que um investimento neste setor pode trazer para a saúde infantil.

### **1.3 Contribuição para o tema**

Em um contexto global, alguns autores como Pearson e Mcphedran (2008), Isunju *et al.* (2011), Van Minh e Hung (2011) e Montgomery e Elimelech (2007) já abordam em seus trabalhos os benefícios socioeconômicos das melhorias no saneamento e enfatizam que a ausência desses serviços acarreta diversas doenças que prejudicam sobretudo a saúde de crianças. Entretanto, as análises frequentemente são orientadas para a redução da mortalidade infantil e para questões nutricionais. Além disso, as soluções propostas pelos autores são o aumento dos investimentos para o saneamento ou políticas voltadas para promoção da saúde como a conscientização da higiene pessoal, sendo desconsiderado o impacto da eficiência dos serviços prestados. Logo, é possível perceber que, mesmo em nível mundial, ainda há espaços sobre o tema que precisam ser explorados.

No Brasil, os trabalhos sobre saneamento também seguem a tendência internacional. Existem diversos autores como Gross *et al.* (1989), Teixeira, Heller e Barreto (2007) e Barreto *et al.* (2010) que tratam dos fatores de riscos na ausência de água tratada, esgotamento e os impactos positivos da melhoria nesses serviços nas condições de saúde de crianças de

algumas regiões específicas do país. Assim como os trabalhos internacionais, os principais resultados enfatizam questões nutricionais e a queda na mortalidade infantil, carecendo de maiores explicações sobre as potencialidades na saúde pública como a redução de internações.

Conforme percebido, a maioria dos trabalhos abordam o saneamento e sua relação com a saúde de crianças em termos de como o acesso a esses serviços pode contribuir para a redução das taxas de mortalidade infantil. Embora seja um indicador relevante, ele não é capaz de capturar efeitos nas condições de saúde de crianças hospitalizadas. Outro aspecto pouco explorado é com relação a eficiência na entrega desses serviços à população. A discussão sobre as formas de melhoria na eficiência se concentra em observar se é a iniciativa pública ou privada que fornece os melhores serviços de água e esgoto, deixando de investigar as externalidades que a provisão eficiente dos serviços pode trazer. Assim, este trabalho também busca contribuir nesta área ao relacionar diretamente os índices de eficiência dos serviços de saneamento com as hospitalizações causadas por DRSI em crianças de até 5 anos de vida.

Além disso, a dependência espacial é um tema que carece de maiores debates sobre que tipo de relação existe entre esses serviços e a saúde da população. Na literatura nacional, embora trabalhos como Rodrigues, Venson e Camara (2019) tenha diagnosticado uma forte dependência espacial entre os serviços de saneamento (água, coleta de lixo e esgoto) pouca importância foi dada aos impactos dessa dependência nas condições de saúde. Scriptore (2016) e Silva *et al.* (2020) mostraram que as condições inadequadas de saneamento provocam efeitos espaciais que podem aumentar a mortalidade infantil e a incidência de internações nos municípios brasileiros. Nesse sentido, o presente trabalho contribui para o tema ao avançar nesta área, trazendo algumas evidências espaciais de como as políticas voltadas ao saneamento podem afetar a eficiência dos serviços e transbordar efeitos sobre as internações por DRSI nos municípios brasileiros.

Em relação ao debate sobre as desigualdades de acesso aos serviços de saneamento, muitos estudos (SAIANI e GALVÃO, 2011; SAIANI e TONETO JR, 2006; VENSON, JACINTO e SBICCA, 2019) exploram o tema a partir de dados agregados, investigando a possibilidade de convergência no acesso, no sentido de redução das desigualdades, ou analisam o cumprimento de metas estabelecidas na legislação. A dificuldade dessas análises agregadas é que, dentro de cada região, existem diversas disparidades entre os próprios municípios. Além disso, os dados empregados são censitários (coletados a cada 10 anos) o que pode fornecer um retrato defasado sobre a situação. Assim, este trabalho busca contribuir para o tema ao realizar uma análise desagregada, em nível municipal, com uma base de dados

anual para que a desigualdade de acesso ao saneamento seja analisada pelo diferencial de atendimento entre os meios urbano e rural, segundo características específicas de cada localidade.

## **2 OBJETIVOS**

### **2.1 Objetivo geral**

O presente estudo tem como objetivo geral avaliar o impacto da eficiência e dos indicadores de saneamento básico, bem como a desigualdade de acesso entre as áreas urbanas e rurais, nas internações e casos causados por DRSI, em crianças entre 0 e 5 anos de idade, nos municípios brasileiros.

#### **2.1.2 Objetivos específicos**

- i) Analisar os transbordamentos da eficiência de custo de serviços de saneamento básico nas internações em crianças menores de 5 anos de idade nos municípios brasileiros;
- ii) Investigar o efeito da desigualdade de acesso ao saneamento, entre os meios urbano e rural, por meio das diferenças de dotação e retorno, sobre os casos diagnosticados de DRSI em crianças com idade inferior a 5 anos, no âmbito municipal.

## **3. REFERENCIAL TEÓRICO**

### **3.1 Conceito de Eficiência e Função de Custo**

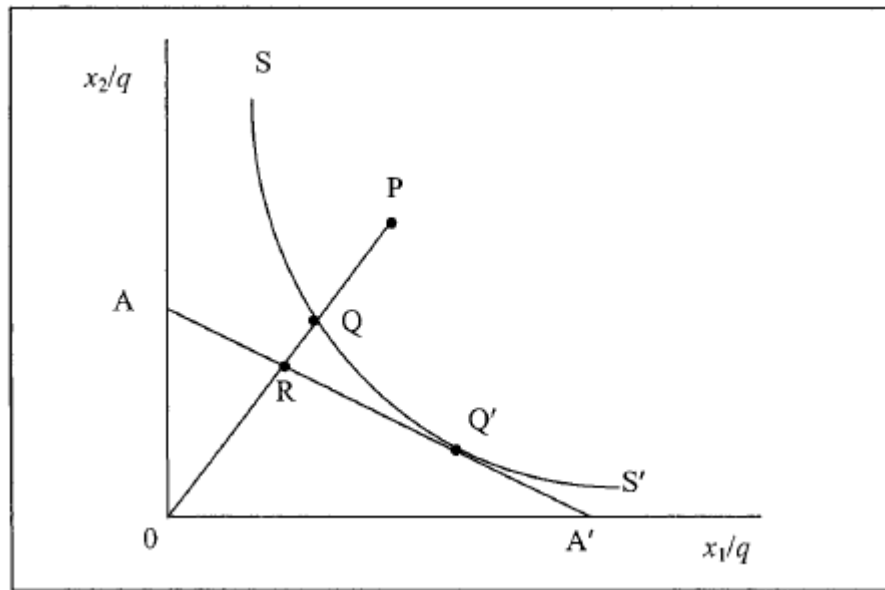
A definição de eficiência na ciência econômica está ligada às condições de operacionalização de um determinado sistema como, por exemplo, uma firma ou indústria. Em uma empresa, a eficiência consiste em empregar os fatores de produção da melhor maneira possível para que os resultados obtidos (bens ou serviços) possam ser otimizados, dada a tecnologia disponível para a produção. A otimização pode ser entendida como um processo de maximização (lucros, bem-estar, entre outros) ou de minimização (custos, perdas, dentre outros) de um resultado específico (FERREIRA, 2005).

No entanto, essa definição de eficiência precisa ser mais prática para que problemas empíricos possam ser modelados. O trabalho de Farrel (1957) foi um dos primeiros a se

apropriar do conceito de eficiência e mostrou que esta medida era composta por dois componentes de eficiência: técnica e alocativa. A eficiência técnica reflete a capacidade da firma em obter o máximo produto, dado o conjunto de fatores de produção (eficiência orientada a produção) ou ainda manter um determinado nível de produção, porém gastando menos insumos (eficiência orientada a insumos) visando, dessa forma, a redução de custo. Já a eficiência alocativa consiste em utilizar os fatores de produção em proporções ótimas, dados os seus preços relativos, ou seja, combina os insumos produtivos de forma a se obter um custo mínimo. Quando essas duas medidas são combinadas, tem-se uma medida de eficiência econômica global.

Para exemplificar a aplicação dos conceitos de eficiência técnica e alocativa, Coelli *et al.* (2005) utilizam o sistema de coordenadas cartesiano para demonstrar como que essas medidas de eficiência podem ser construídas. A Figura 1 ilustra uma situação hipotética em que a eficiência é analisada sob a orientação para o insumo, ou seja, trata-se de uma eficiência de custo que se busca estimar uma fronteira de custo para a firma que produza a um determinado nível com o menor custo.

Figura 1. Eficiência técnica e alocativa.



Fonte: Coelli *et al.* (2005).

A Figura 1 ilustra uma firma que utiliza dois insumos ( $x_1$  e  $x_2$ ) para produzir um único produto ( $q$ ). Para simplificar a análise, a curva isoquanta  $SS'$  foi construída considerando que a firma apresenta retornos constantes de escala. Por definição, essa curva representa as combinações entre os dois insumos para determinar o nível de produção e, portanto, a fronteira de produção. Logo, qualquer ponto sobre a curva  $SS'$  representa uma produção tecnicamente eficiente como em  $Q$  e  $Q'$ . A reta  $AA'$  representa uma linha de isocusto onde existem várias combinações dos insumos produtivos que fornecem o mesmo custo total.

Caso a firma encontre-se no ponto  $P$ , ela não terá nem eficiência técnica e nem alocativa, pois ela estará operando acima da fronteira utilizando insumos além do necessário para produzir uma quantidade específica. O grau de ineficiência técnico no ponto  $P$  é mensurado pela razão  $0P/0Q$  e o grau de ineficiência alocativo é dado por  $0Q/0R$ . Quando a isoquanta tangencia a linha de isocusto  $AA'$  tem-se uma situação em que o ponto  $Q'$  é eficiente tanto do ponto de vista técnico como também alocativo, pois há a máxima produção com o menor custo.

A partir deste exemplo, pode-se adaptar a mesma ideia para uma situação em que se deseja estimar uma fronteira de produção, visando maximizar o produto da firma a partir de seus insumos disponíveis. Entretanto, considerando o contexto do saneamento, faz mais sentido as empresas buscarem minimizar seus custos sem reduzir a produção do que buscar aumentar o nível de produto. Isso ocorre pelo fato de a demanda ser uma variável exógena, ou seja, a procura por serviços de saneamento é determinada fora da empresa. A ideia é que se uma empresa é eficiente em custos, ela terá maiores recursos para realizar investimentos que,

futuramente, retornarão maiores lucros, ou seja, uma operadora que busca minimizar custos está preocupada com sua maximização de lucros. Além disso, conforme apontado por Da Silva e Sousa, Faria e Moreira (2007), a firma que opera no setor de saneamento dificilmente produzirá mais que a demanda, dado o caráter monopolista do setor.

Outra situação que sugere a adequação do problema de minimização de custos está ligada aos índices de atendimento. A partir dos dados disponíveis, não é possível determinar se a firma que opera o saneamento em um município está deixando de atender a demanda em função de não ter capacidade ou se aquela parcela da população que não é atendida está uma área irregular do município, o que inviabilizaria a concessão dos serviços no local. Dada essa complexidade, ao capturar a eficiência pelos custos pode-se identificar, mesmo que indiretamente, quais são aquelas firmas que possuem capacidade de ampliar a cobertura desses serviços dados seus custos, demanda, seu capital e tecnologia disponível (OHIRA e SHIROTA, 2005; SABBIONI, 2008; PONTES, 2019).

Com base no que foi exposto, o modelo teórico para estimar a função de custos das empresas de saneamento supõe que as firmas são minimizadoras de custos conforme a seguinte equação (GARCIA e THOMAS, 2001):

$$C = f(y, w, Z) \quad (1)$$

em que  $y$  representa o produto,  $w$  é o vetor de preços dos insumos usados e  $Z$  denota um vetor com variáveis de controle que podem afetar os custos das firmas. A equação (1) pode ser estimada tanto para o curto quanto para o longo prazo, diferenciando-se apenas pelo componente de custo fixo presente no curto prazo. Entretanto, Pontes (2019) destaca que, para o setor de saneamento, é comum que mesmo no longo prazo o insumo capital seja considerado como quase-fixo<sup>1</sup>, pois sua expansão pode ser altamente custosa e inviável para a firma realizá-la com facilidade. Assim, pode-se estimar a equação (1) considerando características de longo prazo e, então, recuperar a função custo de curto prazo, por meio da minimização dos custos.

Considerando que  $x_v = (x_i)'$  seja um vetor de insumos variáveis, o problema de minimização da função de custo variável condicionado ao capital ( $K$ ) será:

---

<sup>1</sup> Um insumo quase-fixo pode ser entendido como aquele que seu custo total de utilização pode ser parcialmente variável e parcialmente fixo. Ou seja, o capital será quase-fixo em um intervalo de produção  $(y_0, y_1)$  caso  $x_k = 0$  e  $y \in (y_0, y_1) \Rightarrow x_k(y) = x_k$  onde  $x_k$  é uma constante.

$$\min_{x_v \geq 0} \sum_{i'} w_{i'} x_{i'} \quad \text{s.a. } f(y, x_v, x_K; Z) = 0 \text{ e } x_K = 0 \quad (2)$$

O resultado dessa minimização será a recuperação da função de custo no curto prazo<sup>2</sup> que é apresentada na equação (3):

$$\begin{aligned} C_{cp}(y, w_v, w_K; K, Z) &= \sum_{i'} w_{i'} x_{i'}(w_v, y; K, Z) + w_K K \\ &= CV(w_v, y; K, Z) + CF \end{aligned} \quad (3)$$

onde o subscrito *cp* refere-se à função de custo para o curto prazo, CV é o custo variável e CF o custo fixo. Garcia e Thomas (2001) destacam que as propriedades da função de custo de curto prazo são as mesmas da função de longo prazo, ou seja, é não negativa e não decrescente em *y* e em *w* e homogênea de grau de 1, côncava e continuamente diferenciável em *w*. O fato do capital ser considerado como quase-fixo pode implicar que seu uso não necessariamente atingirá o custo o mínimo. Porém, a minimização da equação (2) resolve o problema para todos os fatores de produção e, no longo prazo, esses insumos podem ser ajustados instantaneamente. Assim, para que a função de custo de longo prazo seja recuperada corretamente, o capital tido como quase-fixo deve ser não crescente. Então, para que essa suposição seja satisfeita, no ótimo a condição de primeira ordem deve ser  $\frac{\partial CV(w_v, y; K, Z)}{\partial K} = -w_K$ , onde  $w_K$  é o preço do capital.

A importância em se avaliar a condição de primeira ordem para o capital está relacionada aos conceitos da teoria da firma. Caso a derivada parcial para o capital tenha sinal positivo, isso significa que a firma não está em seu caminho ótimo, sendo necessário que ocorram ajustes em seus insumos de produção para que alcancem o nível ótimo de utilização dos recursos.

### 3.2 Abordagem Econômica para Mensurar a Saúde Infantil

A partir da segunda metade do século XX, os estudos sobre a Economia da Saúde atentaram-se em entender como que a teoria da demanda poderia explicar a busca por cuidados médicos, assim como identificar possíveis falhas de mercado que eventualmente surgissem. Nesse sentido, alguns estudos como de Grossman (1972) emergiram propondo que as análises não incluíssem apenas variáveis relacionadas ao mercado de assistência médica

<sup>2</sup> Uma explicação detalhada sobre as derivações para a recuperação da função custo de curto prazo pode ser obtida em Santos, Lirio e Vieira (2009).

(como o preço e a oferta dos serviços), mas também incorporar o próprio status de saúde, as características dos indivíduos e o ambiente no qual estão inseridos.

Com o surgimento de diversos debates nessa área, a referência seminal para o tema foi proposta por Grossman (1972). Em seu modelo inspirado na teoria do capital humano de Becker (1964), o autor introduziu o conceito de saúde como um bem durável de capital que é herdado, mas que se deprecia ao longo do tempo. Sendo assim, é necessário que o indivíduo realize investimentos em saúde. Este, por sua vez, é tido como uma atividade onde os cuidados médicos são combinados com outros insumos para a “produção” de um novo status de saúde que seja, pelo menos, suficiente para repor a sua deterioração natural.

Quando o contexto de análise envolve a saúde das crianças, a herança em saúde é gerada a partir dos cuidados com o pré-natal, ou seja, a busca por serviços médicos para a manutenção e melhoria da saúde da mãe é responsável por gerar a função de produção de saúde infantil para o bebê ao nascer. Com isso, segundo Reichman *et al.* (2009) alguns fatores teoricamente importantes, mas tipicamente não observados<sup>3</sup>, tornam-se determinantes para explicar os motivos de as crianças nascerem com baixo peso e terem maior propensão de adoecer. Logo, para os autores, a maximização da função de produção da saúde infantil deve ser feita considerando as preferências dos pais, conforme mostra a equação a seguir:

$$U = U(C, H_i, H_p, L) \quad (4)$$

A equação (4) é uma função de utilidade que pode ser entendida como a utilidade em conjunto dos pais. Essa equação diz que a utilidade conjunta dos pais está em função dos bens consumidos ( $C$ ), do status de saúde infantil ( $H_i$ ), condições de saúde dos pais ( $H_p$ ) e outros argumentos relevantes materializados como lazer ( $L$ ). Dessa forma, a maximização da saúde infantil será estabelecida a partir de um processo de otimização em que as decisões/comportamentos dos pais irão afetar a variável de saúde  $H_i$  (REICHMAN *et al.*, 2009).

A dificuldade empírica em implementar um modelo teórico conforme proposto por Reichman *et al.* (2009) está no fato que, das diversas variáveis possíveis que podem afetar a saúde infantil durante o período pré-natal, são poucas aquelas que impactam estatisticamente o recém-nascido (LU *et al.*, 2003). Alguns trabalhos como Almond, Currie e Herrmann

---

<sup>3</sup> Tais fatores tipicamente não observados são o consumo de bebidas alcoólicas, cigarros e outras drogas que interferem na saúde da criança durante o pré-natal, porém como são fontes baseadas em “self-reported” podem não ser totalmente observadas.

(2012) e Almond e Mazumber (2011) destacam a importância dos fatores ligados ao pré-natal para a criança ao nascer, todavia, após seu nascimento, e ao completar seu primeiro ano de vida, são os fatores ligados ao ambiente que se tornam cruciais para o desenvolvimento da saúde infantil. Nesse sentido, as condições ambientais ligadas ao saneamento passam a ganhar destaque nas análises de saúde infantil e modelos que buscam maximizar o argumento  $H_i$  da equação (4) de maneira individual foram desenvolvidos.

Ausburg e Rodriguez-Lesmes (2018) desenvolveram um modelo para entender a relação entre o saneamento local com a saúde infantil, a partir de um quadro conceitual epidemiológico, demográfico e econômico. A base teórica está fundamentada no trabalho de Mosley e Chen (1984), que identificaram que os principais determinantes de sobrevivência e morbidades das crianças estão associados a educação das mães, renda familiar, fatores institucionais (infraestrutura de saúde e de saneamento) e normas e culturas locais. Combinando esses *insights* com a abordagem econômica de Currie (2000), Ausburg e Rodriguez-Lesmes (2018) estimaram uma função de produção de saúde infantil com o objetivo de determinar a quantidade ótima de serviços de saneamento demandada por uma unidade doméstica (ou família). Para tanto, os autores consideram que os pais buscam maximizar a seguinte função objetivo:

$$\sum_{t=1}^T E_t \beta^t U_{ivt} + B(A_{iv,T+1}) \quad (5)$$

Pela equação (5), a família  $i$  no local  $v$  é composta por pais que são altruístas e obtêm utilidade do status de saúde dos seus filhos e da herança,  $B$ , deixada para eles. Então, para um período  $t$  específico a utilidade será dada por:

$$U_{ivt} = U(Q_{ivt}, S_{ivt}, C_{ivt}, L_{ivt}; X_{ivt}, u_{1vt}, \varepsilon_{1vt}) \quad (6)$$

onde  $Q_{ivt}$  é um indicador de saúde infantil,  $C_{ivt}$  é consumo com outros bens e  $L_{ivt}$  é o lazer. A variável  $X_{ivt}$  representa as características observadas dos indivíduos,  $u_{1vt}$  são as características não observadas e  $\varepsilon_{1vt}$  é qualquer choque que possa afetar a utilidade. Além desses termos, fundamentados pelo modelo de Currie (2000), Ausburg e Rodriguez-Lesmes (2018) inserem o termo  $S_{ivt}$  que indica o ganho de utilidade que os pais obtêm por terem acesso a serviços de saneamento para si e suas crianças.

No modelo original de Currie (2000), os pais buscavam maximizar a equação (5) escolhendo a quantidade de  $Q_{ivt}, C_{ivt}, L_{ivt}$  restrito aos seus preços e salários observados. A ideia era que o estoque de saúde infantil  $Q_{ivt}$  poderia ser melhorado/aumentado com investimentos que dependeriam dos preços de mercado. Esses investimentos estavam associados a bens e serviços ligados a saúde da criança (consultas médicas, medicamentos, vacinas, entre outros). Assim, dados os preços desses investimentos, os pais se deparariam com um *trade-off* entre o consumo desses bens com os demais argumentos de sua função de utilidade.

Entretanto, Ausburg e Rodriguez-Lesmes (2018) partem da hipótese de que os agentes, neste caso os pais, já realizam decisões ótimas para todos os bens, exceto em relação a  $Q_{ivt}$  por motivos mais práticos do que teóricos. Dessa forma, os autores propõem a construção de uma função para a saúde infantil partindo dessa premissa. Além disso, existe uma complexidade em definir o *trade-off* proposto por Currie (2000) em termos empíricos, pois são os fatores exógenos à decisão dos pais que mais afetam a condição de saúde infantil e, entre esses fatores, está o saneamento. Assim, a saúde infantil é modelada conforme a equação abaixo:

$$Q_{ivt} = f(Q_{iv,t-1}, S_{ivt}; Z_{vt}, u_{2vt}, \varepsilon_{2vt}) \quad (7)$$

A equação (7) informa que a saúde infantil é uma função das condições de saúde passada em  $t - 1$ , dos indicadores de saneamento  $S_{ivt}$ , de características observadas  $Z_{vt}$  em sua região de residência e por fatores não observados  $u_{2vt}$  e qualquer outro choque não modelável  $\varepsilon_{2vt}$ . Inserido nas características observáveis ( $Z_{vt}$ ), Currie (2000) e Ausburg e Rodriguez-Lesmes (2018) apontam as condições do ambiente no qual a criança está inserida como um fator relevante. Os ambientes urbano e rural podem ser bastantes distintos em relação as suas características, podendo afetar diretamente a saúde infantil.

O efeito marginal esperado do saneamento  $E(\partial Q_{ivt} / \partial S_{ivt})$  deverá ser positivo se  $Q_{ivt}$  mensurar o *status* de saúde em termos mortalidade e morbidade. Em outras palavras, na medida que os indicadores de saneamento melhoram, também se observa melhorias nos indicadores de mortalidade e morbidade. Caso  $Q_{ivt}$  seja um indicador de saúde semelhante ao que Grossman (1972) define, ou seja, uma variável que captura o investimento em saúde como o número (ou valor dispendido) de consulta médicas, então  $E(\partial Q_{ivt} / \partial S_{ivt})$  deverá ser negativo uma vez que a melhora no saneamento trará benefícios ao status de saúde do indivíduo reduzindo a quantidade de consultas, reduzindo a necessidade de um novo

investimento em saúde com remédios e consultas particulares. Segundo Ausburg e Rodriguez-Lesmes (2018), tais resultados são verdadeiros considerando a suposição de linearidade na relação entre  $Q_{ivt}$  e  $S_{ivt}$ .

Com base nos conceitos de eficiência e função de custo, alguns autores (GALIANI, GERTLER e SCHARGRODSKY, 2005; FERRO *et al.* (2014); Da SILVA e SOUZA, FARIA e MOREIRA, 2007) realizaram a aplicação de trabalhos no setor de saneamento e buscaram compreender como que a operação desses serviços poderia impactar nas condições de saúde da população. O déficit em saneamento, a privatização, a regulação e a qualidade são temas que frequentemente aparecem relacionados a saúde sobretudo a infantil. A eficiência surge intrinsecamente dentro de cada um desses tópicos e, implicitamente, manifesta sua importância para as condições de saúde de crianças.

Quando analisada diretamente, a eficiência, se relacionada com indicadores de saúde, revela-se uma variável importante para promover a saúde infantil se ela for capaz de melhorar a qualidade dos serviços e, dessa maneira, um canal de transmissão para os benefícios do saneamento. Entretanto, é importante destacar que nem sempre a eficiência de custo conseguirá ter esse papel. Daí a importância de se investigar como que a privatização, regulação e qualidade podem influenciar na eficiência dos serviços de saneamento e como estes serviços interferem nas morbidades que acometem as crianças.

## **4. METODOLOGIA**

### **4.1 Fronteira de Custo Estocástica para Dados em Painel e Modelagem da Ineficiência**

A análise de eficiência pode ser aplicada utilizando diversos métodos e modelos. Os mais usuais são os modelos de Análise Envoltória de Dados (DEA), que é caracterizado por ser um método não paramétrico, e os modelos de Análise de Fronteira Estocástica (SFA), considerado um método paramétrico. A principal diferença que se observa entre esses métodos é que enquanto no primeiro nenhuma forma funcional é imposta ao conjunto de dados, o segundo supõe alguma forma funcional na sua estimação baseando-se em alguma função estabelecida previamente (MURILLO-ZAMORANO, 2004).

Em ambas as técnicas, podem existir vantagens e desvantagens o que torna difícil apontar qual delas é superior, em termos de se obter os melhores resultados. Todavia, algumas vantagens e desvantagens ficam evidentes tanto quando se utiliza a DEA quanto a SFA. Quando o horizonte no tempo é dispensável ou ainda invariante, a amostra é pequena e as unidades de observação apresentam grande homogeneidade (existem poucos ou nenhum dado

discrepante) a DEA apresenta bons resultados para as medidas de eficiência. Entretanto, seus resultados são altamente sensíveis a *outliers* e por ser tratado como um modelo determinístico todo choque exógeno que afete o gerenciamento dos recursos é desconsiderado e computada uma ineficiência para a unidade em análise (FERREIRA, 2005).

Já na análise de SFA, as principais vantagens incluem: a possibilidade do termo de erro ser dividido em ruído estatístico e componentes de ineficiência, permite realizar análise de causalidade, há pouca interferência de dados discrepantes nas estimações e permite um bom ajuste quando o horizonte no tempo é requisitado. Porém, entre as desvantagens está a necessidade de se impor previamente uma forma funcional para a fronteira, podendo causar a não convergência do modelo estimado. Além disso, é necessário realizar suposição sobre a distribuição das variáveis aleatórias o que pode levar a perda de informação sobre o comportamento natural dos dados em análise (MURILLO-ZAMORANO, 2004).

Com base no exposto, fica evidente que cabe ao pesquisador refletir sobre esses pontos para identificar qual dos métodos tem maior aderência a sua pesquisa. Considerando que o presente trabalho busca realizar uma análise para todos os municípios brasileiros ao longo de vários anos, para medir a ineficiência de custo empiricamente optou-se por utilizar uma fronteira de custo estocástica. Tradicionalmente, a equação (1) pode ser estimada por uma fronteira de custo estocástica conforme especificada por Aigner, Lovell e Schmidt (1977) em seu trabalho seminal:

$$C_{jt} = f(x_{jt}; \beta) + \epsilon_{jt} \quad (8)$$

em que  $x_{it}$  representa as variáveis que influenciam os custos da  $j$ -ésima firma no tempo  $t$  e  $\epsilon_{jt}$  é denominado termo de erro composto, onde  $\epsilon_{jt} = v_{jt} - u_j$ . O termo  $v_{jt}$  é o ruído aleatório do modelo e  $u_j \geq 0$  é denominado como a ineficiência técnica específica de cada firma  $j$  da amostra.

Contudo, a estimação da ineficiência pela equação (8) pode gerar alguns problemas, uma vez que qualquer heterogeneidade<sup>4</sup> presente nos municípios (condições de relevo, propriedades biológicas da água, clima, entre outras) e que pode afetar a produtividade da firma não é isolada no modelo, enviesando seus resultados. Desse modo, vários autores

---

<sup>4</sup> As firmas que operam em municípios que, por ventura, tenham um maior teor de salinidade na água ou ainda aqueles que sofrem com fortes secas podem, por exemplo, se deparar com uma situação em que seus custos com qualidade e potabilidade da água se tornem muito mais caros do que em outras cidades. Isso poderia levar a uma conclusão errônea que os serviços prestados estariam sendo ineficientes, em termos de custos.

propuseram soluções para contornar esse problema. Alguns trabalhos avançaram na modelagem de modelos de fronteira estocástica, pois permitiram que o termo de erro fosse decomposto três componentes (WANG e HO, 2010; CHEN *et al.*, 2014) ou em quatro componentes (COLOMBI *et al.*, 2011; KUMBHAKAR, LIEN e HARDAKER, 2014).

Os modelos que decompõe o erro em quatro componentes podem fornecer melhores resultados, pois, além de carregar o erro idiossincrático do modelo, permite que a ineficiência invariante no tempo (chamada de ineficiência persistente) seja separada da heterogeneidade (efeito firma) e de uma ineficiência transitória, denominada ineficiência variável (KUMBHAKAR, WANG e HORNCASLER, 2015). Isso pode se tornar útil para o contexto da eficiência no setor de saneamento, tendo em vista que as mudanças nas leis do saneamento, bem como o marco regulatório, podem alterar a eficiência de custos das firmas o que torna relevante a distinção entre os efeitos supracitados. Nesse sentido, optou-se por utilizar essa abordagem no presente trabalho seguindo principalmente os métodos de estimação proposto por Kumbhakar, Lien e Hardaker (2014). A forma genérica deste modelo consiste em reescrever a equação (8) como:

$$C_{jt} = \alpha_0^* + f(x_{jt}; \beta) + \alpha_j + \varepsilon_{jt} \quad (9)$$

onde  $\alpha_0^* = \alpha_0 - E(\eta_j) - E(u_{jt})$ ,  $\alpha_j = \mu_j - \eta_j + E(\eta_j)$  e  $\varepsilon_{jt} = v_{jt} - u_{jt} + E(u_{jt})$ . O termo  $\eta_j$  representa a ineficiência persistente da firma, ou seja, capta fatores que interferem na gestão e que são invariantes no tempo. Por suposição, sua distribuição é tal que  $N^+(m, \sigma_{\eta_j}^2)$  o que corresponde uma normal-truncada com média diferente de zero, pois o índice de eficiência por construção não pode ser negativo. O termo  $\mu_j$  com  $N(0, \sigma_{\mu}^2)$  se refere ao efeito não observado específico de cada firma e representa as diferenças de custo causada pela heterogeneidade não observada;  $u_{jt}$  é a ineficiência variável que reflete fontes variantes no tempo e que podem interferir nas boas práticas de gestão. Assim como a ineficiência persistente, sua distribuição segue uma normal-truncada com média não zero. Por fim,  $v_{jt}$  é o termo de erro estatístico do modelo com média zero e variância  $\sigma_v^2$ . É assumido que todos esses termos são independentes e identicamente distribuídos.

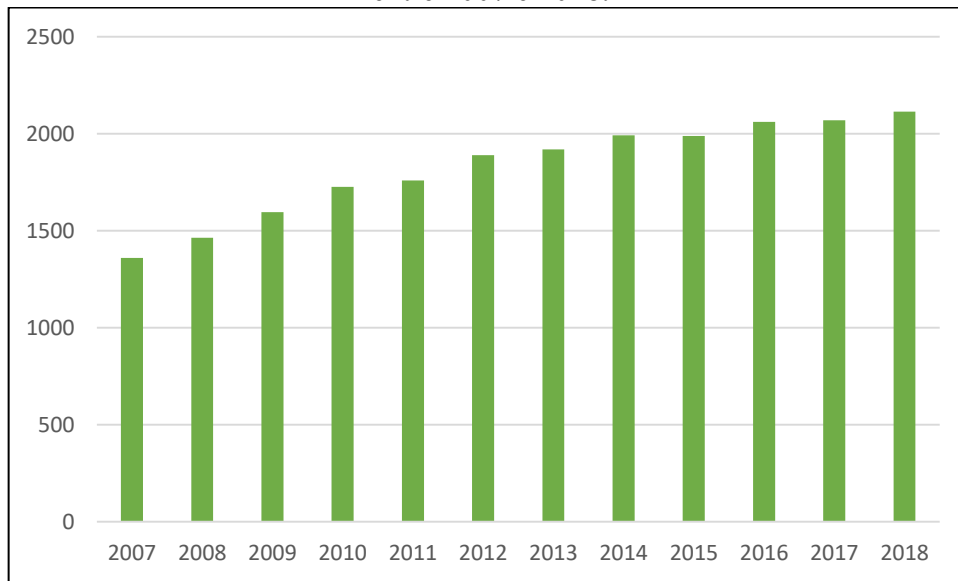
O processo de estimação da equação (9), segundo Kumbhakar, Lien e Hardaker (2014), consiste em primeiramente encontrar as estimativas para o vetor  $\hat{\beta}$  para que sejam obtidos os valores preditos do efeito aleatório da firma ( $\hat{\alpha}_j$ ) e o erro estatístico  $\hat{\varepsilon}_{jt}$  através de uma regressão para dados em painel com efeito aleatório. A seguir, os valores de  $\hat{\varepsilon}_{jt}$  são

utilizados para estimar a ineficiência variável no tempo ( $u_{jt}$ ), utilizando o modelo de fronteira estocástica. O passo seguinte é encontrar a ineficiência persistente  $\eta_j$  a partir do efeito aleatório da firma estimado. Ao final, pelo produto da eficiência técnica persistente e da eficiência variável, é encontrada a eficiência técnica total.

O período de análise para a construção da fronteira compreende o intervalo entre 2007-2018. Isso se deve ao fato de que, no ano de 2007, a Lei 11.445/2007 entrou em vigor, marcando uma nova trajetória de acontecimentos no setor de saneamento. Todas as variáveis utilizadas para a construção da fronteira serão retiradas do Sistema Nacional de Informações sobre Saneamento (SNIS). A Figura 2 exibe o número de prestadores que passaram a reportar suas informações ao longo do período. Apesar de se observar um crescimento do número de empresas no período estudado, não foi possível garantir a participação de todas as firmas devido à ausência de informação em alguns anos, resultando em um desbalanceamento nos dados em painel.

Além disso, Quadro 1 apresenta, em detalhes, as variáveis utilizadas para a construção da fronteira. Todas as variáveis monetárias foram deflacionadas segundo o Índice de Preços por Atacado (IPA). Este índice foi escolhido por ser o principal indicador de reajuste de preços da indústria em geral.

Figura 2. Total de operadoras de serviços de água e esgoto que reportaram suas informações entre 2007 e 2018.



Fonte: Elaboração a partir de dados do SNIS (2020).

Quadro 1. Variáveis que compõe a fronteira de custo estocástica.

Variável	Definição	Unidade
Custo variável (CV)	Compreende o somatório dos custos com pessoal (mão-de-obra própria), produtos químicos, energia elétrica, serviços de terceiros, água importada e despesas fiscais.	R\$/ano
Preço do Trabalho (P <sub>L</sub> )	Consiste na divisão das despesas com pessoal próprio pelo número total de empregados próprios.	R\$/ano
Capital (K)	Medido pela extensão total da malha de distribuição de água e de esgoto coletado.	Km/ano
Produção dos serviços (Q)	Somatório do volume total de água submetida a tratamento com o volume total do esgotamento sanitário coletado.	1.000 m <sup>3</sup> /ano
Custo Médio dos Materiais (CMM)	Calculado dividindo as despesas com produtos químicos para o tratamento pela soma do volume de água produzido.	R\$/ano/1.000m <sup>3</sup>

Fonte: Elaboração própria.

Conforme a Figura 2, pela base de dados SNIS, a quantidade de operadores que reportam suas informações aumenta ao longo do tempo. O problema é que, se for considerado todo o período do estudo (2007-2018), das 2115 firmas que reportaram seus dados, apenas 1360 empresas enviaram relatórios completos com seus dados econômicos e financeiros. Então, caso um modelo seja estimado apenas com essas firmas, uma parcela grande de

municípios ficaria fora das estimativas. Se essas cidades tivessem a pior infraestrutura instalada, os resultados poderiam ser viesados, como destacado por Mattos *et al.* (2019).

Ao observar a Figura 1, não é possível diagnosticar se a perda de informação passada foi relevante para gerar problemas nas estimativas da fronteira estocástica. Se a ausência de informação em períodos anteriores apresentar algum comportamento sistemático, então isso pode gerar um atrito nos dados do painel, provocando estimativas inconsistentes.

Para verificar se há atrito no painel e se está relacionado as variáveis do Quadro 1, em períodos anteriores, aplica-se um teste de atrito, conforme sugere a literatura<sup>5</sup> especializada a fim de verificar se existe viés nos resultados. A descrição do teste e as tabelas com os resultados estão no Apêndice A do trabalho. Os resultados para o teste apontam que o atrito não é aleatório e que está relacionado as variáveis da fronteira em períodos anteriores.

Para corrigir o viés de seleção causado pela perda de informação, Mattos *et al.* (2019) propuseram aplicar um modelo de tomada de decisão relacionado ao relato de informações sobre as variáveis do SNIS ao longo do tempo. Neste modelo, a probabilidade de uma empresa reportar dados no SNIS, no período  $t$ , é uma função dos indicadores de saneamento de períodos passados e variáveis explicativas defasadas. Assim, considera-se  $s_{jt}$  uma variável que assume valor igual 1 quando a operadora  $j$  reporta informação no tempo  $t$ . Levando em conta que esses dados são divulgados dependendo dos indicadores de saneamento no passado mais um termo não observável, a função que indica se as firmas incluem seus dados na base do SNIS tem o seguinte aspecto:

$$s_{jt} = \begin{cases} 1 & \text{se } \varpi_{jt} > -\varphi_0 - \varphi_{jt}\omega_{jt} \\ 0 & \text{se } \varpi_{jt} \leq -\varphi_0 - \varphi_{jt}\omega_{jt} \end{cases} \quad (10)$$

em que  $\omega_{jt}$  representam as variáveis explicativas usadas para a construção da fronteira e suas respectivas defasagens<sup>6</sup> até  $t - 2$  e  $\varpi_{jt}$  é o termo não observável. Para aplicar a correção considerando  $s_{jt}$ , inicialmente, deve ser estimado um modelo *probit* que relaciona o indicador da firma  $j$  reportar seus dados para o município no período  $t$  ( $s_{jt}$ ) tendo como vetor  $\omega_{jt}$ .

---

<sup>5</sup> Nijman e Verbeek (1992).

<sup>6</sup> Para defasagens superiores a dois períodos as estimações dos coeficientes podem não levar a convergência dos verdadeiros parâmetros. Maiores detalhes podem ser consultados em Nijman e Verbeek (1992).

Assim, considerando a forma log-linearizada de uma função Cobb-Douglas<sup>7</sup>, a equação (11) representa a estimação do modelo de fronteira estocástica:

$$\ln CV_{ijt} = X_{ijt}\beta + \alpha_{UF} + \phi_{ANO} + \varepsilon_{ijt} \quad (11)$$

em que  $X_{ijt} = [\ln P_{ijt}, \ln K_{ijt}, \ln Q_{ijt}, \ln CMM_{ijt}, \ln P_{ij,t-1}, \ln K_{ij,t-1}, \ln Q_{ij,t-1}, \ln CMM_{ij,t-1}, \ln P_{ij,t-2}, \ln K_{ij,t-2}, \ln Q_{ij,t-2}, \ln CMM_{ij,t-2}]$ . Além das variáveis discriminadas no Quadro 1, a equação (11) também apresenta uma *dummy* para estados (*UF*), que busca controlar as diferentes características existentes entre os estados brasileiros, e outra *dummy* para o tempo (*ANO*) que visa controlar o efeito do tempo nos custos. Convém destacar que, para preservar a homogeneidade de grau 1 da função de custo nos preços dos insumos, o custo variável total e o preço médio do trabalho foram divididos pelo preço médio da energia elétrica, que foi obtido dividindo as despesas com energia elétrica (R\$/ano) pelo consumo total de energia nas operações dos sistemas de água e esgoto.

Se a empresa não reporta as informações em  $t - 1$ , então, por suposição, a probabilidade de reportar os dados em  $t$  não é afetada. Assim, a esperança condicional do erro da equação (11) dependerá do erro nas estimações do modelo *probit*, ou seja, na probabilidade da firma relatar informações em  $t$  e  $t - 1$ , não dependendo da interação entre esses erros:

$$E(\varepsilon_{ijt} | x_{ijt}, \omega_{ijt}, s_{ijt} = 1, s_{ijt-1} = 1) = \rho_1 \varpi_{ijt} + \rho_2 \varpi_{ij,t-1} \quad (12)$$

Se a equação (12) for aplicada em (11), tomando o valor esperado, é possível encontrar a expectativa condicional dos custos variáveis:

$$E(CV_{ijt} | x_{ijt}, \omega_{it}, s_{it} = 1, s_{it-1} = 1) = \beta_0 + \beta_1 X_{ijt} + \alpha_{UF} + \phi_t + \rho_1 E(\varpi_{it} | \varpi_{it} > -\varphi_0 - \varphi_{it} \omega_{it}) + \rho_2 E(\varpi_{i,t-1} | \varpi_{i,t-1} > -\varphi_0 - \varphi_{it} \omega_{i,t-1}) \quad (13)$$

Utilizando as propriedades assintóticas para a distribuição normal, pode-se assegurar que a expectativa condicional converge para o verdadeiro parâmetro. Assim, reescrevendo a equação (13) sob a forma linear:

---

<sup>7</sup> Essa forma funcional utilizada para representar a tecnologia de produção foi escolhida em virtude de diversos trabalhos que avaliaram a eficiência do setor saneamento como Da Silva e Souza, Faria e Moreira (2007), Ohira e Shirota (2005), Sabbioni (2008), Ferro *et al.* (2014), Ferro e Mercadier (2016) e Pontes (2019) adotarem esta forma, encontrando resultados satisfatórios nas estimações.

$$\ln CV_{ijt} = \beta_0 + \beta X_{ijt} + \alpha_{UF} + \phi_t + \rho_1 \Lambda(\omega_{it}) + \rho_2 \Lambda(\omega_{i,t-1}) + \xi_{ijt} \quad (14)$$

em que  $\Lambda(\omega_{it}) = \frac{\psi(\omega_{it})}{\Phi(\omega_{it})}$  e  $\Lambda(\omega_{i,t-1}) = \frac{\psi(\omega_{i,t-1})}{\Phi(\omega_{i,t-1})}$  são as funções que representam o controle do viés causado pelas firmas que não reportaram informações ao SNIS. Os termos  $\psi(\cdot)$  e  $\Phi(\cdot)$  são, respectivamente, a Função de Densidade de Probabilidade e a Função de Densidade Acumulada, obtidas pela estimação da probabilidade das operadoras reportarem as informações dado o vetor  $\omega$ , respectivamente. Assim, ao final desta etapa,  $\Lambda(\omega_{it})$  e  $\Lambda(\omega_{i,t-1})$  são estimados e utilizados na segunda etapa que consiste em estimar um modelo de regressão linear que relaciona a variação dos custos variáveis com as variações dos regressores, sendo corrigidas pelos termos de correção encontrados.

Após obter os  $\hat{\beta}$  pela estimação da fronteira de custo, pode-se realizar o procedimento Kumbhakar, Lien e Hardaker (2014) para encontrar a ineficiência persistente e variável. Com os valores preditos de  $\varepsilon_{ijt}$ , pode-se estimar a ineficiência variável no tempo pela equação (15):

$$\lambda_{ijt} = \gamma_0 + \gamma_1 \text{fin}_{ijt} + \gamma_3 \text{Leg}_{ijt} + \gamma_4 \text{Reg}_{ijt} + \gamma_5 \text{Leg} * \text{Reg} + w_{ijt} \quad (15)$$

em que *fin* representa um indicador de desempenho econômico-financeiro<sup>8</sup> para a firma *j* no município *i* e tempo *t*. Ambas as variáveis podem variar entre empresas e no tempo, podendo causar uma ineficiência transitória nos serviços prestados. Já *Leg* e *Reg* representam *dummies* para o período que o Plansab entrou em vigor (a partir de 2014) e agência reguladoras, respectivamente. A interação entre eles busca capturar o efeito desses instrumentos, em conjunto, na ineficiência dos serviços prestados.

O procedimento final consiste em estimar a ineficiência persistente. Nesta etapa, foi considerada apenas aquelas variáveis que são invariantes no tempo. Assim,  $m_{ij}$  pode ser estimado pela equação (16):

$$m_{ij} = \delta_0 + \delta_1 \text{emp}_{pub} + \delta_2 \text{emp}_{mista} + \delta_3 \text{aut} + \delta_4 \text{adm}_{direta} + \delta_5 \text{regional} + w_{ij} \quad (16)$$

---

<sup>8</sup> É calculado através da divisão da soma das receitas operacionais diretas com água e esgoto pelas despesas totais para a realização dos serviços.

onde as *dummies*  $adm_{direta}$ ,  $emp_{pub}$  e  $emp_{mista}$  representam as empresas que são administradas pelas prefeituras, totalmente públicas e mistas que possuem participações da iniciativa privada, respectivamente. A variável *aut* apresenta valor 1 caso as operadoras sejam autarquias criadas por lei e que, apesar de serem entidades públicas, possuem autonomia para gerenciar os serviços que estão sob sua responsabilidade. Já a *dummy* regional recebe valor 1 caso a prestadora de serviço seja estritamente regional. Vários autores na literatura recomendam essa distinção entre empresas regionais<sup>9</sup> e locais, além da diferenciação entre o público e o privado, para que o efeito de uma empresa pública seja comparado ao de uma empresa privada.

Ao estimar cada ineficiência, pode-se obter os valores da eficiência técnica transitória e permanente utilizando o procedimento de Battese e Coelli (1988) que consiste em encontrar as eficiências através do residual encontrado nas estimações das ineficiências. Para eficiência transitória aplica-se o método utilizando:

$$EV_{ijt} = \exp(-\lambda_{ijt}) \quad (17)$$

e, modo semelhante, para a eficiência persistente:

$$EP_{ij} = \exp(-m_{ij}) \quad (18)$$

Ao final dessas etapas, pode ser encontrada a eficiência técnica total a partir do produto entre as eficiências transitória e variável. Conforme percebido, a separação das ineficiências ocorre devido a consequência da especificação do método e os componentes transitório e permanente surgem como resultado das estimações. A principal vantagem em utilizar a fronteira de custo estocástica com essa especificação é que ela permite capturar as diversas mudanças que ocorreram neste setor, como as alterações nas leis e no marco regulatório.

---

<sup>9</sup> No contexto do saneamento, uma empresa regional é caracterizada por ser uma empresa de economia mista que presta seus serviços em vários municípios do mesmo estado. Anteriormente, essas empresas eram denominadas Companhias Estaduais de Saneamento Básico (CESB). Por outro lado, as empresas locais são as companhias que possuem natureza jurídica variada, podendo ser autarquias, empresas privadas, entre outras, que prestam serviços exclusivamente para um único município.

#### 4.2 Mensurando o impacto da ineficiência nas condições de saúde infantil

A partir dos resultados obtidos na análise de eficiência dos serviços de saneamento, busca-se relacionar esses resultados com as DRSI. Para tanto, estima-se um modelo em que a variável dependente representa as hospitalizações causadas pelas DRSI em crianças com até 5 anos de vida. Essas doenças podem ser classificadas de acordo com sua forma de contágio e grupo específico de morbidades. O Quadro<sup>10</sup> 2 apresenta um resumo dessas informações.

Quadro 2 – DRSI segundo sua categoria de contágio e grupo de doenças.

Forma de contágio	Grupo	Doenças
Feco-Oral	Infecções diarreicas	Cólera, Salmonela, Amebíase, Isosporíase e outras infecções intestinais.
	Febre tifoide	Tifo e Paratifoide.
	Outros	Hepatite A, Poliomielite, Leptospirose, Ascariíase e Tricuriase.
Inseto vetor	Procriação na água	Filariose, Malária, Chagas, Dengue, Zika, Chikungunya, Febre Amarela e Leishmaniose.
Água contaminada	Penetração da pele	Esquistossomose
	Ingestão	Cisticercose, Teníase e outras infecções causadas por helmintos.

Fonte: Cairncross e Feachem (1990) e WHO (2009). Adaptado.

Assim, definindo DRSI como a taxa de internações por essas doenças (Quadro 2) por mil habitantes no município  $i$  para o período  $t$ , tem-se a equação (19):

$$DRSI_i = f(EV, EP, X_i) \quad (19)$$

É importante destacar que muitas informações sobre o saneamento não possuem continuidade ao longo do tempo, fazendo com que a amostra de municípios ao longo do tempo varie muito. Assim o modelo estimado corresponde a dois períodos de análise, a saber, 2007 e 2018. Isso foi necessário para manter a compatibilidade dos bancos de dados obtidos. A escolha para o ano 2007 se deve ao fato que foi, neste ano, que Lei Federal 11.445/2007 foi

<sup>10</sup> Códigos de doenças ligadas ao saneamento inadequado, segundo a Classificação Internacional de Doenças (CID-10) são: i) grupo A: 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 23, 27, 28, 32, 36, 37, 46, 48.4, 48.8, 71, 75, 80, 90, 92, 95, 97; ii) grupo B: 02, 05, 08, 15, 25, 26, 33, 50, 51, 52, 53, 55, 57, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 77, 78, 79, 80, 81, 82, 83, 86, 95, 96.

sancionada, dando início a grandes mudanças devido ao novo marco legal do saneamento. As informações sobre saneamento são extraídas do SNIS, enquanto as informações sobre os casos de DRSI podem ser obtidas através dos registros do Sistema de Informação Hospitalar (SIH), mantido pelo Departamento de Informática do SUS (DATASUS).

A equação (19) mostra que a taxa de internações por DRSI pode ser modelada como uma função da eficiência variável (*EV*) e persistente (*EP*) dos serviços de saneamento, conforme explicitado pelo modelo de Kumbhakar, Lien e Hardaker (2014), mais uma matriz de variáveis de características de controle, denominado  $X_i$ . A composição da equação (19) busca materializar a estrutura teórica da equação (7), trazendo, além da capacidade na entrega eficiente desses serviços, características ligadas a renda (IFDM-ER), infraestrutura de saúde (Cobertura ESF) e ao estoque de saúde infantil (Cobertura Vacinal). O Quadro 3 apresenta uma descrição detalhada das variáveis explicativas utilizadas.

Quadro 3. Variáveis explicativas para construção do modelo que relaciona eficiência e saúde infantil.

Variáveis	Definição	Sinais esperados
EV	Medida de eficiência variável das firmas de saneamento em cada município, obtidas segundo o método da fronteira estocástica.	–
EP	Medida de eficiência persistente das firmas de saneamento em cada município, obtidas segundo o método da fronteira estocástica.	–
Cobertura Vacinal	Percentual de cobertura vacinal do calendário da Sociedade Brasileira de Imunização. As vacinas recomendadas para todas as crianças até o quinto ano de vida são: BCG, Hepatite A e B, DTP, Poliomielite, neumocócica, Meningocócica B e Tríplice Viral, como <i>proxy</i> do status de saúde das crianças.	–
Cobertura ESF	Percentual da área total atendidas pelos serviços oferecidos nos postos de Estratégia de Saúde da Família (ESF) como uma <i>proxy</i> de infraestrutura de saúde.	–
IFDM-ER	Índice Firjan de Desenvolvimento Municipal, dimensão Emprego e Renda como uma <i>proxy</i> de renda do município.	–

Fonte: Elaboração própria.

Como é possível notar, a equação (19) não considera o status de saúde infantil em  $t - 1$  por dois motivos. O primeiro é que o modelo teórico considera como unidade de observação o indivíduo ao passo que no modelo empírico foi possível observar apenas o agregado de indivíduos por município. O segundo é que o objetivo do modelo não é captar efeitos dinâmicos e, por considerar períodos de estimação distantes (como explicitado adiante),

considerar uma informação de saúde passada muito anterior ao período de estimação não faria sentido para os dados de seção cruzada.

#### 4.2.1 Análise Espacial

A estimação da equação (19) pode ser feita através de uma abordagem não espacial, como outros trabalhos (MATTOS *et al.*, 2019; SAIANI e AZEVEDO, 2018; CRUZ e RAMOS, 2012) já realizaram. Entretanto, ignorar os possíveis efeitos espaciais pode causar problemas de inconsistência e ineficiência nas estimações realizadas, conforme já destacado por Silva *et al.* (2020) e Scriptori (2016). Dado que a análise econométrica é sobre unidades geográficas, muito provavelmente existirão interações socioeconômicas e interdependência espacial entre as regiões, tornando a abordagem espacial mais adequada uma vez que as correlações espaciais entre as observações  $i$  e  $j$  estarão presente na modelagem.

Para diagnosticar a existência de efeitos espaciais, é essencial que se realize uma Análise Exploratória de Dados Espaciais (AEDE). Essa etapa é importante, pois, além da autocorrelação espacial, qualquer outra heterogeneidade pode ser identificada. Segundo Anselin (1988), a autocorrelação espacial ocorre quando o valor de uma variável, observada na região  $i$  afeta o resultado dessa mesma variável, porém em uma região  $j$ . Por outro lado, a heterogeneidade espacial ocorre quando os dados das unidades espaciais são muito distintos e explicam o mesmo fenômeno.

O grau de vizinhança é um ponto fundamental para se realizar uma análise de autocorrelação espacial. Essa vizinhança pode ser estabelecida de várias formas e servirá para a construção de uma estrutura, chamada matriz de pesos espaciais, que descreve a dependência entre as unidades analisadas. Essa matriz é formada com base na contiguidade das observações e pode ser definida segundo sua vizinhança, distância geográfica ou por fatores socioeconômicos. Dessa forma, a lógica por trás de sua construção segue em atribuir um peso  $w_{ij} = 1$  se as regiões  $i$  e  $j$  são contíguas e 0 caso contrário (ALMEIDA, 2012).

Existem algumas matrizes de pesos espaciais, como a convenção de rainha e torre, que são mais comuns de serem utilizadas. A matriz de pesos do tipo rainha considera, além das fronteiras com extensão diferente de zero, os vértices existentes como próximos. Já a matriz do tipo torre leva em consideração apenas as fronteiras físicas entre as regiões observadas (ALMEIDA, 2012).

Quando a distância geográfica é utilizada para definir os pesos espaciais, a matriz de contiguidade será formada por uma quantidade  $k$  de vizinhos mais próximos, de tal maneira

que para ser considerado dentro dessa vizinhança uma distância  $d$  deve ser estabelecida. Então, as regiões  $i$  e  $j$  serão consideradas vizinhas se  $d_{ij} \leq d_i(k)$ , onde  $d(k)$  é uma distância de corte. Logo, todos os  $k$  vizinhos que atendam essa condição irão receber um peso  $w_{ij} = 1$  na matriz de contiguidade (ALMEIDA, 2012).

Como percebido, há diversas formas para a construção da matriz de pesos espaciais e, para evitar erros subjetivos na escolha, foi utilizado o teste de autocorrelação espacial de Moran. O coeficiente  $I$  de Moran, para o caso univariado, pode ser calculado sob a forma matricial, como (ALMEIDA, 2012):

$$I = \frac{Z'WZ}{Z'Z} \quad (20)$$

em que  $Z$  representa os valores da variável de interesse (DRSI) padronizada e  $W_Z$  os valores médios dessa mesma variável padronizada, porém nas regiões vizinhas e definidas a partir de uma matriz de pesos espaciais  $W$ . Neste estudo, o coeficiente de Moran testará a hipótese nula de aleatoriedade espacial das DRSI. A rejeição dessa hipótese implica na dependência espacial dos efeitos das DRSI sobre as regiões vizinhas.

O valor calculado pelo coeficiente de Moran varia entre  $[-1; 1]$ . Resultados positivos revelam que há uma similaridade entre os valores das variáveis estudadas, ou seja, autocorrelação positiva significa que se ocorrerem altos valores da variável de interesse em uma região, também ocorrerá o mesmo em suas regiões vizinhas. Por outro lado, um resultado negativo indica uma dissimilaridade entre os valores das regiões estudadas, isto é, para alto valores na variável de interesse em uma região  $i$  suas vizinhas tendem a apresentar baixos valores nessa mesma variável.

É importante destacar que o  $I$  de Moran apresentado analisa uma autocorrelação global, ou seja, fornece padrões de correlações que o conjunto de dados como um todo podem possuir, se agrupando sistematicamente ou apresentando comportamento disperso. Conforme destaca Anselin (1988), a existência de padrões globais de autocorrelação espacial não implica que, a nível local, será observado o mesmo comportamento. Podem ocorrer casos em que a nível global não seja detectado qualquer associação espacial para a variável de interesse e essa ausência de autocorrelação espacial global acaba por provocar uma camuflagem em padrões de comportamento local. Ou ainda pode existir uma situação em que um diagnóstico de autocorrelação global (por exemplo, positivo) oculte a existência de outros padrões locais de autocorrelação.

A partir dessas considerações, torna-se importante também uma análise dos padrões espaciais locais e, na presente pesquisa, *I* de Moran para o caso bivariado será aplicado para realizar tal análise. Também conhecido por ser uma estatística LISA (*Local Indicator Spatial of Association*), o *I* de Moran local busca verificar se o valor de uma variável, em uma determinada região, afeta os valores de outra variável em regiões vizinhas. Neste caso, o *I* de Moran local pode ser expresso como:

$$I_{Z_1Z_2} = z_{1i}W_{Z_2i} \quad (21)$$

em  $Z_1$  é uma variável padronizada de interesse que, para o presente trabalho, são os níveis de eficiência dos serviços de saneamento;  $Z_2$  representa outra variável padronizada defasada na região vizinha, aqui representada pelas DRSI nessas localidades. Por fim,  $W_{Z_2}$  é uma matriz de pesos espaciais. Como se trata de um índice local, somente os vizinhos de  $i$ , definidos pela matriz de pesos espaciais, são considerados no cálculo.

Idealmente, para uma análise espacial local revelar padrões de associação nas diferentes regiões é necessário que se calcule a estatística de Moran ano a ano<sup>11</sup> e, para cada período calculado, seja apresentado um mapa de diagnóstico onde pode ser visualizado quais tipos de autocorrelação espacial existem no conjunto de dados.

Considerando a possibilidade da existência dos efeitos espaciais, a sua manifestação pode ocorrer de diversas formas. Para o presente estudo, um primeiro efeito espacial esperado ocorre quando uma observação  $j$  sofre um choque na variável dependente e este transborda efeitos para um vizinho  $i$ . No contexto do saneamento, isso significa que alterações na taxa de internações por DRSI de um local pode afetar as regiões vizinhas. Observando que diversas DRSI possuem natureza infectoparasitária e muitos agentes transmissores são insetos que podem facilmente se deslocar entre várias regiões, principalmente entre as regiões fronteiriças, é razoável esperar que tal efeito seja capturado na estimação. Assim, para capturar esse efeito na estimação, pode-se estimar um modelo de defasagem espacial do tipo SAR, conforme a equação (22):

$$DRSI_i = \rho WDRSI_j + \alpha_1 EV + \alpha_2 EP + X_i\beta + \varepsilon_i; \quad |\rho| < 1 \quad (22)$$

---

<sup>11</sup> Infelizmente, devido a significativa perda de informação na amostra, não foi possível realizar uma análise comparativa ano a ano que garantisse a homogeneidade da amostra, isto é, havia mudanças excessivamente bruscas de municípios que entravam e saíam do cálculo do *I* de Moran local devido aos dados missing. Com o intuito de minimizar essa perda de informação, optou-se por fazer uma análise para o ano inicial e final da amostra.

Na equação (22),  $\rho$  é o parâmetro autorregressivo espacial que busca mensurar a influência média que os vizinhos têm nas observações do vetor DRSI. O componente  $WDRSI_j$  é o termo de defasagem espacial que representa, para cada observação, a taxa de interações por DRSI das observações vizinhas.

Outro efeito espacial esperado é devido a heterogeneidade espacial, isto é, podem ocorrer diferentes respostas nas regiões a partir de um determinado choque<sup>12</sup>. Dado que as condições de saneamento e de saúde, em termos de sua infraestrutura, são heterogêneas nos municípios brasileiros é justificável que exista esse efeito espacial no presente estudo e, como consequência, surgirá na estimação do modelo um termo de erro heterocedástico que carrega essa heterogeneidade espacial. Segundo Anselin (2003), esse tipo de modelo é considerado um modelo de erro autorregressivo espacial (SEM) de alcance global e pode ser estimado conforme a equação (23):

$$DRSI_i = \alpha_1 EV + \alpha_2 EP + X_i\beta + \varepsilon; \quad \varepsilon = \lambda W\varepsilon + \xi \quad (23)$$

Na equação (23),  $|\lambda| < 1$  é o coeficiente autorregressivo espacial para a defasagem do termo de erro e  $\xi$  é o termo de erro não-correlacionado e homocedástico. Com essa especificação, os erros do modelo são associados com as observações através de uma média dos erros nas regiões vizinhas mais um componente de erro aleatório. Dessa forma, a influência sobre a variável dependente não ocorre apenas em função de um determinado choque na região específica, mas também de *spillovers* de choques de regiões vizinhas e vizinhas das vizinhas, até exaurir a ponderação dada pela matriz  $W$ .

Por fim, um terceiro componente espacial que pode surgir é através das variáveis exógenas do modelo. Quando isso ocorre, significa que essas variáveis possuem efeitos que transbordam espacialmente, impactando regiões vizinhas de onde o choque aconteceu. Neste trabalho, é razoável supor que a infraestrutura de saúde e a eficiência dos serviços de saneamento de uma região afete um vizinho. Por exemplo, regiões em que há mais leitos hospitalares naturalmente absorvem pacientes de localidades vizinhas. Se a região recebe indivíduos acometidos por DRSI, então as condições de saneamento dos vizinhos estariam exercendo uma externalidade negativa nessa região ao impactar as interações ali ocorridas. Por outro lado, o inverso também ocorrer, ou seja, se as condições de saneamento e saúde

---

<sup>12</sup> Um exemplo de choque que pode afetar as DRSI em regiões vizinhas são vazamentos nas redes de esgoto, que podem percorrer quilômetros, tendo potencial para afetar vários municípios.

melhoram nos vizinhos, uma determinada região pode se beneficiar desses avanços, pois não mais receberá a demanda por internações referentes as DRSI daquela região vizinha. De acordo com Anselin (1988), esse tipo de dependência espacial não afeta todo o sistema, o que significa que esse transbordamento nas variáveis explicativas é local e pode ser modelado conforme a equação (24):

$$DRSI_i = W\theta_1EV + W\theta_2EP + W\theta_3X_i + \alpha_1EV + \alpha_2EP + X_i\beta + \varepsilon_i \quad (24)$$

O modelo apresentado na equação (24) é conhecido por ser um modelo regressivo cruzado espacial (SLX), onde  $\theta_k$  são os parâmetros dos vetores das variáveis explicativas defasadas espacialmente a serem estimados e, conjuntamente, estão associados a uma matriz de ponderação.

Em relação ao método de estimação, quando a presença de autocorrelação espacial é detectada as estimativas feitas via MQO se tornam viesadas, pois não conseguem capturar os efeitos espaciais. O modelo SLX, por exemplo, pode apresentar problemas de multicolinearidade entre as variáveis explicativas e suas defasagens espaciais. Além disso, o modelo SLX se estimado por MQO pode levar a problemas de viés de variável relevante omitida. Essa omissão, segundo Elhorst (2014), ocorreria pelo fato da matriz de variáveis explicativas e suas defasagens estarem correlacionadas, fazendo com que  $E(\hat{\beta}) \neq \beta$ .

De forma semelhante, Elhorst (2014) aponta alguns problemas que podem ser encontrados quando um modelo do tipo SEM é estimado via MQO. O autor faz destaque a matriz de variância-covariância, que se torna diferente daquela encontrada em um modelo de regressão clássico, tornando as estimativas dos erros-padrão ineficientes.

Em vista desses e outros problemas, Elhorst (2014) propõe que os modelos econométricos espaciais em painel sejam estimados por meio do método da máxima verossimilhança. Considerando um modelo espacial com erro autoregressivo com  $N$  observações, a função de máxima verossimilhança a ser maximizada será:

$$\begin{aligned} \text{Log}L = & -\frac{N}{2}\text{Log}(2\pi\sigma^2) + \text{Log}|I_N - \lambda W| \\ & - \frac{1}{2\sigma^2} \sum_{i=1}^N \left\{ y_i - \lambda \left[ \sum_{i=1}^N w_{ij}y_i \right] - \left( x_i - \lambda \left[ \sum_{j=1}^N w_{ij}x_i \right] \right) \beta \right\}^2 \end{aligned} \quad (25)$$

onde  $I_N$  é uma matriz identidade de dimensão  $n \times n$ . As demais variáveis expostas na equação seguem a mesma especificação dos modelos SAR, SEM e SLX. A equação (25) pode ser resolvida maximizando de tal forma que o estimador  $\beta$  seja encontrado através da solução da condição de primeira ordem:

$$\beta = ([X - \lambda(I_T \otimes W)X]'[X - \lambda(I_T \otimes W)X])^{-1} \times [X - \lambda(I_T \otimes W)X]'[Y - \lambda(I_T \otimes W)Y] \quad (26)$$

em que  $I_T$  e operador  $\otimes$  representam, respectivamente, uma matriz identidade  $t \times t$  e o produto de Kronecker, indicando o produto entre duas matrizes que resultará em uma matriz de bloco.

É importante destacar que a interpretação dos coeficientes estimados pelos modelos espaciais possui diferença na forma como são analisados os coeficientes de uma regressão tradicional<sup>13</sup>. O fato de permitir transbordamentos espaciais, faz com que os “betas” tenham efeitos diretos e indiretos. Os efeitos diretos são os efeitos observados na variável dependente da própria observação  $i$  quando ocorre uma variação na variável explicativa  $x_i$ , conforme já conhecido pela econometria clássica. Já os efeitos indiretos são os transbordamentos das variáveis dependentes e independentes sobre a variável dependente. Por exemplo, a variação de um ponto percentual da eficiência variável de uma cidade  $j$  pode provocar uma variação nas taxas de DRSI no município  $i$ . Ao somar esses dois efeitos, tem-se o efeito total da variável explicativa.

Para detectar se as correlações espaciais incluídas no modelo são significativas, as estatísticas do teste LM-Erro e LM-Lag conjuntamente com os critérios de informação de Akaike (AIC) e bayesiano (BIC) são empregadas para a escolha de qual modelo tem melhor adequação ao conjunto de dados.

O teste LM possui sua versão tradicional e outra mais robusta. Ambas as versões possuem sua formulação matemática e hipóteses similares. A diferença reside em um fator de correção incluído nas versões robustas para contornar problemas de má especificação<sup>14</sup>. A principal modificação para os testes robustos é que é considerada a influência da autocorrelação na variável dependente (ou seja,  $\rho \neq 0$ ), quando se trata do teste LM-Erro, e considerada a influência de um processo de erro espacial autorregressivo (isto é,  $\lambda \neq 0$ ), quando se trata do teste LM-Lag.

<sup>13</sup> O tratamento matricial desses efeitos pode ser encontrado em LeSage e Pace (2009).

<sup>14</sup> Os testes LM-Erro e LM-Lag consideram em suas versões tradicionais o parâmetro de defasagem da variável dependente e dos erros, respectivamente, iguais a zero. Caso tais parâmetros sejam diferentes de zero, os testes tendem a rejeitar suas hipóteses nulas com maior frequência, ocasionando perda de poder no teste, ou seja, mesmo quando a hipótese nula é verdadeira os resultados dos testes levam (erroneamente) a sua rejeição. As principais diferenças, em detalhes, de cada versão dos testes podem ser vistas em Florax, Holmer e Rey (2003)

### **4.3 Medindo o efeito da desigualdade urbano-rural no acesso ao saneamento na saúde infantil.**

O déficit no acesso aos serviços de saneamento pode variar substancialmente entre as regiões do país e ainda segundo as características do próprio estado. Ao considerar essa situação, alguns autores (SAIANI; TONETO JR, 2006; SILVEIRA, 2013; VENSON, JACINTO e SBICCA, 2019) exploraram a temática a fim de diagnosticar os fatores determinantes que aumentam/diminuem a probabilidade de redução dos déficits em saneamento. Porém, além de encontrar os motivos para o déficit, também é igualmente importante considerar a possibilidade do seu efeito nas condições da saúde da população.

Como já mencionado anteriormente, o Brasil é um país de grandes disparidades regionais e locais, o que leva a um cenário de desigualdade em várias áreas sociais, como no caso do saneamento. Desse modo, considerar que os efeitos do déficit de saneamento sejam semelhantes, na média, para um determinado conjunto de observações pode mascarar a existência de efeitos distintos ao longo da distribuição dos dados em termos de direção (positivo ou negativo) e intensidade. Em termos estatísticos, é natural supor a existência de uma heterocedasticidade na distribuição dos dados observados.

Assim, propõe-se estimar o efeito do déficit de acesso aos serviços de saneamento na saúde infantil através da abordagem da regressão quantílica (RQ). Ao analisar os diferentes quantis da variável de interesse, é possível identificar como que esta se modifica em resposta a um conjunto de variáveis explicativas. A principal vantagem em se utilizar esse método é que será possível caracterizar toda a distribuição condicional das hospitalizações infantis, a partir de um conjunto de regressores, fornecendo um olhar mais completo da relação entre as variáveis.

A RQ pode seguir uma abordagem condicional (RQC) ou incondicional (RQI). Dado que o objetivo desta seção é analisar o efeito dos diferenciais de acesso ao saneamento entre os meios urbano e rural, a RQI será utilizada como forma de decompor em cada quantil da distribuição de interesse os efeitos das covariadas sobre os casos de DRSI. Apesar da RQC ampliar as possibilidades de explorar as covariadas ao longo de toda distribuição, Firpo, Fortin e Lemieux (2009) ressaltam algumas limitações que os parâmetros estimados podem ter sobretudo quanto a sua interpretação e generalização. Segundo os autores, os coeficientes condicionados a cada quantil não podem ser generalizados para a toda a distribuição dos quantis pelo fato da distribuição não condicional de  $Y$  não corresponder ao quantil condicional

encontrado pelo valor esperado do quantil<sup>15</sup>. Dessa forma, é necessário conhecer toda a distribuição da variável  $Y$  dado  $X$ . Além disso, o efeito marginal da variável independente sobre a dependente é dado como o efeito de  $X$  sobre a distribuição condicional de  $Y$ , o que diz pouco sobre o possível efeito de  $X$  sobre a distribuição incondicional de  $Y$ .

Para contornar esse problema, Firpo, Fortin e Lemieux (2009) propõem um método de regressão quantílica incondicional (RQI) que deriva do método de decomposição de Oaxaca-Blinder, isto é, ao invés de decompor efeitos apenas para a média da distribuição, o método amplia essa decomposição para qualquer estatística distributiva como quantis, mediana, variância, entre outras, a partir da estimação de uma regressão da função de influência recentrada (*Recentred Influence Function – RIF*).

Para encontrar a *RIF* e aplicar a decomposição quantílica, inicialmente é necessário obter a influência que cada observação individual pode ter sobre o valor da estatística de interesse (quantil). A ferramenta analítica para essa identificação é a função de influência (*IF*), que pode ser representada para o  $p$ -ésimo quantil por (FIRPO, FORTIN e LEMIEUX, 2009):

$$IF(Y; q_p) = \frac{p - I\{Y \leq q_p\}}{f_Y(q_p)} \quad (27)$$

em que  $q_p$  se refere ao quantil  $p$  da distribuição incondicional de  $Y$  (casos de DRSI),  $f_Y(q_p)$  é função de densidade de probabilidade (FDP) de  $Y$  para o quantil  $p$  e  $I\{Y \leq q_p\}$  é uma variável indicadora que delimita a ocorrência de  $Y$  até o limite do quantil, ou seja, assume valor igual a 1 quando  $Y \leq q_p$  e zero caso contrário.

Se a estatística distribucional for adicionada na IF, tem-se a *RIF* que pode ser tratada como a contribuição de uma observação individual para a estatística distributiva avaliada (quantil). Seguindo Firpo, Fortin e Lemieux (2009), a *RIF* para qualquer quantil  $p$  será:

$$RIF(Y, q_p) = q_p + IF(Y, q_p) \quad (28)$$

Considere que existam dois grupos de forma que  $T = 1$  indica os indivíduos da zona rural e  $T = 0$  os indivíduos da zona urbana. Para garantir que a *RIF* seja capaz de decompor os

---

<sup>15</sup> Esse fato decorre da Lei das Expectativas Iteradas (LEI) que, no caso de uma regressão linear clássica, é verificada. Firpo, Fortin e Lemieux (2009) mostram, em detalhes, como fazer valer a LEI para o contexto dos quantis por meio da Regressão Quantílica Incondicional.

efeitos corretamente, é necessário construir uma distribuição contrafactual ( $F_C$ ) de  $Y_0|T = 1$ , ou seja, a distribuição que teria prevalecido sob a estrutura das DRSI na zona urbana, mas com as características observadas e não observadas da zona rural. Para tanto, é necessário garantir os pressupostos de *Ignorability*<sup>16</sup> e *Overlapping Support*<sup>17</sup>. Satisfeitas essas suposições, é possível obter a distribuição contrafactual aplicando o Teorema da Transformação Integral de Probabilidade Inversa para a função de distribuição condicional  $F_{CY}^{-1}(p|X)$ . Em linhas gerais, esse teorema se apropria do cálculo da função de cada quantil que tem sua FDA definida em  $[0,1]$  e aplica métodos computacionais para inverter a função de interesse, obtendo, assim  $F_C$ .

A computação da *RIF* consiste em, primeiramente, estimar o quantil  $\hat{q}_p$  conforme o seu estimador original proposto por Koenker e Basset (1978) e obter a densidade estimada da variável dependente. Em um segundo momento, com as densidades estimadas é possível recuperar a distribuição contrafactual  $F_{CY}^{-1}(p|X)$  e, assim, será encontrada uma *RIF* para cada observação amostral. Segundo Firpo, Fortin e Lemieux (2009), se for considerada a hipótese dependência linear entre a  $\widehat{RIF}$  e as covariadas, é possível estimar o efeito parcial incondicional das variáveis explicativas sobre a dependente a partir de um estimador de MQO, como mostra a equação (29):

$$\hat{\gamma} = (\sum_{i=1}^N X'_{it} X_{it})^{-1} \sum_{i=1}^N \widehat{RIF}(Y_{it}; \hat{q}_p) X_{it} \quad (29)$$

em que  $N$  indica a amostra de cada grupo estudado (zonas urbana e rural) e  $i$  indica cada observação da amostra. A equação (29) mostra que, em linhas gerais, o método proposto por Firpo, Fortin e Lemieux (2009) consiste em substituir a variável de interesse por uma reponderação da influência que cada observação tem sobre a distribuição, isto é, a *RIF*. Com isso, o estimador de MQO pode ser aplicado, considerando a hipótese de dependência linear, e os coeficientes estimados  $\hat{\gamma}$  irão representar o impacto que uma mudança na distribuição das variáveis explicativas tem sobre a distribuição incondicional da variável de interesse.

---

<sup>16</sup> A hipótese de Ignorabilidade diz que se  $(X, T, \varepsilon)$  tem uma distribuição conjunta, então para  $\forall x \in X$  o termo de erro  $\varepsilon$  deve ser independente de  $T$ , dado  $X = x$ . O desenvolvimento algébrico desse pressuposto está em Firpo, Fortin e Lemieux (2009).

<sup>17</sup> A suposição de Sobreposição requer que exista tal sobreposição nas características observáveis entre os grupos, no sentido de que não há valor de  $x$  em  $X$  de modo que ele seja observado apenas no grupo 1 (zona rural).

Uma vantagem em se utilizar a RQI aplicando o RIF-MQO<sup>18</sup> é que seus resultados podem ser usados diretamente para a realização da decomposição quantílica, seguindo a abordagem proposta por Oaxaca-Blinder. Para este trabalho, a decomposição quantílica será aplicada para verificar quais fatores contribuem para os diferenciais de casos de DRSI entre as zonas urbanas e rurais dos municípios brasileiros.

Tradicionalmente, a decomposição de Oaxaca-Blinder é aplicada para a média. Firpo, Fortin e Lemieux (2009) generalizam essa decomposição para qualquer estatística da distribuição. Assim, pode-se representar o diferencial total entre as distribuições como:

$$\Delta^v = v(F_{y1}) - v(F_{y0}) \quad (30)$$

em que  $v(\cdot)$  representa a estatística da distribuição (quantil) dos casos de DRSI para os grupos urbano (0) e rural (1). A decomposição dos autores propõe dividir  $\Delta^v$  em um componente que estará relacionado à diferenças de características observáveis dos municípios e um componente associado a diferenças na estrutura de DRSI entre os dois tipos de grupos. Para que esses dois componentes sejam separados corretamente, será utilizada a distribuição contrafactual ( $F_c$ ) encontrada anteriormente para simular uma distribuição de DRSI, compondo a estrutura de casos de DRSI na zona rural e de características da zona urbana. Ou seja,  $\Delta^v$  pode ser reescrito sob a seguinte forma:

$$\Delta^v = [v(F_{y0}) - v(F_{yc})] + [v(F_{yc}) - v(F_{y1})] \quad (31)$$

O primeiro termo da decomposição entre colchetes da equação (31), é o efeito coeficiente, referindo-se claramente aos efeitos da diferença entre os coeficientes estimados das covariadas para a zona urbana em relação a rural. O segundo termo entre colchetes da equação mostra o efeito composição (também chamado de dotação) que mostra a diferença dos casos de DRSI da zona urbana perante aos casos da zona rural devido a características específicas (dotações) de cada área. Se for aplicado o valor esperado das RIF em cada termo da equação (31), têm-se:

---

<sup>18</sup> Termo usado por Firpo, Fortin e Lemieux (2009) para designar o procedimento de estimações dos coeficientes  $\hat{\rho}$ .

$$\begin{aligned}\Delta^v &= \{E[RIF(Y_0; v_0)|X, T = 0] - E[RIF(Y_1; v_c)|X, T = 0]\} \\ &\quad + \{E[RIF(Y_1; v_c)|X, T = 0] - E[RIF(Y_1; v_1)|X, T = 1]\} \\ \Delta^v &= [X_0\beta_0 - X_c\beta_c] + [X_c\beta_c - X_1\beta_1]\end{aligned}\quad (32)$$

Se a hipótese de linearidade for verdadeira, então a expectativa condicional da estatística contrafactual em que se assume a distribuição de características das zonas urbanas e a estrutura de casos de DRSI da zona rural pode ser estimada de forma que seu resultado seja  $X_c B_c = X_0 \beta_1$ . Com isso pode-se reescrever a equação (32) para que o efeito composição e o efeito coeficiente se tornem mais explícitos:

$$\Delta^v = \underbrace{(\beta_0 - \beta_1)X_0}_{\text{efeito coeficiente}} + \underbrace{(X_0 - X_1)\beta_1}_{\text{efeito composição}} \quad (33)$$

Neste estudo, foi considerada a hipótese de dependência linear para apresentar os principais resultados. Firpo, Fortin e Lemieux (2009) ressaltam que na ausência do pressuposto da linearidade, não se pode concluir que  $X_c B_c = X_0 \beta_1$ , pois haverá chances das estimativas não refletirem os verdadeiros efeitos em função de um possível erro de especificação. Contudo, os autores comentam que isso não anula os procedimentos envolvidos na decomposição. Para contornar esse problema, é necessário que a equação (29) seja computada considerando uma matriz de pesos<sup>19</sup>, que pondera a RIF a partir de um *Propensity Score*, *logit* ou *probit* das funções de distribuições acumuladas dos grupos envolvidos e da distribuição contrafactual. Para a presente pesquisa, este exercício foi feito apenas como robustez do modelo principal e seus resultados encontram-se no apêndice B.

O conjunto de variáveis explicativas escolhidas para compor o estudo está baseado em trabalhos anteriores (SAIANI; TONETO JR, 2006; VENSON, JACINTO e SBICCA, 2019) que buscaram identificar os determinantes dos casos de DRSI nos meios urbano e rural. No geral, a presente pesquisa buscou associar essas variáveis com a saúde infantil. O Quadro 4 apresenta a descrição de todas variáveis utilizadas e seus sinais esperados. Todas as variáveis foram extraídas do site do Departamento de Informática do SUS (DATASUS) e, para a

---

<sup>19</sup> Neste caso, a equação (29) poderia ser reescrita como:  $\hat{\gamma} = (\sum_{i=1}^N \widehat{W}_t X'_{it} X_{it})^{-1} \sum_{i=1}^N \widehat{W}_t RIF(Y_{it}; \widehat{q}_p) X_{it}$  onde o subscrito  $t$  refere-se aos grupos urbano, rural e a distribuição contrafactual. A construção de  $\widehat{W}_t$  consiste em uma série de propriedades e pressupostos estatísticos acerca das distribuições de probabilidade, além de propriedades de Teoria Assintótica que vão além deste trabalho. Felizmente, a demonstração da construção dessa matriz encontra-se em Firpo, Fortin e Lemieux (2018) e em Rios-Avila (2020).

presente análise, utilizou-se um período de cobertura de 18 anos, iniciando em 1998 e finalizando em 2015.

Quadro 4. Variáveis explicativas para o modelo de RQI.

Variável	Definição	Sinal
DRSI	Variável dependente que registra os casos de DRSI por mil habitantes relatados pelas famílias ao Programa de Saúde da Família. Os dados utilizados consideram os casos por desidratação, infecções diarreicas e infecções respiratórias.	
Condição de água insegura	Variável explicativa de interesse que indica o percentual das famílias em relação ao total que obtém acesso a água por fontes seguras inseguras de abastecimento, ou seja, caminhões pipa, poços, fontes e bicas comunitárias. O fornecimento do abastecimento a partir de uma rede geral é, segundo os órgãos internacionais, como a ONU, e nacionais (IBGE), a forma mais segura de consumir água.	+
Condição de esgoto inseguro	Variável explicativa de interesse que indica o percentual de famílias em relação ao total que possuem o esgotamento sanitário a céu aberto ou por meio de fossa rudimentar (sem tratamento séptico). O esgotamento sanitário a partir de uma rede coletora geral é, segundo os órgãos internacionais, como a ONU, e nacionais (IBGE), a forma mais segura de excretar os resíduos domésticos do esgoto e, por isso, deve ser utilizado como padrão de acesso.	+
$DRSI_{t-1}$	Variável de controle que identifica o número de casos diagnosticados de DRSI por mil habitantes em $t - 1$ .	+
Taxa de Analfabetismo	Variável de controle que indica o percentual de analfabetos existentes nas famílias. É uma <i>proxy</i> do nível educacional para os municípios.	+
Plano de Saúde	Variável de controle que indica o percentual de famílias em relação ao total com acesso a algum tipo de plano de saúde ativo. É uma <i>proxy</i> para o nível de renda do município.	-
Moradia	Variável de controle que captura a proporção de casas construídas com todos os materiais (madeira, taipa de pilão, palafitas e papelão prensado), exceto alvenaria. É uma <i>proxy</i> para condições inadequadas de moradias.	+

Fonte: elaboração própria.

Todas as variáveis explicativas do quadro 4 são computadas para ambos os grupos (urbano e rural). Além dessas variáveis, foram incluídas *dummies* para controlar os efeitos específicos de cada estado. É importante destacar que a variável DRSI utilizada nesta parte da pesquisa é diferente daquela usada na análise de eficiência. Isso é devido a sua extração ser de outra base de dados, que também pertence ao DATASUS. Aqui a base utilizada refere-se ao Programa de Saúde da Família (PSF), que realizava visitas domiciliares em áreas urbanas e

rurais mensalmente nos domicílios da sua área de cobertura. Daí a importância em usar tal base, pois esta é a única que contém dados desagregados que consegue capturar as morbidades provenientes da zona rural com certo grau de precisão.

Dentre as possíveis variáveis de condições de água, disponíveis nos dados do DATASUS, optou-se por manter nas regressões somente a variável que refletisse as condições inseguras, ou seja, a variável de água incluída representa o percentual de famílias que são abastecidas por fontes como bicas e poços comunitários, carros pipa, água da chuva e manancial. As demais variáveis de água<sup>20</sup> quando incluídas, causavam problemas de multicolinearidade severa.

De forma semelhante, nas condições de esgoto, optou-se por manter as variáveis que mensurassem o percentual de famílias com esgoto a céu aberto e que utilizam fossa, pois a variável que indicava o percentual atendido com rede coletora é muito superior no meio urbano. Em locais rurais, os prestadores instalam interceptores em um determinado local para recolher o esgoto *in natura* e direcioná-lo para a rede pública, dado que a distância entre as casas na zona urbana pode ser grande o suficiente para inviabilizar, em termos de custos, uma rede coletora por tubulação.

## 5. ANÁLISE DOS RESULTADOS

### 5.1 Eficiência de Custo no Saneamento e Saúde Infantil

#### 5.1.1 Análise Descritiva

Inicialmente, foi realizada uma análise descritiva das variáveis de estudo para um melhor conhecimento do seu comportamento. Na Figura 3 é apresentado o comportamento, ao longo do tempo, das variáveis principais variáveis interesse, isto é, a média da taxa de internações por DRSI e a média dos custos variáveis das firmas de saneamento. Os outros resultados descritivos para as demais variáveis, encontram-se na Tabela 1.

---

<sup>20</sup> Substituir a variável de condição insegura de água por uma fonte segura, como abastecimento via rede pública, provocaria naturalmente um gap entre o meio urbano e rural uma vez que nas zonas rurais é comum que as prestadoras de serviços de água perfurem poços artesianos e instalem um reservatório. Por sua vez, também é comum no meio urbano que localidades mais altas ou mais afastadas utilizem esse meio de abastecimento, podendo não separar adequadamente os efeitos da desigualdade. Assim, uma variável que contabilize apenas abastecimento inseguro é mais adequado para capturar os déficits de acesso a esses serviços.

Figura 3. Evolução dos custos variáveis das empresas de saneamento e interações por DRSI.



Fonte: Elaboração própria a partir de dados da pesquisa.

A Figura 3 mostra a tendência de queda ao longo dos anos das DRSI, embora apresente alguns pontos de pico. É interessante notar que quando os custos variáveis eram mais modestos ocorreram movimentos mais acentuados de alta e que das interações. Ao passar por um período de custos mais elevados, a tendência das interações permaneceu de queda, entretanto, com menores oscilações.

Outro ponto relevante na Figura 3 é que a elevação dos custos ocorre a partir de 2013-2014, período em que o Plansab é aprovado e, entre outros aspectos, a regulação passa a ter maior destaque no setor de saneamento, sugerindo, conforme destacado por Pontes (2019), que as operadoras precisaram aumentar suas despesas correntes, na tentativa de se adequar as novas normas que estavam sendo estabelecidas. Após o esforço inicial, a queda abrupta dos custos pode ser explicada, de certo modo, pela postergação dos prazos a serem cumpridos pelas firmas para atender todas as exigências das leis e normas do saneamento. A seguir, a Tabela 1 apresenta as outras estatísticas de descritivas.

Tabela 1. Principais estatísticas descritivas da análise de eficiência de custo e saúde infantil.

Variáveis	N	Média	Desvio padrão	Mínimo	Máximo	Coefficiente de Variação (%)
Painel A						
ln Custo variável	19.961	14,63	1,70	8,37	22,20	11,68
ln Preço do trabalho	19.418	4,09	1,07	-8,29	11,87	26,21
ln Capital	21.849	4,82	1,34	-1,20	10,64	27,79
ln Custo médio dos materiais	19.418	4,30	1,54	-9,19	17,81	35,89
ln Produção	21.849	6,99	1,92	-3,91	14,11	27,49
Painel B						
DRSI	15.784	73,61	116,68	0,33	4139,94	158,51
IFDM-ER	15.784	0,6051	0,1425	0,1456	0,9523	23,54
Cobertura Vacinal (%)	15.784	96,01	25,73	0	431,44	26,79
Cobertura ESF (%)	15.784	80,21	24,38	0	100	30,39

Fonte: Dados da pesquisa.

No painel A da Tabela 1, é possível notar pelo coeficiente de variação a grande heterogeneidade que as variáveis utilizadas para a construção da fronteira possuem, ainda que tais variáveis estejam em sua forma logaritmizada. Isso é uma característica comum entre as operadoras de saneamento nos municípios brasileiros. Segundo Saiani e Azevedo (2018), a produção, utilização de insumos e os custos possuem grandes diferenças devido a existência de operadoras regionais, locais e microrregionais. Apesar do presente estudo não analisar o efeito do porte da empresa na saúde infantil, a amplitude das variações é explicada, em grande parte, pelo tamanho de cada município onde se encontra as operadoras.

O painel B da Tabela 1 também apresenta heterogeneidade no conjunto de dados, principalmente nas taxas internações por DRSI. Segundo Leivas *et al.* (2015), essa heterogeneidade, referente à saúde infantil, é devida as disparidades regionais e até mesmo locais que existem na infraestrutura de saúde brasileira. Entretanto, é interessante destacar que a cobertura vacinal das crianças, em alguns municípios, superou os 100%. Isso indica que a meta estipulada foi ultrapassada e mais crianças que o “necessário” foram imunizadas contra as principais doenças indicadas pela SBI.

### 5.1.2 Estimação da fronteira estocástica e da ineficiência do serviço de saneamento básico no Brasil entre 2007 e 2018

A seguir, a Tabela 2 apresenta os resultados especificados na equação (11), conforme o modelo de Kumbhakar, Lien e Hardaker (2014). As colunas com números ímpares são as estimações realizadas desconsiderando o viés de seleção, ou seja, não foi considerado no

modelo que informações passadas possam afetar as estimativas dos coeficientes encontrados. Por outro lado, as colunas pares incorporam a correção para o viés de seleção e, desse modo, foram adicionados ao modelo os resultados passados para as variáveis explicativas. A coluna que corresponde ao modelo (I) apresenta os resultados obtidos sem o controle dos efeitos por estados e de ano. Nos modelos (II) e (III) são acrescentadas tais *dummies* uma de cada vez e, finalmente, no modelo (IV) tem-se a estrutura completa e, portanto, objeto de análise deste trabalho, para se analisar os efeitos permanentes e transitórios sobre a ineficiência das empresas.

Tabela 2. Resultados para as estimações da fronteira de custo estocástica (2007 a 2018).

Função de custo	Modelo (I)		Modelo (II)		Modelo (III)		Modelo (IV)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln P <sub>L</sub>	0,3824*** (0,0076)	0,4412 (0,3998)	0,3916*** (0,0077)	0,5843 (0,3991)	0,0114*** (0,0043)	0,7423*** (0,2080)	0,0132*** (0,0043)	0,0981*** (0,0232)
ln K	0,7211*** (0,0129)	3,005*** (0,4753)	0,7317*** (0,0129)	0,5674 *** (0,1806)	0,4085*** (0,0080)	0,0816 (0,2445)	0,4177*** (0,0081)	0,9282*** (0,0386)
ln CMM	0,5101*** (0,0044)	0,4901*** (0,1836)	0,5107*** (0,0043)	0,5366*** (0,1832)	0,0284*** (0,0031)	0,2897*** (0,0959)	0,0287*** (0,0032)	0,1315*** (0,0156)
ln Q	0,1330*** (0,0084)	-1,2633*** (0,3699)	0,1282*** (0,0085)	-1,3761*** (0,3687)	0,1607*** (0,0047)	0,5044*** (0,1903)	0,1637*** (0,0048)	0,1443*** (0,0215)
Dummy UF	Não	Não	Não	Não	Sim	Sim	Sim	Sim
Dummy Ano	Não	Não	Sim	Sim	Não	Não	Sim	Sim
ln P <sub>L</sub> (t - 1)	-	-2,304 (1,8451)	-	-2,7099 (1,8417)	-	2,9420*** (0,9550)	-	0,8569*** (0,3046)
ln K (t - 1)	-	13,5787** (6,6798)	-	15,8827** (6,6626)	-	16,2498*** (3,8233)	-	-0,3072 (0,9912)
ln CMM (t - 1)	-	14,8742** (7,2938)	-	11,0188 (7,2884)	-	-10,2692*** (3,3712)	-	-0,0396*** (0,0095)
ln Q (t - 1)	-	-29,0833*** (3,3381)	-	-37,1542*** (6,5211)	-	-12,8419*** (2,2844)	-	-0,0255*** (0,0096)
ln P <sub>L</sub> (t - 2)	-	-16,3889*** (1,5857)	-	-16,3428*** (1,5828)	-	-10,9433*** (3,6302)	-	-0,5697*** (0,0099)
ln K (t - 2)	-	35,0796*** (6,5377)	-	37,1544*** (6,5211)	-	-4,0863*** (0,8410)	-	-0,4760 (0,6345)
ln CMM (t - 2)	-	-3,4597** (1,3490)	-	-3,4319** (1,3433)	-	3,5790*** (0,7341)	-	-0,6618 (0,6376)
ln Q (t - 2)	-	-31,0684*** (4,3113)	-	-28,0886*** (3,3142)	-	-2,7244 (1,8413)	-	-1,6425** (0,6933)
Constante	-6,4868*** (0,0608)	25,5654*** (5,1732)	5,8698*** (0,6152)	23,5449*** (5,1706)	10,5221*** (0,0461)	29,5681*** (2,8239)	11,5791*** (0,6353)	28,1738*** (1,7163)
<b>Ineficiência Variável</b>								
fin	-0,0010*** (0,0001)	-0,0001*** (0,00001)	-0,0011*** (0,0001)	-0,0001*** (0,00001)	-0,0001*** (5,67e-06)	-0,0001*** (9,07e-06)	-0,0001*** (5,80e-06)	-0,0001*** (8,06e-09)
Leg	0,6036*** (0,0232)	0,4451*** (0,0228)	0,5771*** (0,0239)	0,4255*** (0,0229)	0,0823*** (0,0112)	0,0987*** (0,0112)	0,0803*** (0,0112)	0,0971*** (0,0110)
Reg	0,2331*** (0,1159)	0,1757*** (0,0133)	0,2206*** (0,0115)	0,1689*** (0,0134)	0,1011*** (0,0055)	0,1269*** (0,0065)	0,0991*** (0,0055)	0,1260*** (0,0068)
Leg*Reg	-0,3383*** (0,0245)	-0,2616*** (0,0240)	-0,3102*** (0,0246)	-0,2431*** (0,0241)	-0,1088*** (0,0118)	-0,1231*** (0,0118)	-0,1060*** (0,0183)	-0,1212*** (0,0111)

(continua...).

Tabela 2. Resultados para as estimações da fronteira de custo estocástica (2007 a 2018) (continuação).

	Modelo (I)		Modelo (II)		Modelo (III)		Modelo (IV)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Ineficiência Persistente</b>								
Empresa Privada (base)								
Empresa Pública	0,0995 (0,0758)	0,0278 (0,0827)	-0,0929 (0,0671)	0,1043 (0,0739)	-0,3574*** (0,0813)	-0,4389*** (0,0835)	0,2863 (0,0772)	0,3580 (0,8227)
Empresa Mista	0,5427*** (0,0154)	0,4731*** (0,0176)	0,2206*** (0,0135)	0,2250*** (0,0157)	0,3403*** (0,0169)	0,3251*** (0,0173)	0,1708*** (0,0159)	0,1768*** (0,0164)
Autarquia	0,1519*** (0,0268)	0,1250*** (0,0299)	0,1020*** (0,0234)	0,0724*** (0,02672)	-0,1272*** (0,0294)	-0,1425*** (0,0294)	-0,0321 (0,0275)	-0,0568 (0,0810)
Adm. mun. direta	-0,4153*** (0,0287)	-0,5430*** (0,0317)	-0,5486*** (0,0251)	-0,6473*** (0,0283)	-1,1950*** (0,0314)	-1,1712*** (0,0312)	-1,147*** (0,029)	-1,119*** (0,0298)
Regional	-0,1881*** (0,0271)	-0,1933*** (0,0303)	-0,0049 (0,0237)	-0,0808*** (0,0270)	-0,3620*** (0,0305)	-0,2949*** (0,0305)	-0,2041*** (0,0283)	-0,1696*** (0,0289)
Observações	16.376	16.376	16.376	16.376	16.376	16.376	16.376	16.376
R <sup>2</sup>	0,7040	0,6972	0,8057	0,8033	0,7473	0,7337	0,8200	0,8111

Nota: \* p<0.10; \*\* p<0.05; \*\*\* p<0.01. Erros padrão entre parênteses. Para analisar os coeficientes encontrados para as *dummies* de natureza jurídica, tem-se como base o desempenho da empresa privada.

Fonte: Dados da pesquisa.

Analisando os resultados da função custo do modelo mais completo (IV), evidencia-se que o custo variável (variável explicada) apresenta aumentos conforme os regressores no período  $t$  também aumentam, indicando a monotonicidade da função estimada, condição essencial para uma Cobb-Douglas. Além disso, todas as variáveis foram significativas e as elasticidades de custo do produto encontradas são inferiores a um, apontando para a existência de economias de escala no curto prazo.

Outro fato a ser destacado é que não houve mudanças nos sinais das variáveis no período  $t$  entre o modelo sem correção de viés (coluna 7) e o que possui a correção (coluna 8). Isso revela que, apesar da existência de seletividade na amostra, o seu efeito nos coeficientes da função de custo estimada não é grande o suficiente para provocar mudanças qualitativas e causar interpretações errôneas.

Uma outra condição necessária para garantir que a função estimada de curto prazo seja correspondente a função de custo no longo prazo é a significância do coeficiente relacionado ao capital. Se as empresas estivessem em seu ótimo, então a derivada parcial da função de custo variável em relação ao capital deveria ser negativa, conforme mostram Garcia e Thomas (2001). Entretanto, os resultados apresentados revelam um resultado positivo e significativo para a elasticidade do custo em relação ao capital, ou seja, as operadoras estão fora do seu caminho ótimo.

Em relação às variáveis defasadas do modelo 8, entre aquelas que obtiveram significância, apenas o preço do trabalho em  $t - 1$  apresentou coeficiente positivo o que significa que aumentos dos preços passados afetam (aumentam) os custos futuros com mão de obra. As demais variáveis que foram significativas exibiram coeficientes negativos, evidenciando que quando ocorrem aumentos em períodos passados em tais regressores o custo variável no presente diminui.

Vale destacar a lógica por trás dos efeitos negativos das variáveis defasadas do custo médio dos materiais (CMM) e do volume de água e esgoto coletado (Q). É muito comum que as autarquias, empresas públicas e até as de economia mista, realizem licitações para a compra de materiais de consumo e outros materiais ligados a operação do saneamento e, dentre eles, os produtos químicos. Então, em algumas situações, pode ocorrer uma aquisição desses produtos muito superior ao que seria necessário, seja para aproveitar o preço licitado, seja por questões de definição da própria licitação. Com isso, o volume desses produtos em estoque aumentaria, provocando um efeito negativo nos custos variáveis corrente.

O efeito negativo observado da coleta de esgoto defasada [ $\ln Q (t - 1)$ ] pode estar atrelado, principalmente, ao volume de água produzido. Em municípios onde a água utilizada

para o abastecimento não é proveniente de águas que são transportadas pela gravidade até as ETAs, é muito comum a construção de reservatórios e caixas d'água em pontos estratégicos para garantir o abastecimento. Em períodos de maior precipitação, esses reservatórios são acionados para armazenar o volume adicional que chega para tratamento, o que pode ocasionar uma economia com custos de energia elétrica para ativar bombas, casas de máquinas e outros equipamentos ligados a produção de água.

Ainda na Tabela 2, são apresentados os resultados para a ineficiência variável no tempo. No modelo mais completo (modelo IV), os resultados de todas as variáveis incluídas para captar efeitos variáveis na ineficiência se mostraram significativos. Apesar de significativo, o efeito observado do indicador de desempenho econômico-financeiro (*fin*) é muito pequeno. Isso significa que mesmo que as operadoras de saneamento melhorem suas receitas operacionais o seu poder de redução na ineficiência transitória é relativamente baixo.

Por outro lado, as demais variáveis que representam o ambiente jurídico-regulatório, em que se encontra o saneamento, apresentaram resultados mais expressivos, indicando que o cenário no qual as empresas estão inseridas parece ser mais importante que a operação em si para melhores resultados sobre a ineficiência variável.

Inicialmente, a variável *Leg*, que assume valores iguais 1 a partir do ano 2014 (ano quando foi condicionada a transferência de recursos federais aos municípios somente se houvesse PMSB) e a variável *Reg*, que assume valores iguais a 1 quando o município é regulado por alguma agência, se mostraram incapazes de reduzir a ineficiência variável individualmente, pois seus coeficientes foram positivos e significativos. Isso significa que os municípios que possuem apenas agência reguladora, em média, têm uma ineficiência transitória maior que aqueles que não possuem, muito provavelmente, conforme destaca Borja (2014), pelo fato do baixo desempenho técnico em atender e executar as normas impostas. Além disso, Pontes (2019) destaca que com o marco regulatório em 2007 as agências reguladoras poderiam estar criando novas normativas para serem atendidas e isso, no curto prazo, estaria afetando os custos e aumentando a ineficiência.

O fato da legislação aumentar a ineficiência variável poderia estar ligado ao modo como a mesma é implementada. Em alguns estudos, como em Medeiros e Rodrigues (2019), apesar de positivo, o efeito das políticas de saneamento não foi significativo para impactar a eficiência das operadoras. Por outro lado, Pontes (2019) observou um efeito negativo, porém sem significância estatística para sua amostra. No geral, o presente estudo e os demais citados convergem ao expor a dificuldade que este instrumento possui em gerar externalidades positivas para o setor, pois, em muitas situações, a condicionalidade de recebimento de

recursos federais para o setor poderia está acelerando a criação dos PMSB, mas com pouca coordenação, provocando dificuldades na execução do que foi planejado.

Entretanto, quando há regulação nos municípios nos anos em que a transferência de recursos federais se torna condicional a existência de PMSB os resultados da Tabela 3 indicam que, conjuntamente, essas variáveis são capazes de reduzir a ineficiência de custo das empresas de saneamento. Tais resultados são semelhantes aos encontrados em Pontes (2019) e reforçam a importância de a regulação ser realizada tomando como base o plano existente, para que o PMSB deixe de ser apenas obrigatório e se torne um instrumento relevante para melhorar os serviços de água e esgoto nos municípios brasileiros.

A última parte da Tabela 3 apresenta os resultados referente aos efeitos persistentes sobre a ineficiência de custo. Constatou-se pelos resultados da Tabela 3 que as operadoras que são sociedades mistas, sejam com administração pública ou privada, aumentam a ineficiência se comparadas com a empresa privada, enquanto empresas que estão sob a responsabilidade da administração pública direta (prefeituras) podem reduzir a ineficiência dos serviços, pois seus coeficientes foram negativos e significativos.

As demais variáveis relacionadas a natureza jurídica, apesar de não se mostrarem significativa, mantiveram o sinal esperado dos seus coeficientes. Tais resultados se assemelham com alguns trabalhos (Da Silva e Souza; De Faria; Moreira, 2007; Sabbioni, 2008) e, por outro lado, difere de tantos outros, evidenciando que ainda não há um consenso na literatura que aponte qual tipo de empresa (pública, privada ou mista) é mais eficiente para fornecer os serviços de saneamento.

Os resultados do presente estudo, caminham no sentido contrário aos achados de Seroa da Motta e Moreira (2006) e Ferro e Mercadier (2016) que apontavam que as operadoras privadas tinham melhor desempenho se comparadas com autarquias e empresas públicas. Um dos motivos que pode ter causado essa diferença decorre do fato do modelo adotado. Apesar dos autores utilizarem a fronteira estocástica como forma de estimação na ineficiência dos custos, suas abordagens não consideravam a separação dos efeitos transitórios e persistentes e os efeitos que poderiam ter sobre a ineficiência. Assim, a separação desses efeitos, como no presente estudo, pode trazer maior refinamento para as estimações e isolar a parcela de importância que cada um tem sobre a ineficiência gerada.

Em relação a comparação entre empresas regionais e locais, os resultados da Tabela 3 mostram que as empresas regionais são capazes de entregar os serviços de saneamento de forma mais eficiente se comparado aos operadores locais. Esses resultados corroboram com os trabalhos de Tupper e Resende (2004), Sabbioni (2008) e Pontes (2019) que destacaram a

importância que as economias de escala têm sobre o desempenho dessas companhias ao permitirem que insumos como o custo operacional e os gastos com mão de obra possam ser mitigados quando as firmas atendem diversas localidades.

Após estimar as ineficiências para a média da distribuição e obter os coeficientes da Tabela 2, é possível calcular a eficiência variável e permanente para cada operadora de saneamento e, através do produto entre essas duas eficiências, encontrar uma eficiência total. A Tabela 3 apresenta a média para essas variáveis.

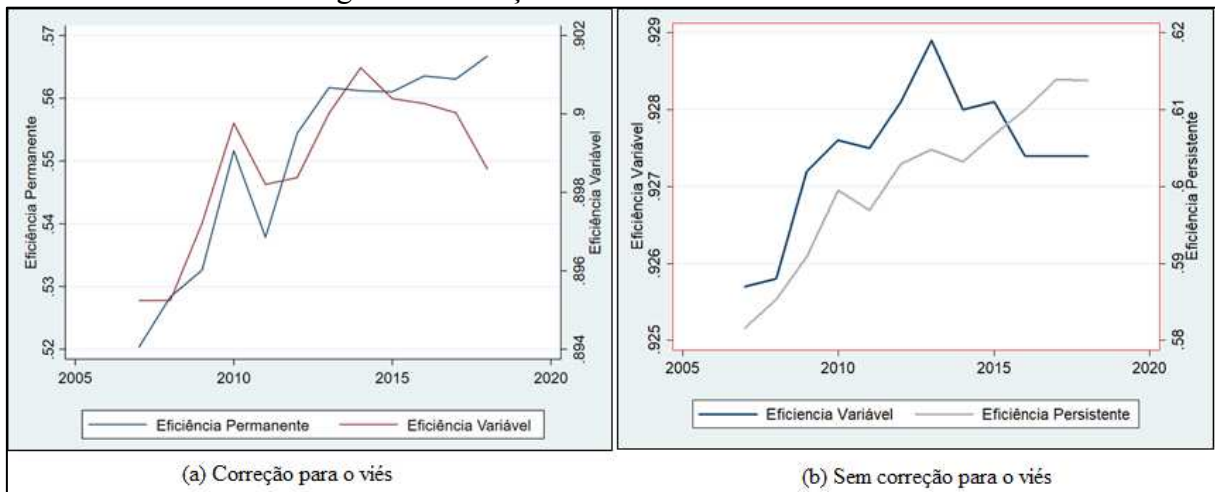
Tabela 3. Média da eficiência das empresas prestadoras do serviço de saneamento.

<b>Eficiência</b>	<b>Média</b>	<b>Desvio Padrão</b>	<b>Mínimo</b>	<b>Máximo</b>
Variável	0,8468	0,0289	0,4982	0,9691
Persistente	0,5534	0,1721	0,1176	0,9362
Total	0,4889	0,1526	0,0888	0,8869

Fonte: Dados da pesquisa.

Pelos resultados da Tabela 3 é possível notar que a eficiência variável é superior a persistente. Isso significa que, no curto prazo, o desempenho das firmas tem sido melhor e, ao longo dos anos, problemas como a gestão das empresas acabam influenciando a eficiência do setor, reduzindo os níveis de eficiência. A seguir, a Figura 4 apresenta a evolução da média de ambas as eficiências ao longo dos anos. Na Figura (a) é apresentado os resultados da eficiência estimados pela coluna 8, com correção de viés de seleção, do modelo (IV) da Tabela 2 e, a Figura (b), os resultados pelo modelo da coluna 7, sem correção de viés de seleção.

Figura 4. Evolução das médias das eficiências.



Fonte: Dados da pesquisa.

Analisando a Figura 4, nota-se que tanto a eficiência variável, quanto a persistente, crescem rapidamente a partir de 2008, ou seja, logo após o marco regulatório do saneamento. Isso significa que o marco regulatório, conjuntamente com os esforços para aumentar os investimentos no setor, podem ter contribuído para esse aumento verificado. Entre 2011 e 2012, verifica-se uma queda abrupta que pode ser explicada pelo fim dos investimentos iniciais do PAC e, logo em seguida, a eficiência retoma sua tendência de crescimento. Entretanto, a partir de 2014, ano em que o PLANSAB entra em vigor, somente a eficiência persistente permanece com uma tendência de crescimento. Tal como já apresentado nos resultados da Tabela 2, a Figura 6 evidencia que impor apenas uma legislação sem considerar mecanismos de regulação apropriados, pode afetar eficiência dos serviços prestados, sobretudo no curto prazo.

Em linhas gerais, os resultados da estimação da fronteira estocástica evidenciaram que as firmas não estão em seu caminho ótimo e que possuem uma eficiência variável superior a persistente, revelando que o desempenho das empresas é melhor no curto prazo. Problemas de gestão e o ambiente jurídico e regulatório são fatores que podem contribuir ou piorar a situação da eficiência no setor de saneamento. Como demonstrado, firmas regionais e a adequação da regulação ao instrumento de planejamento municipal (PMSB) podem trazer melhorias para os serviços de saneamento, desde que aplicados em conjunto.

### 5.1.3 Análise Exploratória de Dados Espaciais (AEDE)

Para realizar a AEDE, foram testadas algumas matrizes de pesos e aquela que alcançasse o melhor ajuste aos dados da amostra seria aplicada em toda análise espacial. A

partir do cálculo do I de Moran para cada ano do período analisado (2007 a 2018), cada tipo de matriz (rainha, torre e distância inversa) foi testada e verificada a significância da estatística de Moran calculada. Ao final do período, uma média simples foi calculada para determinar quais das matrizes havia obtido o maior I de Moran.

Dentre as matrizes espaciais, a do tipo rainha com os 20 vizinhos de primeira ordem mais próximos foi a que apresentou as melhores correlações espaciais. Essa matriz de contiguidade foi construída considerando uma normalização em suas linhas, de tal forma que a soma final de cada linha fosse igual 1. Em outras palavras, vizinhos mais distantes possuíam um peso inferior aos vizinhos mais próximos. A Tabela 4 apresenta os resultados para a autocorrelação global da variável de interesse.

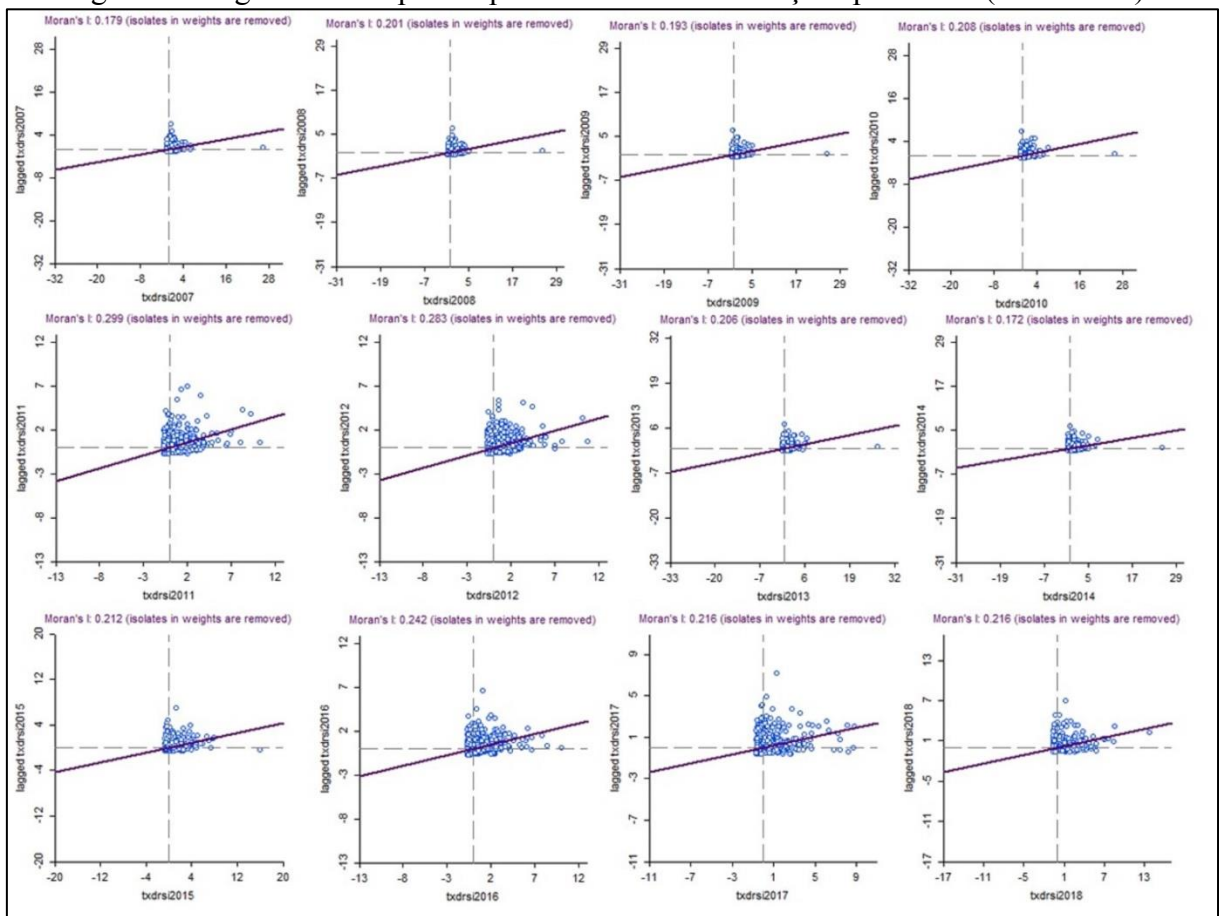
Tabela 4. Autocorrelação espacial global para as taxas de internações por DRSI.

<b>Ano</b>	<b>I de Moran Univariado</b>	<b>P-valor</b>
2007	0,179	0,001
2008	0,201	0,001
2009	0,193	0,001
2010	0,208	0,001
2011	0,299	0,001
2012	0,283	0,001
2013	0,206	0,002
2014	0,172	0,001
2015	0,212	0,004
2016	0,242	0,001
2017	0,216	0,001
2018	0,360	0,003
<b>Média</b>	<b>0,230</b>	

Fonte: Dados da pesquisa.

Os resultados da Tabela 4 indicam que a autocorrelação global da taxa de internações por DRSI foi significativa para todos os anos. A partir destes achados, é possível rejeitar a hipótese de aleatoriedade espacial o que significa que as taxas de DRSI se autocorrelacionam entre os municípios vizinhos. Além disso, o I de Moran foi positivo para todo o período, revelando que municípios que possuem elevados níveis de internação também estão cercados por vizinhos que se encontram na mesma situação. Esses resultados convergem para a literatura recente (VIEIRA, COSTA e JACINTO, 2020; SILVA *et al.*, 2020) que verificaram a autocorrelação espacial global positiva sobretudo para a região do nordeste do Brasil. A seguir, a Figura 5 apresenta o diagrama de dispersão de Moran que ilustra uma correlação do tipo alto-alto dos resultados obtidos na Tabela 4.

Figura 5. Diagrama de dispersão para as taxas de internações por DRSI (2007-2018).



Fonte: Dados da pesquisa.

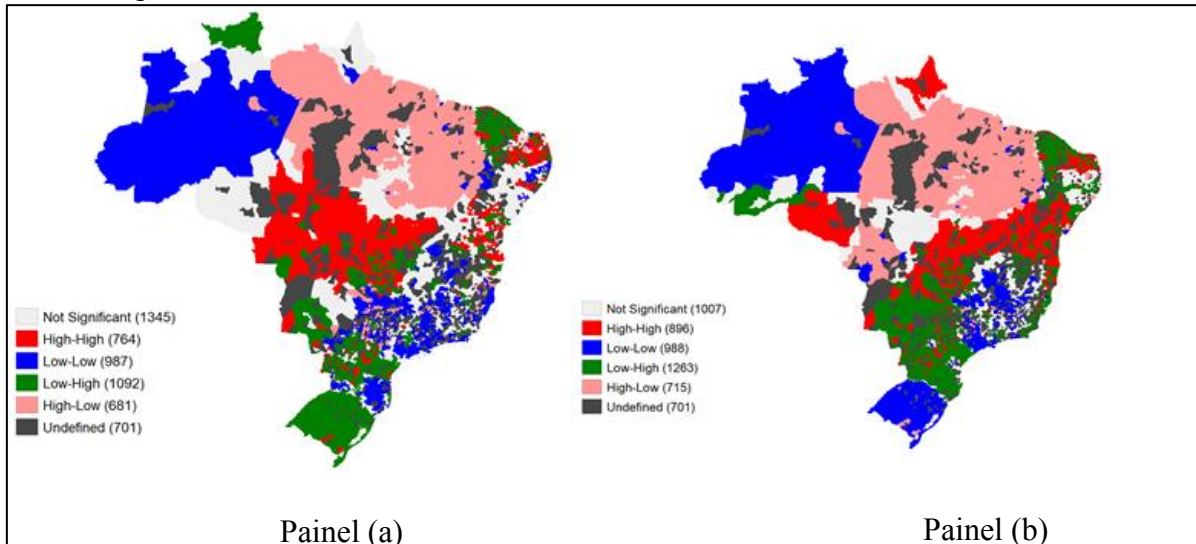
O fato da autocorrelação global ser significativa e persistente ano a ano pode revelar que o esforço em melhorar as condições de saneamento da população, no entanto, pode ainda não ter causado o efeito estrutural desejado para todas as regiões, pois melhorias no saneamento em um dado município poderia evidenciar efeitos positivos em um local vizinho, de maneira que suas taxas de internações por DRSI diminuíssem.

Para afirmar se esse efeito vem ocorrendo ou não, se faz necessário uma análise espacial local que envolva as taxas de internações por DRSI e variáveis ligadas ao saneamento, neste caso, a eficiência. Uma análise local é essencial tendo em vista a grande heterogeneidade existente em todo o território brasileiro. A ferramenta utilizada para essa análise foi o I de Moran local bivariado.

As Figuras 5 e 6 apresentam a distribuição espacial das aglomerações das taxas de internações por DRSI e os níveis de eficiência variável e persistente encontrados nos anos de 2007 e 2018. O objetivo de avaliar o primeiro e o último ano da amostra está em verificar se, com o avanço do tempo, ocorreram mudanças nos padrões de autocorrelação espacial entre

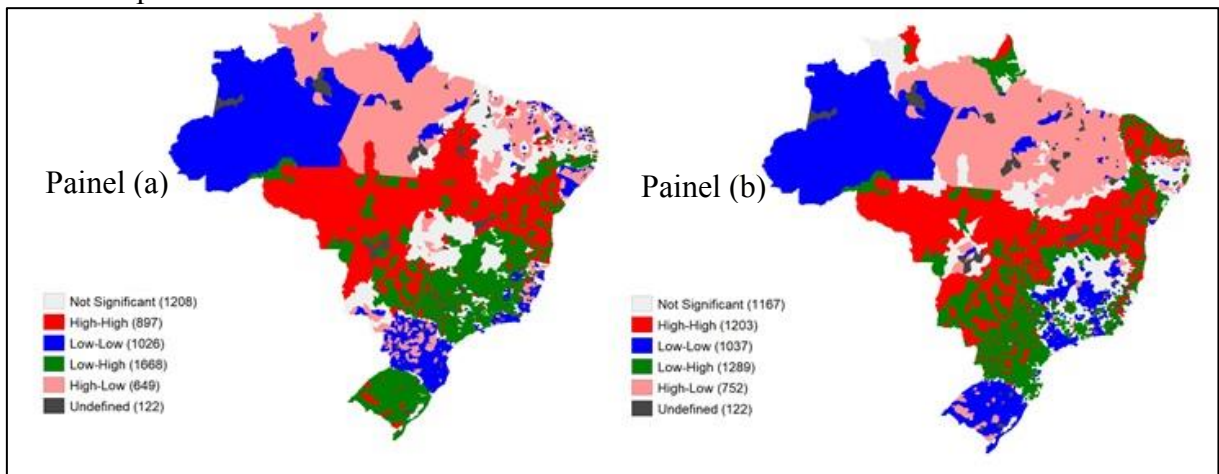
essas duas variáveis. Em ambas as figuras o nível de significância para as autocorrelações é de pelo menos 10% e as regiões na cor preta e branca representam a ausência de dados para os municípios e a não significância da relação analisada, respectivamente.

Figura 5. Distribuição espacial das taxas de internações por DRSI e os níveis de eficiência variável e permanente em 2007.



Fonte: Elaboração própria a partir de dados da pesquisa.

Figura 6. Distribuição espacial das taxas de internações por DRSI e os níveis de eficiência variável e permanente em 2018



Fonte: Elaboração própria a partir de dados da pesquisa.

Os mapas localizados no Painel (a) de ambas as figuras relacionam as DRSI com a eficiência variável, enquanto os mapas do Painel (b) associam as internações com a eficiência permanente. Na Figura 5, é possível notar que o padrão de autocorrelação espacial que as eficiências variável e permanente apresentam com as taxas de internações por DRSI são muito semelhantes, exceto em uma parte da região centro-oeste e sul do Brasil, para o ano de 2007. Isso significa que, na maioria dos municípios, observa-se o mesmo padrão na relação entre

eficiência variável e taxas de internações por DRSI e eficiência permanente e taxas de internações por DRSI.

Em relação aos tipos de autocorrelação espacial, nota-se pelas figuras 5 e 6 que nas regiões norte e nordeste a maioria dos municípios apresentou uma relação do tipo Alto-Baixo (cor rosa), e Baixo-Baixo (cor azul) para 2007 e 2018. Uma correlação do primeiro tipo indica que regiões com elevados valores de taxas de internações, estão rodeadas por municípios com baixos valores de níveis de eficiência variável e permanente. Por outro lado, o segundo tipo revela que uma região com baixos valores nas taxas de internações por DRSI possuem vizinhos com baixos níveis de eficiência variável e permanente.

A autocorrelação Alto-Baixo é um resultado esperado para as regiões norte e nordeste, pois, conforme estudos recentes (VIEIRA, COSTA e JACINTO, 2020; SILVA *et al.*, 2020), são nessas regiões onde municípios com altas taxas de internações por DRSI estão cercados por vizinhos que também possuem elevados níveis de hospitalizações. Os resultados do presente estudo, além de confirmar tais achados, complementam as pesquisas anteriores ao detectar que a autocorrelação espacial das internações por DRSI podem se associar com baixos níveis de eficiência no setor de saneamento.

Por outro lado, a autocorrelação Baixo-Baixo observada em grande parte da região norte contrariou o que era esperado para os municípios dessa localidade. Os baixos níveis de eficiência para o norte já eram aguardados, pois, historicamente, é uma região que carece de maiores níveis de investimentos no setor o que acaba engessando o progresso dos serviços e conseqüentemente a sua eficiência. Entretanto, a consequência esperada da baixa eficiência seria um maior número de internações por DRSI conforme o resultado de vários trabalhos (UHR, SCHMECHEL e UHR, 2016; MATTOS *et al.*, 2019; FONSECA e VASCONCELOS, 2017; FERREIRA *et al.*, 2016) comentam.

As razões que podem explicar, pelo menos em parte, esse resultado controverso, estão ligadas a duas questões. A primeira é que os estudos anteriormente citados, seja em análise espacial ou convencional, relacionam os impactos do saneamento na saúde com *proxies* de eficiência (nível de água tratada, indicadores de qualidade, entre outros). Embora essas variáveis tratem, de alguma forma, de níveis de eficiência, não são indicadores construídos para estimar uma fronteira de eficiência de custo. Assim, apesar da maioria dos municípios possuírem uma relação negativa entre eficiência e internações por DRSI, existem particularidades locais que podem afetar essa associação que, em estudos anteriores, não estavam sendo captadas por não usarem um indicador de eficiência.

O segundo motivo pode estar relacionado a característica da região norte. Alguns estudos evidenciam (CESAR *et al.*, 2005; JUSTINO e ANDRADE, 2020) que muitos municípios da região não possuem uma infraestrutura instalada minimamente adequada para realizar internações hospitalares, conforme protocolo do SUS. Além disso, trabalhos como Medeiros, Costa e Cardoso (2021) relatam que a região norte é a que tem o maior déficit de leitos de internações para cada 100 mil habitantes. Isso provoca, na maioria dos casos menos graves, como é a situação das DRSI, que o tratamento ocorra na residência dos próprios moradores. Por exemplo, uma infecção diarreica pode ser tratada em casa com hidratação e medicamentos orais ao invés de mobilizar recursos e leitos hospitalares, deixando esses leitos para casos mais graves, como fraturas, cirurgias, entre outros. Outro fator que pode motivar o tratamento de DRSI fora de hospitais são as Unidades de Saúde Básicas Fluviais<sup>21</sup> (UBSF) que realizam o acompanhamento de saúde em regiões com maior carência de infraestrutura. A junção desses fatores pode, de algum modo, ter influenciado na relação entre eficiência e as taxas de internações por DRSI.

A autocorrelação espacial do tipo Alto-Alto (na cor vermelho) foi predominante na região centro-oeste e em parte do nordeste (Bahia, Rio Grande do Norte e Ceará), tanto em 2007 quanto em 2018, sendo que a quantidade de municípios nesta classificação aumentou de um ano para outro. Esse tipo de associação mostra que municípios com altas taxas de internações por DRSI são vizinhos de locais que detêm altos níveis de eficiência variável e permanente. Alguns fatores epidemiológicos da região centro-oeste, além do provimento dos serviços de água e esgoto, podem explicar a relação observada entre as variáveis. Uma característica dessa região, segundo Buhler *et al.* (2014), é que ela possui a maior taxa de internação de crianças indígenas, embora não detenha a maior população dessa origem. Alguns fatores culturais como o término precoce da amamentação tornam as crianças dessa região mais suscetíveis a qualquer tipo de doença. Adicionando essa característica com os baixos níveis de coleta e tratamento de esgoto nas regiões metropolitanas, tem-se um cenário que favorece os altos índices de internações por DRSI.

Além disso, é importante destacar que os elevados níveis de eficiência em alguns locais para a região nordeste e centro-oeste podem representar uma prestação de serviço que está no limite de sua operação. Alguns trabalhos (TOMÉ, 2017; BORJA, 2014) destacam a carência de investimentos para o setor de saneamento nessas regiões, sendo a região sudeste a mais privilegiadas no recebimento de recursos. Com isso, as empresas buscam reduzir seus

---

<sup>21</sup> Detalhes sobre o funcionamento do programa podem ser encontrado em <<https://aps.saude.gov.br/ape/ubsf>>

custos não para alcançar maiores lucros, mas, principalmente, para serem capazes de manter a operação dos sistemas de água e esgoto. Em municípios menores, onde a prestação ocorre, na maioria dos casos por autarquias, faltam recursos para ampliação dos serviços e, dessa forma, buscar a eficiência é, na verdade, um processo de cortar despesas sem uma gestão otimizada. Logo, se uma operadora reduz seus gastos e mantém o nível de atendimento, ela alcançou melhores níveis de eficiência de custo. Porém, se a estrutura atual é limitada e atende um percentual baixo de residências e com pouca qualidade no tratamento da água, a eficiência de custo pode não está produzindo o efeito esperado na saúde infantil, pois a falta de recursos impede que uma boa gestão reflita em menores taxas de internações por DRSI.

Há ainda nas Figuras 5 e 6 a autocorrelação espacial do tipo Baixo-Alto (cor verde) que foi predominante nas regiões sul e sudeste e em uma pequena parte do centro-oeste. Esse tipo de autocorrelação espacial indica que os municípios com baixas taxas de internações por DRSI são vizinhos de municípios com altos níveis de eficiência. Esse resultado é interessante, pois confirma que são principalmente as regiões sul e sudeste onde os efeitos espaciais da eficiência possuem influência sobre as taxas de internações por DRSI, indicando que nesses locais os níveis de investimento alcançaram níveis que proporcionam as operadoras a gerenciarem seus custos de forma que a busca pela eficiência traga externalidades positivas para a saúde infantil.

Considerando os resultados apresentados, a AEDE se mostrou uma boa ferramenta para definir os tipos de autocorrelação espacial que podem existir entre as taxas de internações por DRSI e os níveis de eficiência. Entretanto, é importante destacar que outros fatores que podem afetar as hospitalizações por DRSI como, a estrutura de saúde dos municípios, nível de desenvolvimento, entre outras variáveis, não foram observadas nessa análise e, portanto, para detectar o efeito dos níveis de eficiência com maior exatidão é necessária uma análise de regressão, que será executada adiante.

#### **5.1.4 Efeitos da eficiência do serviço de saneamento básico sobre a saúde infantil**

Com as estimativas das eficiências obtidas e verificada a dependência espacial, foram cruzados esses dados com as informações sobre saúde infantil, juntamente com algumas variáveis de controle, com o intuito de analisar o efeito da ineficiência variável e permanente sobre as internações por DRSI. Na Tabela 5, encontra-se as estimativas para os modelos de regressão espacial além dos resultados utilizando o MQO.

Tabela 5. Resultados econométricos e espaciais para o efeito da eficiência sobre as hospitalizações de crianças (2007 e 2018).

Variáveis	Modelo									
	MQO		SEM		SAR		SLX		SAC	
	2007	2018	2007	2018	2007	2018	2007	2018	2007	2018
EP	-5,4614*** (0,9838)	-5,775*** (1,2658)	-7,4762*** (1,0896)	-1,6785** (0,7521)	-8,1528*** (1,0674)	-1,5893** (0,9250)	-6,7769*** (1,1035)	-1,5205* (0,9405)	-7,6010*** (1,0959)	-6,1324*** (1,2461)
EV	-27,7128*** (11,9360)	-1,2047 (2,1118)	-16,0582* (9,4263)	-5,2309*** (1,2801)	-17,3914* (9,4014)	5,7127*** (1,2577)	-19,4970** (9,4005)	-4,3355*** (1,2993)	-16,2829* (9,4280)	-1,4183* (0,8010)
IFDM-ER	-17,1727*** (1,1021)	-3,1074** (1,2447)	-1,7891* (1,0207)	-2,8592** (1,2450)	-1,9389* (1,0981)	-3,0700** (1,2637)	-1,7839* (1,0001)	-3,3194*** (1,2371)	-1,8193* (1,1022)	-3,2795*** (1,2272)
Cobertura Vacinal	-0,5464*** (0,0640)	-0,0007 (0,0061)	-0,0390 (0,0551)	0,0019 (0,0614)	-0,0591 (0,0553)	-0,0047 (0,0613)	-0,0575 (0,0552)	-0,0107 (0,0613)	-0,0427 (0,0553)	-0,0104 (0,0611)
Cobertura ESF	-0,3837*** (0,0620)	-0,1192* (0,0739)	-0,0501* (0,0285)	-0,1051* (0,0544)	-0,0363* (0,0216)	-0,1077* (0,6451)	-0,0608* (0,02567)	0,0784* (0,0479)	-0,04827* (0,0194)	0,1121 (0,0732)
Constante	492,9166*** (210,5258)	52,2948 (113,4687)	160,8017** (89,2456)	57,0432 (113,1251)	167,8465** (88,8412)	54,2117 (112,74)	195,8078** (88,8971)	85,4645 (112,8396)	161,7867** (89,2103)	51,0575 (111,9181)
Erro_defasado	-	-	0,5691*** (0,0941)	0,3566*** (0,0816)	-	-	-	-	0,4791*** (0,1404)	-0,4307** (0,2144)
DRSI_defasada	-	-	-	-	0,3087*** (0,0612)	0,3741*** (0,0679)	-	-	0,0809 (0,0923)	0,6604*** (0,1530)
EP_defasada	-	-	-	-	-	-	-34,6284 (31,0087)	-43,0071 (41,0221)	-	-
EV_defasada	-	-	-	-	-	-	-20,7012*** (3,7210)	-24,3788*** (4,5821)	-	-
Dummy UF	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	4.869	5.448	4.869	5.448	4.869	5.448	4.869	5.448	4.869	5.448
R <sup>2</sup>	0,0843		0,4901	0,2398	0,4818	0,2420	0,4941	0,2466	0,4903	0,2420
AIC	51309,44		51287,82	61065,28	51288,21	61052,17	51278,74	61036,26	51279,05	61050,09
BIC	51517,14		51598,50	61289,79	51508,89	61276,68	51505,91	61267,37	51506,22	61281,19
Teste Wald			36,56***	17,15***	25,41***	30,35***	36,84***	48,12***	36,74***	37,96***
Teste LM-Lag	2840,6***									
Teste LM-Erro	2922,6***									
Teste LM*-Lag	112,4***									
Teste LM*-Erro	200,66***									
Erro										

Fonte: Dados da pesquisa.

Obs: \*\*\* p-valor&lt;0,01; \*\* p-valor&lt;0,05; \* p-valor 0,10.

A estimação por MQO é importante, pois através dos resíduos desse modelo é que os testes LM-Erro e LM-Lag são implementados. Iniciando as análises pela estatística de Wald verifica-se que, em todos os modelos espaciais, o teste de Wald foi significativo, o que demonstra que o modelo restrito (MQO) não é adequado para realizar as estimações. Com isso, todos os modelos espaciais estimados tornam-se candidatos a explicar a relação entre eficiência de custo no saneamento e saúde infantil.

Para determinar qual o modelo é o mais adequado as estatísticas do teste LM-Erro e LM-Lag, em suas versões tradicionais e robustas (LM\*), conjuntamente com os critérios de informação de Akaike (AIC) e bayesiano (BIC), são empregadas para a escolha de qual modelo tem melhor adequação ao conjunto de dados.

Esses resultados também são apresentados na Tabela 5 e, como os resultados observados nos testes LM\*-Erro e LM\*-Lag foram significativos, é evidenciado a existência de autocorrelação nos erros e na variável dependente. Quando ambos os testes são significativos, recomenda-se a estimação de um modelo que contenha ambos os efeitos (neste caso, o modelo SAC). Entretanto, convém destacar que, conforme descrito na estratégia empírica, há como pressuposto a presença de efeitos espaciais nas variáveis independentes e, portanto, também foi estimado<sup>22</sup> um modelo SLX.

Dentre os modelos espaciais estimados, a especificação SLX foi a que obteve o melhor ajuste, pois os valores encontrados a partir critérios de AIC e BIC foi o menor dentre as demais especificações em ambos os anos. Considerando que em 2007 foi o início do marco regulatório do saneamento, seria mais adequado analisar os coeficientes do modelo SLX para o ano de 2018 uma vez que existe uma maior possibilidade deste ano capturar os efeitos que as mudanças no setor podem ter sobre a saúde infantil. Entretanto, como as demais especificações também obtiveram significância estatística, alguns de seus principais resultados serão analisados, no intuito de fornecer maiores explicações sobre as diversas maneiras que os efeitos do saneamento podem transbordar.

Os resultados da Tabela 5 mostraram que em todos os modelos espaciais os coeficientes estimados para a eficiência variável (EV) e permanente (EP) apresentaram significância estatística e os sinais esperados, embora o nível de significância observada tenha variado entre os anos e modelos.

---

<sup>22</sup> Autores como Almeida (2012), Florax *et al.* (2003) recomendam um passo a passo econométrico para se determinar qual modelo deve ser estimado. Entretanto, Anselin (2003) pondera que esses procedimentos devem ser executados quando o objetivo é encontrar a natureza da autocorrelação espacial, ou seja, quando a AEDE é inconclusiva ou pouco se conhece da relação teórica entre as variáveis. Portanto, nesta pesquisa não se considerou estimar outros modelos espaciais tendo em vista os resultados da AEDE e, principalmente, pelas expectativas teóricas levantadas na estratégia empírica.

No modelo SEM, além das eficiências significativas, o termo de erro autorregressivo também apresentou significância, indicando que choques positivos/negativos em regiões vizinhas de um determinado local podem transbordar e afetar as DRSI de outra região. Apesar de variáveis distintas utilizadas, Silva *et al.* (2020) também identificaram esses efeitos espaciais, isto é, a melhoria nos serviços de água e esgoto beneficiavam a redução da taxa de mortalidade infantil no município e em seus vizinhos. Desse certo modo, é possível apontar para uma convergência entre os resultados dos trabalhos.

Analisando os resultados encontrados para o modelo SLX, constata-se, novamente, a significância das eficiências. Considerando que os resultados da fronteira estocástica apontaram que são as empresas regionais, autarquias ou aquelas administradas pelo município aquelas de detêm a capacidade de aumentar a eficiência permanente, esse resultado sugere que a privatização, sobretudo local, dos serviços de saneamento pode não surtir os efeitos esperados na saúde da população, conforme aponta o estudo de Saiani e Azevedo (2018).

Pelo lado da eficiência variável, sua defasagem espacial no modelo SLX também se mostrou significativa e com capacidade de reduzir as interações por DRSI. Isso significa que boas práticas que promovam uma melhora na eficiência variável em um determinado município também podem contribuir para as condições de saúde infantil em uma cidade vizinha do local onde ocorreram os ganhos de eficiência.

Esse resultado é relevante, pois, complementa os estudos anteriores que envolvem efeitos espaciais do saneamento na saúde infantil (SILVA *et al.*, 2020; VIEIRA, COSTA e JACINTO, 2020; SCRIPTORE e AZZONI, 2018; MARCONATO *et al.*, 2020). Tais trabalhos apontam, no geral, que para reduzir interações infantis por DRSI o investimento no setor de saneamento deveria aumentar para que os serviços ofertados fossem capazes de atingir toda a população. De fato, ainda existem muitas regiões que carecem de água tratada e esgotamento sanitário. Entretanto, diante do cenário econômico atual que envolve a baixa capacidade de investimento público e a insegurança jurídica que desestimula a iniciativa privada em aportar recursos ao setor, o presente trabalho aponta uma alternativa viável que seja capaz de melhorar o desempenho das firmas e ainda promover melhorias para a saúde infantil.

Essa alternativa seria aproveitar a estrutura atual, analisando os problemas gerenciais juntamente com um melhor ambiente jurídico para que a legislação e a regulação se tornem mais eficazes e, assim, obter ganhos de eficiência nos serviços que impactarão não somente na entrega de água e esgoto para a população, mas também reduzirão as taxas por interações relacionadas às DRSI.

Ainda em relação aos resultados da Tabela 5 o modelo SAR também apresentou significância estatística, tanto em 2007 quanto em 2018, para o termo defasado espacialmente tanto, neste caso, as taxas de internações por DRSI. O sinal positivo encontrado para a variável dependente defasada era esperado e vai ao encontro de resultados empíricos de outros trabalhos. Seja em uma análise por distrito como em Scriptore e Azzoni (2018) ou por microrregião, conforme o trabalho de Viera, Costa e Jacinto (2020), o efeito da vizinhança é observado. Esses resultados, somado aos encontrados neste estudo, que analisa os efeitos espaciais a nível municipal, reforçam o alcance global espacial que as taxas de DRSI podem carregar. A partir desses achados, fica evidente que as políticas públicas voltadas para a saúde infantil precisam considerar ações extremamente específicas, como no caso dos distritos, até medidas mais abrangentes que englobem microrregiões.

Por fim, na Tabela 5, estão os resultados estimados para o modelo SAC. Como foi possível perceber, as eficiências também se mostraram significativas, porém, dentre os termos espaciais, apenas o erro apresentou significância estatística em ambos os anos ainda que a variável dependente defasada espacialmente tenha obtido o sinal esperado em 2007, porém sem significância. Apesar do efeito em conjunto do termo de erro espacial autorregressivo e variável dependente defasada não ter sido significativa em 2007, seus efeitos isoladamente (modelos SAR e SEM) se mostraram adequados para capturar efeitos espaciais na relação entre DRSI e eficiência de custo.

Dentre as variáveis de controle, os resultados da Tabela 5 apontaram que a *proxy* de renda (IFDM-ER) e a cobertura do programa ESF foram significativos e com o sinal esperado em todos os modelos espaciais, exceto para o SAC em 2018. Com isso, torna-se claro que é necessário garantir os serviços médicos em locais onde a demanda é maior para que ocorra um trabalho de prevenção das DRSI. Além disso, o efeito significativo observado e com sinal esperado pelo IFDM-ER evidencia como que as desigualdades socioeconômicas podem afetar a saúde infantil. Vieira, Costa e Jacinto (2020) também mostraram como que a renda e outras variáveis demográficas podiam afetar as internações por doenças infecciosas intestinais. Em linhas gerais, se as condições de pobreza são severas, segundo os autores, maior o efeito sobre a saúde.

Conforme mencionado na estratégia empírica, a interpretação dos coeficientes de um modelo espacial é distinta daquela realizada em modelos de regressão convencional, sendo necessário uma análise separada dos efeitos marginais dos coeficientes encontrados. Nesse sentido, a Tabela 6 apresenta os resultados dos efeitos diretos, indiretos e total para os modelos espaciais estimados.

Tabela 6. Efeitos diretos, indiretos e totais dos modelos espaciais estimados.

<b>Variável dependente: DRSI</b>					
<b>Efeitos diretos</b>					
<b>Variáveis</b>	<b>Ano</b>	<b>SAR</b>	<b>SEM</b>	<b>SLX</b>	<b>SAC</b>
EP	2007	-8,1607*** (1,0684)	-7,4762*** (1,0896)	-6,7769*** (1,1035)	-7,6015*** (1,0961)
	2018	-5,7218*** (1,2597)	-5,2309*** (1,2800)	-4,3355*** (1,2993)	-6,1324*** (1,2461)
EV	2007	-17,4081** (9,4104)	-16,0582* (9,4263)	-19,4970** (9,4005)	-16,2840* (9,4291)
	2018	-1,5918 (12,5280)	-1,6785* (0,9631)	-1,5204 (0,8754)	-1,4183 (1,3942)
IFDM-ER	2007	-1,9408* (1,0992)	-1,7891* (1,1020)	-1,7839* (1,0956)	-1,8194* (1,1023)
	2018	-3,0749** (1,2387)	-2,8592** (1,2450)	-3,3194*** (0,0613)	-3,2795*** (1,2272)
Cobertura vacinal	2007	-0,0592 (0,0553)	-0,0390 (0,0551)	-0,0575 (0,0552)	-0,0428 (0,0553)
	2018	-0,0047 (0,0614)	-0,0019 (0,0614)	-0,0107 (0,0613)	0,0104 (0,0611)
Cobertura ESF	2007	-0,0363 (0,0566)	-0,0501 (0,0568)	-0,0608 (0,05677)	-0,0482 (0,0568)
	2018	-0,1079 (0,0741)	-0,1051 (0,7445)	-0,07849 (0,0741)	-0,1121 (0,0732)
<b>Efeitos indiretos</b>					
EP	2007	-0,6094*** (0,1519)	-	-7,8145*** (1,3076)	-0,1413 (0,1679)
	2018	-0,5923*** (0,1754)	-	-10,8314*** (1,7313)	-1,2203*** (0,4431)
EV	2007	-1,2999* (0,7556)	-	-4,6716*** (0,8397)	-0,3028 (0,3972)
	2018	-0,1647 (1,2483)	-	-6,1398*** (1,1540)	-0,2822 (2,3808)
IFDM-ER	2007	-0,1449* (0,0876)	-	-	-0,3380 (0,0448)
	2018	-0,3183** (0,1428)	-	-	-6,5265** (3,1360)
Cobertura vacinal	2007	-0,0044 (0,0042)	-	-	-0,0007 (0,0014)
	2018	-0,0004 (0,0063)	-	-	-0,0027 (0,0122)
Cobertura ESF	2007	-0,0027 (0,0042)	-	-	-0,0008 (0,0014)
	2018	-0,0111 (0,0079)	-	-	-0,0223 (0,0158)

(Continua...).

Tabela 6. Efeitos diretos, indiretos e totais dos modelos espaciais estimados (continuação).

		<b>Efeitos totais</b>			
EP	2007	-8,7701*** (1,1550)	-7,4762*** (1,0896)	-14,5915*** (1,5013)	-7,7428*** (1,1490)
	2018	-6,3141*** (1,3943)	-5,2309*** (1,2801)	-15,1670*** (1,9117)	-7,3842*** (1,5802)
EV	2007	-18,7081* (10,1174)	-16,0582* (9,4263)	-24,1686** (9,4660)	-16,5868* (9,6206)
	2018	-1,7566 (13,3007)	-1,6785* (0,9631)	-7,6604 (12,0877)	-1,7078 (14,3979)
IFDM-ER	2007	-2,0857* (1,1817)	-1,7891* (1,1020)	-1,7839* (1,0956)	-1,8533* (1,1247)
	2018	-3,3932** (1,36816)	-2,8592** (1,2450)	-3,3194*** (1,2371)	-3,94904*** (15,016)
Cobertura vacinal	2007	-0,0636 (0,0595)	-0,0390 (0,0551)	-0,0575 (0,0552)	-0,0435 (0,0565)
	2018	-0,0052 (0,0678)	-0,0019 (0,0614)	-0,0107 (0,0613)	-0,0125 (0,0736)
Cobertura ESF	2007	-0,0391 (0,0608)	-0,0501 (0,0568)	-0,0608 (0,05677)	-0,0491 (0,0579)
	2018	-0,1191 (0,0817)	-0,1051 (0,7445)	-0,0784 (0,0741)	-0,1350 (0,0851)

Fonte: Elaboração própria a partir de dados da pesquisa.

A partir da Tabela 6 nota-se que os efeitos diretos, indiretos e totais para as variáveis de eficiência foram significativos (exceção para o modelo SAR em 2018) e com os sinais esperados. Esses resultados apontam a relevância que se tem para a saúde infantil alcançar a eficiência neste setor. Analisando especificamente os resultados para os efeitos diretos, tomando como exemplo o modelo SLX, é possível perceber que o aumento em 1% na EP promove uma redução média de, aproximadamente, 6,7 internações por mil habitantes por motivos de DRSI em 2007 e 4,3 casos por mil habitantes para 2018. Já para a EV, se esta aumentar 1% terá uma magnitude maior nas internações, reduzindo, em média, 19,4 casos por mil habitantes em 2007 e 1,5 casos por mil habitantes em 2018.

O efeito marginal direto com sinal negativo para as eficiências evidencia que, diretamente, o esforço das empresas do setor de saneamento, em buscar melhorar sua gestão, provoca externalidades positivas para a saúde infantil. Neste caso, a busca pelo lucro (no caso empresas mistas e privadas) e otimização dos custos (para autarquias e outras empresas públicas) trazem melhorias a população.

Analisando os efeitos marginais indiretos encontrados também se percebe a significância dos coeficientes das eficiências, mantendo-se os sinais esperados. Os efeitos indiretos geram mais debates, pois ainda não há um consenso estabelecido na literatura sobre os transbordamentos que a eficiência do saneamento pode ter sobre a saúde infantil. Neste estudo, os resultados apontam para a existência de transbordamentos, ou seja, se uma cidade

melhora seus níveis de eficiência, então esse fato tende a melhorar as condições de saúde infantil uma vez que as internações por DRSI se reduzem nos municípios vizinhos. Tomando como exemplo novamente o modelo SLX, se a EP aumentar em 1%, em média, as internações por DRSI reduzem 7,8 casos por mil habitantes em 2007 e 10,8 casos por mil habitantes em 2018 para os municípios vizinhos. Já a EV, seus resultados indicaram que o aumento da eficiência em 1% reduz as internações por DRSI em 4,6 casos por mil habitantes, na média, em 2007 e em 6,1 casos por mil habitantes em 2018.

É importante lembrar que os efeitos de transbordamentos do modelo SLX são localizados, não afetando todo o sistema. Isso significa que os efeitos dos coeficientes analisados irão incidir principalmente sobre os vizinhos diretos definidos na matriz de pesos espaciais, neste caso, uma matriz de 20 vizinhos de primeira ordem como especificado na seção 5.1.3.

Analisando a literatura para confrontar os resultados deste estudo o que se observa é que alguns trabalhos como em Justino e Andrade (2020) e Silva *et al.* (2021) comentam que o investimento no setor seria uma boa forma de se reduzir indicadores de mortalidade infantil, ainda que as inversões nas regiões pudessem ocorrer de forma desigual, e com isso provocar externalidades para outras regiões. Scriptori (2016) avaliou diretamente os efeitos espaciais dos domicílios possuírem acesso adequado a água e, em seus resultados, apenas o efeito direto foi significativo, indicando que ações na melhoria dos serviços de água possuem efeito espacial somente na própria região.

Observando os resultados desses trabalhos é que a relevância de avaliar os serviços de saneamento pela ótica da eficiência ganha destaque. Indicadores de água e esgoto se mostram relevantes quando o cenário é extremamente precário, com pouca cobertura dos serviços. Quando a situação de saneamento tem avanços é comum, em alguns casos, como em Scriptori (2016) e Justino e Andrade. (2020), que variáveis como o acesso adequado a água ou qualquer outro indicador ligado a provisão dos serviços perca sua significância estatística, embora continue sendo relevante para a manutenção do status de saúde da população. A não significância pode indicar que para o saneamento continuar contribuindo com externalidades positivas para a saúde infantil é necessário que outros fatores sejam melhorados. Nesse caso, a eficiência pode surgir como um fator extra, pois, conforme já mencionado, uma empresa com uma gestão eficiente de custos pode destinar mais recursos para novos investimentos ou finalizar os que estão pendentes.

Dentre as variáveis de controle apenas o IFDM-ER apresentou efeitos diretos, indiretos e totais significativo. No modelo SLX, se o IFDM-ER aumentar em 1% o efeito total

sobre as internações por DRSI podem reduzir, em média, 1,78 casos por mil habitantes no ano de 2007 e, para o ano de 2018, seu efeito tem maior magnitude, reduzindo 3,31 casos por mil habitantes em média.

Como esse o IFDM-ER representa uma *proxy* dos níveis de emprego e renda seria natural imaginar que um maior nível de renda ocasione menos internações por DRSI, pois, com maiores rendimentos, as condições do ambiente em que se encontra o domicílio são favoráveis para a redução de doenças infecciosas. A higiene, o tipo de construção, forma de afastamento de esgoto, entre outros fatores, estão altamente correlacionados com a renda e diversos estudos (LEIVAS *et al.*, 2011; TEIXEIRA, HELLER e BARRETO, 2007) já apontavam como uma importante variável para as condições de saúde infantil. Assim, os resultados do presente estudo complementam os demais ao detectar efeitos espaciais que uma *proxy* de renda tem sobre hospitalizações infantis.

## **5.2 Desigualdade urbano-rural no saneamento e seus efeitos na saúde infantil**

Nesta seção, as análises são voltadas para a investigação das características urbanas e rurais e como que estas podem afetar a saúde infantil, a partir dos casos de DRSI, identificados pelo Programa de Saúde da Família. É importante destacar que os casos DRSI apresentados nesta seção são diferentes daqueles mencionados na seção anterior. Enquanto na seção anterior os casos de DRSI refletem as autorizações de internações hospitalares, os casos aqui relatados são aqueles diagnosticados em postos de saúde, podendo gerar ou não internações. Por este motivo, é natural que as estatísticas produzidas nesta seção sejam distintas e não devem ser comparadas.

A análise econométrica foi realizada a partir das estimações de regressões quantílicas incondicionais, que utilizam a Função de Influência Recentrada (RIF) para decompor, em cada quantil de interesse, os efeitos explicados e não explicados das variáveis independentes sobre as diferenças de casos de DRSI, nas zonas urbana e rural. Convém mencionar que, quando um efeito positivo é encontrado, então isso significa um aumento nos casos de DRSI e, portanto, reflete uma piora nas condições de saúde infantil. Dessa maneira, podem ocorrer situações em que a redução do diferencial de casos de DRSI, entre os meios urbano e rural sejam devido ao aumento desses casos.

### **5.2.1 Análise Descritiva**

A análise das principais estatísticas descritivas foi separada entre urbano e rural, para uma caracterização de ambos os grupos a serem comparados, posteriormente, pela

decomposição quantílica. A seguir, a Tabela 7 exibe algumas estatísticas descritivas para a variável dependente e a Tabela 8 apresenta os principais resultados descritivos das demais variáveis utilizadas<sup>23</sup>. Dentre os 5.565 municípios brasileiros, foi possível utilizar 77% desse total, para compor a amostra urbana e, para o caso rural, devido a maior ausência de informações, foram considerados 52% dos municípios para integrar a amostra. Portanto, os dados consideram municípios que possuíam informações completas para o período estudado (1998 a 2015).

Tabela 7. Estatísticas descritivas dos casos de DRSI/mil hab. segundo os meios urbano e rural.

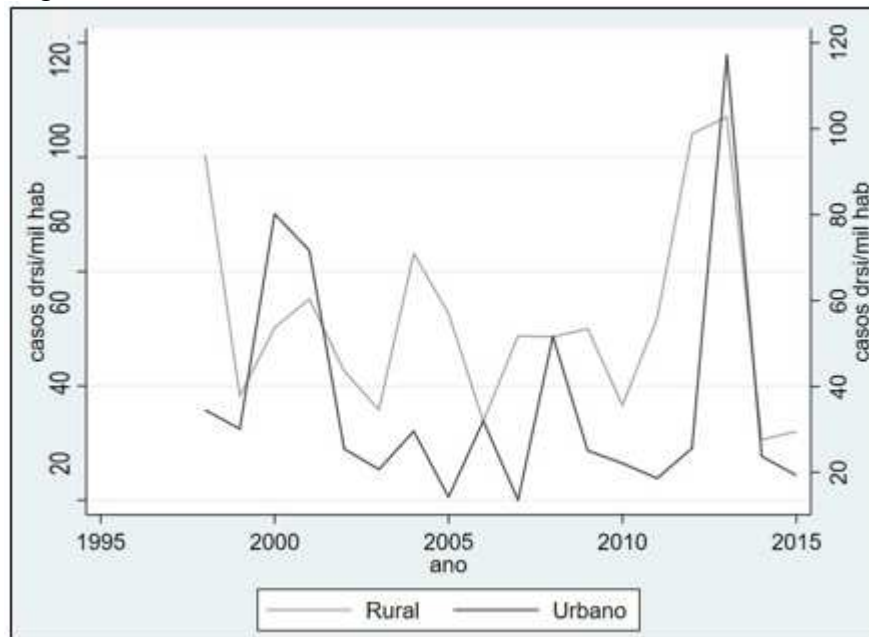
<b>Grupo</b>	<b>Observações</b>	<b>q10</b>	<b>q25</b>	<b>q50</b>	<b>Média</b>	<b>q75</b>	<b>q90</b>	<b>q99</b>
Urbano	77.041	0,32	0,78	2,12	7,66	4,99	9,73	41,66
Rural	46.758	0,47	1,10	2,88	10,28	6,80	13,28	64,61

Fonte: Dados da pesquisa.

Em relação à variável de interesse, a média dos casos de DRSI, entre os meios urbano e rural, está de 3 vezes (urbano) a 5 vezes (rural) maior se comparada com sua mediana, em termos absolutos. Isso evidencia que existem valores discrepantes potenciais que influenciam o valor da média dos casos. Com isso, uma análise que considere apenas a média pode ser problemática e gerar resultados inconsistentes ou ineficientes. Ao observar o número de casos entre os quantis, é possível notar que o distanciamento de casos entre os meios urbano e rural se inicia principalmente a partir do quantil 75. Para observar o comportamento dos casos de DRSI ao longo dos anos, a Figura 7 ilustra, para as zonas urbana e rural, a evolução média dos casos no período analisado.

<sup>23</sup> É importante destacar que, preliminarmente, os dados referentes ao Distrito Federal e ao estado do Acre foram excluídos da análise, tendo em vista a elevada falta de informações para as variáveis do meio rural.

Figura 7. Casos médios de DRSI nas zonas urbana e rural entre 1998-2015.



Fonte: Dados da pesquisa.

A partir da Figura 7, é possível ter maior clareza do comportamento dos casos de DRSI entre o meio urbano e rural. Excetuando-se no começo do período, a zona rural permaneceu a frente da urbana na média de casos registrados, se aproximando no final do período em que ambas apresentam um crescimento acentuado entre 2012-2013 até a queda abrupta em 2014. Não há, todavia, um evento específico que explique totalmente essa alta repentina nos casos de DRSI. No entanto, alguns fatos podem ajudar a explicar tal comportamento. O primeiro é que, em 2013, o Brasil passou por um surto de dengue em todo o território nacional aumentando, dessa maneira, casos de diarreia e desidratação. Dados do DATASUS<sup>24</sup> mostram uma média de 11,1 casos de dengue por mil habitantes neste período, evidenciando o grande número de casos, uma vez que as diretrizes do SUS consideram 3 casos por mil habitantes o valor para se determinar surto de alguma doença.

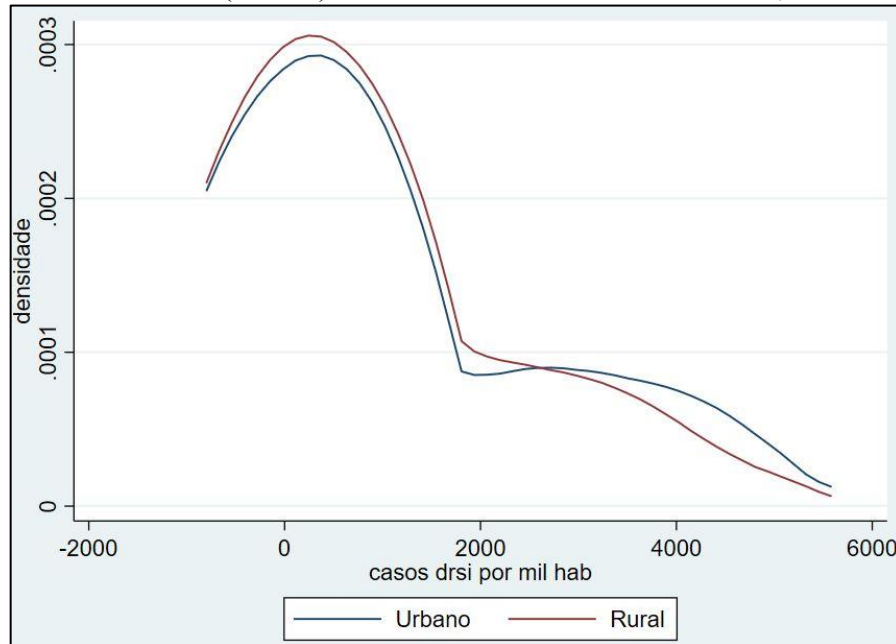
Outro fato, que também pode ter contribuído para o crescimento acentuado das DRSI, foi a queda observada na Figura 6 da eficiência de custo no setor de saneamento no período 2011-2012. Embora no meio rural há pouca participação de empresas de saneamento, a zona urbana é dependente dessas firmas, para que os serviços de saneamento contribuíssem para a manutenção da saúde, principalmente a infantil.

Embora a análise da média ao longo dos anos possa fornecer informações sobre o comportamento dos casos de DRSI em crianças com até 5 anos de vida, na presença de valores discrepantes esses valores, mesmo que observados por um longo período, podem ser

<sup>24</sup> Disponível em: <<https://portalsinan.saude.gov.br/dengue>>.

influenciados e não retratam com exatidão a realidade observada nos municípios. Assim, é necessária uma investigação para além da média. A Figura 8 apresenta a densidade dos casos de DRSI entre os meios urbano e rural.

Figura 8. Densidade (Kernel) das DRSI nas zonas urbana e rural, 1998 a 2015.



Fonte: Elaboração própria a partir de dados da pesquisa.

Os gráficos de densidade são úteis, pois possibilitam o conhecimento da distribuição dos dados. A partir da Figura 8, fica evidente a assimetria à direita existente na distribuição dos casos de DRSI tanto no meio urbano quanto no rural devido aos *outliers*. Além disso, a distribuição dos casos de DRSI no meio rural permanece acima do meio urbano até o limite de 2200 casos/mil hab. Isso revela que há mais casos discrepantes na zona urbana se comparado ao meio rural. As áreas urbanas que estão periféricas aos grandes centros possuem características tais como, alta densidade demográfica, saneamento e moradias precárias, baixos níveis de renda, entre outras, que podem favorecer o surgimento de valores discrepantes de casos de DRSI em crianças. Segundo Silveira (2013) regiões com essas características estão mais suscetíveis a surtos de casos de DRSI, o que explica, em partes, uma densidade maior em áreas urbanas quando os casos de DRSI ultrapassam os 2200/mil hab.

Com intuito de complementar a análise descritiva, a Tabela 8 exhibe algumas estatísticas descritivas das variáveis explicativas utilizadas no modelo, separando os resultados entre a zona urbana e rural<sup>25</sup>.

<sup>25</sup> Em alguns casos, quando as médias dos grupos são muito semelhantes, é comum que se apliquem testes de diferença de médias para evidenciar se as médias dos grupos são estatisticamente diferentes. No entanto, os

Tabela 8. Estatísticas descritivas das variáveis de saneamento para os meios urbano e rural (1998 a 2015).

<b>Zona Urbana</b>				
<b>Variáveis (%)</b>	<b>Observações</b>	<b>Média</b>	<b>Desvio Padrão</b>	<b>Coefficiente de Variação (%)</b>
Taxa de Analfabetismo	77.041	17,61	10,14	57,63
Condição de água (insegura)	77.041	3,44	8,33	241,73
Condição de água (segura)	77.041	81,26	20,75	25,53
Condição de Esgoto (fossa)	77.041	53,81	36,24	67,35
Condição de Esgoto (céu aberto)	77.041	9,51	13,26	139,47
Plano de saúde	77.018	7,79	8,77	112,62
Moradias de taipa	77.041	4,69	10,92	221,92
Moradias de madeira e material aproveitado	77.041	15,46	24,67	159,57
Moradias inadequadas	77.041	18,01	23,94	132,91
<b>Zona Rural</b>				
Taxa de Analfabetismo	46.760	27,47	14,77	53,76
Condição de água (insegura)	46.706	13,49	19,75	146,39
Condição de água (segura)	46.756	33,53	26,52	79,10
Condição de Esgoto (fossa)	46.758	57,73	28,33	49,06
Condição de Esgoto (céu aberto)	46.758	32,20	26,44	82,12
Plano de saúde	46.758	3,37	6,80	201,95
Moradias de taipa	46.758	11,97	18,59	155,30
Moradias de madeira e material aproveitado	46.758	2,89	6,09	210,72
Moradias inadequadas	46.758	30,65	30,58	99,74

Fonte: Dados da pesquisa.

Os resultados da Tabela 8 apresentam elevados coeficientes de variação para as variáveis estudadas, mostrando que, tanto a zona urbana como a rural, possuem grande variabilidade em seus conjuntos de dados. Além disso, como apontado por Silveira (2013), é possível perceber que a zona rural possui resultados mais desfavoráveis, uma vez que o percentual de moradias inadequadas, esgoto a céu aberto e abastecimento de água por fonte insegura são maiores no meio rural.

Apesar disso, é na zona urbana que se observa a maior heterogeneidade amostral. Isso pode indicar a existência de lugares no meio urbano, como regiões periféricas de grandes centros, com condições de saneamento semelhantes ou menos adequadas que aquelas encontradas no meio rural. De acordo com Venson, Jacinto e Sbicca (2019), não é raro que

---

valores encontrados para os grupos pesquisados estão muito distantes, não sendo necessário avaliar tal suposição.

essas regiões tenham condições de saneamento inadequadas. Segundo os autores, nesses locais a infraestrutura é precária e, somada ao contexto social e econômico das famílias mais pobres que ali residem, criam um cenário desfavorável para o combate de casos de DRSI, pois, famílias mais pobres, com menos renda e baixa escolaridade podem ter mais resistência em procurar auxílio médico para prevenir as doenças, deixando ocorrer sinais de agravo para buscar ajuda em hospitais, postos de saúde, entre outros tipos de atendimento.

Ainda em relação aos resultados da Tabela 8, verifica-se que é na zona rural que existe o maior percentual de moradias inadequadas, variável que agrega o percentual de todas as moradias que não são construídas por alvenaria. Entretanto, ao subdividir essa variável, fica constatado que o meio urbano possui um maior percentual (15,46% do total de moradias) de casas construídas com material aproveitado como madeira, papelão, plásticos, entre outros. Por outro lado, na zona rural, as casas feitas com taipa e pilão representam 11,97% e são predominantes dentro da categoria de moradias inadequadas. Dado que existem essas diferenças nas especificidades das moradias inadequadas, para captar melhor esse efeito, é mais interessante utilizar um indicador que agregue todas essas condições, pois assim se teria um retrato da realidade da inadequação quanto a salubridade nessas regiões e não captaria só as especificidades das variáveis separadas.

Os resultados da Tabela 8 fornecem apenas um retrato sobre o comportamento na média para as variáveis explicativas. Como a análise proposta segue uma abordagem quantílica, é interessante que o comportamento médio dessas variáveis seja avaliado a partir dos quantis de distribuição da variável de interesse que, neste caso, são os casos de DRSI por mil habitantes diagnosticados nos PSF's. Nesse sentido, a Tabela 9 exhibe as médias para cada variável explicativa segundo os limites de cada quantil observado na Tabela 7 dos casos de DRSI. A lógica em avaliar essas estatísticas descritivas é verificar se as variáveis independentes aumentam ou se reduzem conforme a distribuição de casos das doenças ocorrem ao longo do quantis.

Tabela 9. Médias das variáveis explicativas em cada quantil de distribuição dos casos de DRSI.

<b>Zona Urbana</b>						
<b>Variáveis (%)</b>	<b>q10</b>	<b>q25</b>	<b>q50</b>	<b>q75</b>	<b>q90</b>	<b>q99</b>
Taxa de Analfabetismo	13,78	14,71	15,77	16,71	17,30	17,61
Condição de água (insegura)	1,54	1,82	2,34	2,84	3,19	5,06
Condição de água (segura)	83,87	83,63	83,10	82,26	81,72	73,71
Condição de Esgoto (fossa)	5,89	6,43	7,13	8,07	8,84	13,78
Condição de Esgoto (céu aberto)	42,54	44,32	47,81	51,15	52,85	60,86
Plano de saúde	10,49	10,15	9,35	8,88	8,03	7,80
Moradias inadequadas	18,72	16,89	16,86	16,99	17,41	28,93
<b>Zona Rural</b>						
Taxa de Analfabetismo	22,65	23,09	24,55	26,22	27,12	27,05
Condição de água (insegura)	7,99	8,44	10,16	11,74	12,79	15,77
Condição de água (segura)	35,54	34,98	34,65	34,20	33,88	33,75
Condição de Esgoto (fossa)	23,60	24,10	26,37	29,15	31,03	33,58
Condição de Esgoto (céu aberto)	60,03	61,31	60,92	59,56	58,47	56,73
Plano de saúde	5,81	5,43	4,66	3,95	3,57	2,87
Moradias inadequadas	27,54	28,52	28,72	28,90	29,60	38,74

Fonte: Dados da pesquisa.

A Tabela 9 mostra que, de uma maneira geral, quando existem mais casos de DRSI, tanto na zona urbana quanto na rural, as características ambientais tendem a ser mais precárias. Isso pode ser visualizado, por exemplo, nas condições de água insegura onde seu percentual aumenta gradualmente na medida em que se avança para os quantis superiores da distribuição de casos de DRSI. Outras variáveis, como o plano de saúde e as moradias inadequadas, também podem sinalizar para um outro tipo de deterioração, neste caso, a da renda familiar. São nos quantis onde estão a maior quantidade de DRSI diagnosticadas que se percebe os menores percentuais médios dessas variáveis, sugerindo, conforme mencionado por Almond, Currie e Hermann (2012), que as características ambientais possuem grande relevância para melhorar o *status* de saúde infantil.

No geral, o que pode ser observado, a partir dessas estatísticas descritivas, é que a distribuição dos casos de DRSI entre os meios urbano e rural, ao longo dos quantis, pode ter sido afetada pelas diferentes características entre eles, que são mais desfavoráveis no meio rural. Ademais, observou-se também uma variabilidade das variáveis analisadas dentro das zonas urbana e rural, conforme indicado pelo coeficiente de variação. Assim, uma análise que contemple somente a média dos casos não será totalmente conclusiva sobre as diferenças existente entre os grupos. Logo, se faz necessário investigar, por meio da decomposição quantílica, como os casos de DRSI em crianças de até 5 anos modificam-se, em resposta ao

conjunto de variáveis explicativas, identificando quais fatores contribuem para os diferenciais de casos entre as zonas urbanas e rurais dos municípios brasileiros.

### 5.2.2 Resultados Econométricos

As regressões quantílicas (RQI) foram realizadas considerando os quantis 10, 25, 50, 75 e 90, para representar toda a distribuição incondicional dos casos de DRSI. Os resultados para a zona urbana são apresentados na Tabela 10, sendo que a primeira coluna de resultados apresenta o modelo estimado por MQO, seguida das estimações para cada um dos quantis aqui utilizado. O que pode ser observado é que se o modelo considerado fosse o MQO, os resultados poderiam levar a uma conclusão errônea de insignificância dos coeficientes das variáveis. Isso evidencia como que uma análise na média pode ser afetada quando existem grandes heterogeneidades entre as observações.

Em geral, dentre as variáveis com significância, os resultados encontrados no modelo estimado por MQO parecem se assemelhar com alguns outros estudos (SILVEIRA, 2013; SAIANI e TONETO JUNIOR, 2006), pois foi observada a influência que o ambiente (moradias inadequadas), a educação (taxa de analfabetismo) e esgotamento a céu aberto têm sobre a saúde infantil. Embora a *proxy* de renda não tenha sido significativa para a análise na média, essa é outra variável relevante para que as condições de saúde infantil melhorem ou não.

Tabela 10. Resultados econométricos para a zona urbana, média e quantis da distribuição incondicional das DRSI (1998 – 2015).

Variáveis	MQO	q10	q25	q50	q75	q90
Condição de água (insegura)	0,00384 (0,0405)	0,00061* (0,00030)	0,0018** (0,0004)	0,0049** (0,0018)	0,0177*** (0,0030)	0,0469*** (0,0138)
Condição de esgoto (céu aberto)	0,0306* (0,0152)	0,0014*** (0,00016)	0,0015*** (0,0002)	0,0049*** (0,0005)	0,0067*** (0,0009)	0,0146*** (0,0025)
Condição de esgoto (fossa)	0,0413 (0,0281)	0,0020*** (0,0003)	0,0024*** (0,0004)	0,0123*** (0,0012)	0,0348*** (0,0024)	0,0980*** (0,0086)
DRSI <sub>t-1</sub>	0,00007 (0,0005)	0,00004 (0,00007)	0,00003 (0,00006)	0,00008 (0,0001)	0,0001 (0,0003)	0,0003 (0,0005)
Plano de saúde	-0,1340 (0,0805)	-0,0023*** (0,0005)	-0,0074*** (0,0009)	-0,0211*** (0,0020)	-0,0394*** (0,0039)	-0,0440*** (0,0096)
Taxa de analfabetismo	-0,0544* (0,024)	0,0017** (0,0005)	0,0024** (0,0008)	0,0025 (0,0017)	0,0070 (0,0040)	0,0215* (0,010)
Moradias	0,0848* (0,0371)	0,0012*** (0,0002)	0,0035*** (0,0003)	0,0096*** (0,0009)	0,0273*** (0,0017)	0,0544*** (0,0054)
Constante	6,550 (5,77)	0,4001*** (0,0183)	1,180*** (0,0436)	3,2662*** (0,1140)	7,9955*** (0,2616)	14,4923*** (0,9419)
Efeito fixo tempo	Sim	Sim	Sim	Sim	Sim	Sim

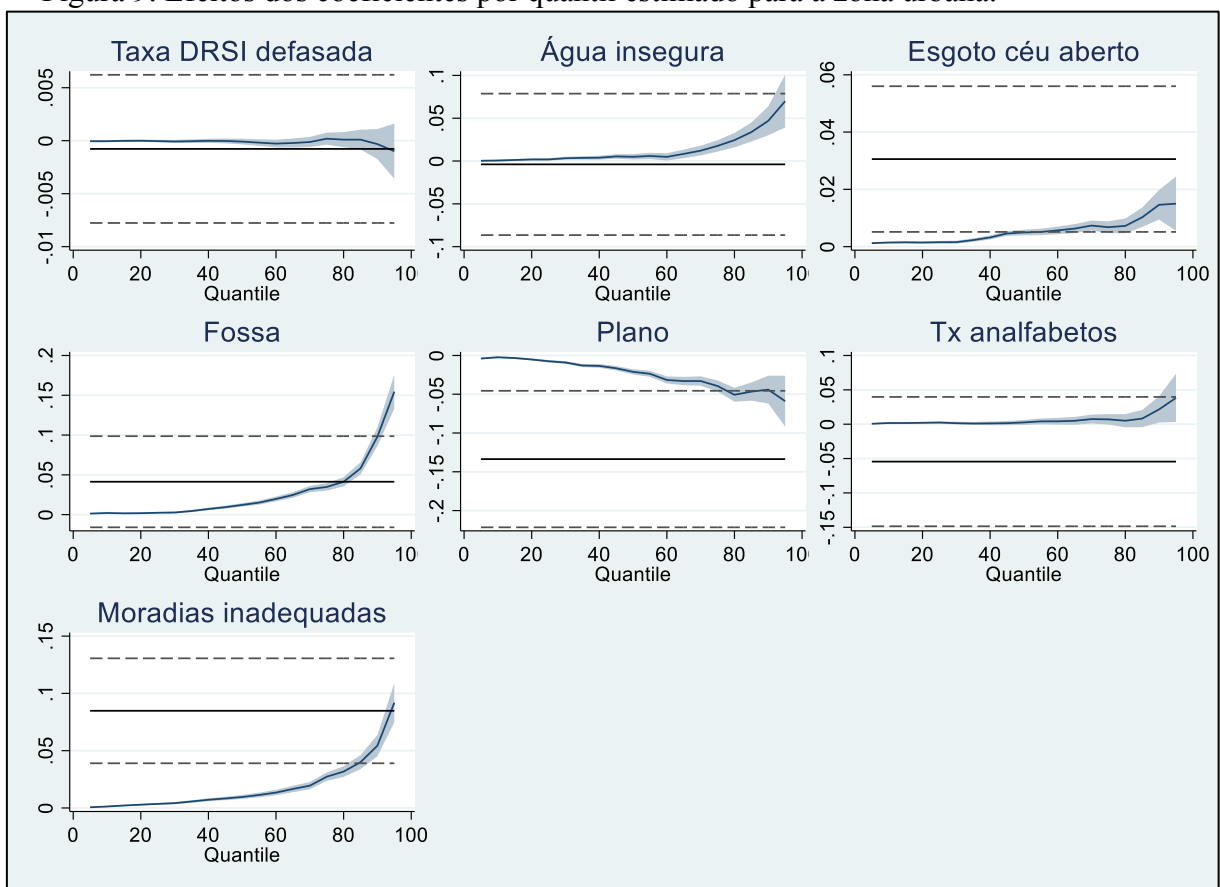
Efeito fixo UF	Sim	Sim	Sim	Sim	Sim	Sim
R <sup>2</sup>	0,0026	0,0835	0,1213	0,1495	0,1211	0,0635
Observações	77098	77098	77098	77098	77098	77098
Estatística F	8,63***	149,26***	226,29***	288,27***	225,81***	111,15***
Média RIF	-	0,3267	0,7851	2,1257	4,9976	9,7343

Fonte: Dados da pesquisa.

Obs: \*\*\*p<0,001; \*\*p<0,01; \*p<0,05. Erros padrão obtidos pelo procedimento bootstrap.

Em relação às estimações das RQI's, os resultados da Tabela 10 mostram que a maioria das variáveis mantiveram sua significância, ao longo dos quantis, ao nível de 0,1% de significância. Além disso, também é possível perceber que os efeitos dos coeficientes aumentam a medida em que os quantis da distribuição incondicional de casos de DRSI aumentam. A Figura 9 ilustra esse comportamento e compara, para cada variável, o quão distante ficaram os coeficientes estimados, em cada quantil, dos resultados obtidos pelo MQO, destacados pela linha cheia na horizontal. As linhas tracejadas na horizontal e as áreas sombreadas, que acompanham o comportamento dos coeficientes estimados em cada quantil representam, respectivamente, os intervalos de confiança para as estimações por MQO e RQI.

Figura 9. Efeitos dos coeficientes por quantil estimado para a zona urbana.



Fonte: Elaboração própria a partir de dados da pesquisa.

As variáveis ligadas à água e ao esgoto apresentaram o comportamento esperado, isto é, seus sinais foram positivos e significativos ao longo dos quantis selecionados. Como pode ser observado na Figura 9, o efeito das condições de água insegura, nos primeiros quantis, são praticamente os mesmos. Nos quantis 10 e 25, o aumento em 1 ponto percentual na proporção total de domicílios com condições de inseguras de água provocam um aumento na distribuição incondicional dos casos de DRSI em 0,00061 e 0,0018 casos por mil habitantes, enquanto nos quantis com mais casos (75 e 90), o crescimento em 1 ponto percentual na proporção dos domicílios com condições inseguras de água reflete em um aumento 0,0177 e 0,0469 casos por mil habitantes na distribuição incondicional de casos de DRSI.

Entretanto, é a partir da mediana que os efeitos das condições de água insegura crescem rapidamente, indicando que, em regiões onde se observa mais casos de DRSI em crianças, a falta de uma fonte de água segura tem um efeito negativo mais intenso na saúde infantil. Isso explica de alguma forma a seletividade em algumas políticas voltadas para o saneamento. Conforme Saiani e Oliveira (2018) apontam, regiões com maior crescimento econômico e demográfico tendem a receber com mais frequência recursos para políticas de saneamento, pois, além da necessidade, são áreas vantajosas politicamente. Porém, o governo parece sofrer com certa “miopia” ao adotar esse perfil para as políticas de saneamento, ou seja, o gestor público não consegue visualizar que, no longo prazo, tais medidas podem favorecer um aumento na desigualdade de acesso aos serviços de água e esgotamento sanitários seguros. Com isso, a população de regiões com menor crescimento e densidade demográfica é obrigada a buscar soluções próprias para o tratamento de água e destinação final do esgoto.

Destaca-se também que a variável condição de esgoto (fossa) obteve uma maior influência sobre os quantis de distribuição dos casos de DRSI, em relação ao esgoto a céu aberto. É natural que problemas de esgoto a céu aberto ocasionem muito malefícios para a saúde infantil, como já apontado por Silveira (2013). Entretanto, os resultados da Tabela 10 podem apontar para um perigo “oculto” em se utilizar esgotamento do tipo de fossa, sobretudo rudimentar. Ao perceber o esgoto a céu aberto, é comum evitar tais regiões, pois ele é visível. O problema da fossa é que ela pode apresentar infiltrações subterrâneas e contaminar os lençóis de água de toda uma região e, considerando um abastecimento de água inseguro via poços ou bicas, torna-se ainda mais grave seus efeitos negativos.

Ainda em relação às variáveis de esgoto, verifica-se que, mesmo que o esgotamento por fossa exerça a maior influência a partir da mediana da distribuição dos casos de DRSI, nos quantis inferiores seus efeitos são muito semelhantes. Se for considerado o quantil 10, para o

esgoto a céu aberto, o aumento em 1 ponto percentual nessa variável afeta a distribuição incondicional dos casos de DRSI em 0,0014 casos por mil habitantes e, quando as condições de esgoto são concedidas por meio de fossa, o acréscimo de 1 ponto percentual de tal variável modifica a distribuição dos casos de DRSI em 0,0020 casos por mil habitantes.

O maior efeito no quantil 90 das variáveis relacionadas ao esgoto e a água pode estar ligado à dinâmica que as DRSI em crianças com até 5 anos possuem com diversos aspectos que permeiam o saneamento. A quantidade de DRSI aumenta onde a média das variáveis de controle é pior, isto é, aumentam-se os níveis de analfabetismo, moradias inadequadas e reduz o percentual de planos de saúde ativos. Isso mostra que, conforme o meio no qual a criança está inserida se deteriora, a exposição a condições de água insegura e esgotamento inadequado acelera seus efeitos, que já são intensificados pela própria dinâmica do contágio das DRSI.

Em relação aos resultados encontrados para as variáveis de controle, foram obtidos os sinais esperados. A *proxy* de renda, plano de saúde, mostrou-se capaz de reduzir os casos de DRSI, em todos os quantis. Embora não seja tão comum associar essa variável com a renda, o fato de uma família possuir plano de saúde ativo é um indicativo de uma maior renda e se esta é designada exclusivamente para cuidados de saúde, então é razoável acreditar que percentuais maiores de famílias com plano, podem contribuir para diminuir casos de DRSI em crianças de até 5 anos. Assim como nas variáveis de água e esgoto, seu efeito, em termos absolutos, é maior para os últimos quantis. Enquanto o aumento em 1 ponto percentual de famílias com plano ativo provoca uma redução na distribuição incondicional das DRSI em 0,0023 casos por mil habitantes no quantil 10, para o quantil 90, essa mesma variação modifica (reduz) a distribuição incondicional dessas em doenças 0,044 casos por mil habitantes, um efeito 19 vezes maior se comparado ao primeiro quantil analisado.

Outro resultado relevante para o meio urbano é o efeito da variável de moradia inadequada. Assim como as demais variáveis de controle, nos quantis mais baixos seu efeito modifica pouco a distribuição incondicional dos casos de DRSI, embora sejam significativos estatisticamente. Analisando seu efeito para o quantil 90, é esperado que o aumento em 1 ponto percentual de moradias inadequadas modifique a distribuição de casos DRSI em 0,054 casos por mil habitantes. O que merece ser destacado neste resultado para esse quantil é que seu efeito só não foi superior a condição de esgoto (fossa). Isso mostra que o ambiente no qual a família reside tem grande influência sobre a saúde infantil.

Em relação a *proxy* de educação, seus resultados não foram significativos em todos os quantis e muito próximos de zero na maioria dos casos, ainda que os sinais encontrados foram

os esperados. Apesar de a educação ser uma variável relevante para a saúde infantil, o percentual de analfabetos nas famílias captura apenas se existe alguém no domicílio que possui essa característica, podendo ser avós, tios, irmãos ou pais. Caso o indivíduo analfabeto não seja o pai ou mãe, é razoável supor que seu conhecimento sobre higiene e opinião sobre como cuidar das crianças do domicílio tenha pouca influência na saúde infantil, explicando a pouca influência observada de tal variável.

A seguir, na Tabela 11 são apresentados os resultados das RQI's para a zona a rural. É possível perceber semelhanças e diferenças com os resultados do meio urbano. Um fato comum é que os coeficientes dos últimos quantis, que foram significativos, também apresentaram os maiores efeitos sobre as DRSI. Isso pode significar que, independente do ambiente (rural ou urbano) no qual as crianças vivem, os quantis com maiores casos tendem a apresentar efeitos mais fortes para os coeficientes estimados.

Tabela 11. Resultados econométricos para a zona rural, média e quantis da distribuição incondicional das DRSI (1998 – 2015).

Variáveis	MQO	q10	q25	q50	q75	q90
Condição de água (insegura)	-0,0576* (0,027)	0,00012 (0,0004)	0,00064 (0,0006)	0,0014 (0,0017)	0,0034 (0,0031)	0,0155 (0,0069)
Condição de esgoto (céu aberto)	-0,0072 (0,032)	-0,00078 (0,0005)	-0,00063 (0,0006)	0,0045** (0,0015)	0,0126*** (0,0026)	0,0272*** (0,0066)
Condição de esgoto (fossa)	0,0433 (0,0462)	0,0005 (0,0007)	0,00018 (0,0007)	0,0034 (0,0018)	0,010** (0,0028)	0,0295*** (0,0071)
DRSI <sub>t-1</sub>	0,0147 (0,0152)	0,00005 (0,00009)	0,00008 (0,00009)	0,0001 (0,00019)	0,0007 (0,0004)	0,0017 (0,0015)
Plano de saúde	0,0284 (0,123)	-0,0121*** (0,0012)	-0,0193*** (0,0017)	-0,0391*** (0,0034)	-0,0535*** (0,0064)	-0,0835*** (0,0147)
Taxa de analfabetismo	0,0108 (0,0444)	0,00074 (0,0007)	0,0017* (0,0083)	0,0050*** (0,0020)	0,0053 (0,0063)	0,0073 (0,0091)
Moradias	0,0668* (0,0307)	0,0015*** (0,0004)	0,0028*** (0,0004)	0,0085*** (0,0012)	0,0203*** (0,0027)	0,0438*** (0,0068)
Constante	42,524 (58,315)	0,9507*** (0,0744)	2,3105*** (0,1016)	6,2310 (1,0338)	6,5889*** (2,8365)	11,2550*** (5,8010)
Efeito fixo tempo	Sim	Sim	Sim	Sim	Sim	Sim
Efeito fixo UF	Sim	Sim	Sim	Sim	Sim	Sim
R <sup>2</sup>	0,0024	0,0622	0,1114	0,1391	0,1099	0,0571
Observações	46860	46860	46860	46860	46860	46860
Estatística F	7,33***	66,05***	124,84***	160,92***	123,03***	60,34***
Média RIF	-	0,4720	1,100	2,8883	6,8085	13,286

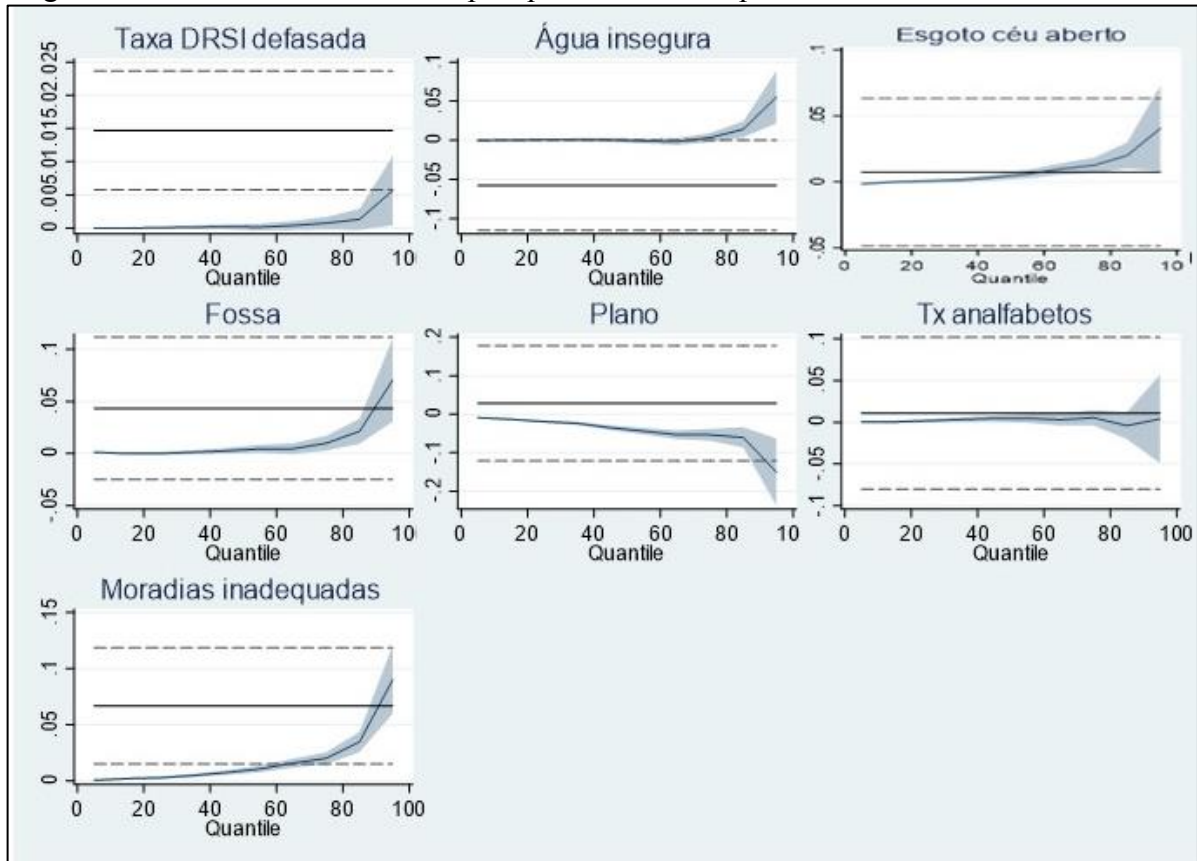
Fonte: Dados da pesquisa.

Obs: \*\*\*p<0,001; \*\*p<0,01; \*p<0,05. Erros padrão obtidos pelo procedimento *bootstrap*.

Por outro lado, uma diferença observável na Tabela 11 está relacionada à significância dos coeficientes estimados. Enquanto na Tabela 10 a grande maioria das variáveis foi significativa, em todos os quantis da distribuição, os resultados na Tabela 11 indicam que

apenas a partir da mediana da distribuição é ocorreu a predominância da significância estatística em algumas variáveis. Um dos fatores que pode ter contribuído para tal resultado é a redução da amostra para o meio rural. Porém, destaca-se que os sinais encontrados para as variáveis significativas foram os esperados, assim como na zona urbana. A Figura 10 ilustra o comportamento ao longo dos quantis para as variáveis explicativas.

Figura 10. Efeitos dos coeficientes por quantil estimado para a zona rural.



Fonte: Elaboração própria a partir de dados da pesquisa.

O comportamento das variáveis explicativas, ao longo dos quantis, no meio rural foi semelhante ao que foi observado na zona urbana, conforme pode ser visualizado na Figura 10. Entretanto, algumas diferenças merecem ser discutidas.

O efeito da condição de água insegura sobre os casos de DRSI em crianças com até 5 anos na zona urbana se mostrou relevante, pois, seus coeficientes eram significativos em todos os quantis, embora nos quantis inferiores houvesse pouca influência, em termos absolutos. Para a zona rural, o que se observou foi a insignificância dos seus efeitos, tendo em vista que em nenhum quantil os coeficientes foram significativos. Isso pode ser explicado pela capacidade de adaptação das famílias rurais ao ambiente em que estão inseridas. Historicamente, os domicílios rurais não possuem acesso aos serviços de saneamento com frequência, sendo necessário realizar tratamentos alternativos para a água que é obtida por

fontes inseguras. Medidas como, por exemplo, ferver a água utilizada para beber e cozinhar, parecem ser de alta relevância para a saúde das crianças quem convivem nesse meio, como bem destacado por alguns autores (JEULAND *et al.*, 2013; TEIXEIRA, HELER e BARRETO, 2007).

As variáveis que refletem o esgotamento sanitário apresentaram significância estatística somente nos últimos quantis da distribuição de casos de DRSI. A influência dos efeitos das condições de esgoto a céu aberto e fossa ficaram muito próximos, tanto no quantil 75 como no 90, diferente do que foi observado para a zona urbana onde o efeito das condições de esgoto por fossa foi muito superior. Além disso, o crescimento dos efeitos dos coeficientes de cada variável entre os quantis também foi menor. Enquanto no quantil 75 o aumento em 1 ponto percentual do esgotamento a céu aberto modificava a distribuição de casos de DRSI em crianças em 0,012 casos por mil habitantes, no quantil 90 esse efeito foi de 0,027 casos por mil habitantes. De forma semelhante, no quantil 75, o aumento em 1 ponto percentual do esgotamento por fossa modificava a distribuição dos casos de DRSI em crianças em 0,010 casos por mil habitantes, ao passo que no quantil 90 a mesma variação levava a uma mudança de 0,029 casos por mil habitantes.

Isso mostra que, para as zonas rurais, a transmissibilidade das DRSI entre as crianças parece ter uma dinâmica mais lenta mesmo que as condições de saneamento se deteriore. A distância entre os domicílios, a menor densidade demográfica, o menor convívio social com pessoas de outras regiões, entre outras características, pode ter contribuído para que as diferenças observadas entre os últimos quantis não fossem muito diferentes. Entretanto, alguns controles como plano de saúde ativo e moradias inadequadas ainda podem potencializar o surgimento dos casos de DRSI em crianças com até 5 anos uma vez que seus coeficientes foram significativos em todos os quantis e com efeitos crescentes, em valores absolutos.

Em relação às variáveis de controle plano de saúde e moradias inadequadas o que se observou foi que, para a zona rural, ter um plano de saúde ativo na residência está relacionado negativamente com o status de *saúde* infantil nos quantis analisados e, conforme ocorre a elevação dos casos de DRSI, as características ligadas aos domicílios familiares se tornam mais influentes sobre as condições de saúde infantil.

Em linhas gerais, a partir dos resultados das RQI's, pode-se apontar diferentes níveis de influência das características observáveis nos meios urbano e rural, sobre a distribuição de casos de DRSI. Entretanto, esses resultados evidenciam apenas como que as DRSI se relacionam com as características específicas de cada localidade. Para avançar no debate

sobre como a desigualdade de acesso aos serviços de saneamento pode influenciar a saúde infantil, é necessário considerar o que aconteceria na zona rural caso esta possuísse as características de zona urbana. Em outras palavras, a construção de uma distribuição contrafactual pode auxiliar no processo de identificação dos fatores que explicam as diferenças na distribuição de casos de DRSI entre os dois grupos estudados. Para tanto, na Tabela 12 são apresentados os resultados da decomposição quantílica<sup>26</sup>.

Tabela 12. Decomposição das diferenças entre as zonas urbano e rural em cada quantil da distribuição de casos de DRSI em crianças de até 5 anos.

	q10	q25	q50	q75	q90
Grupo 1 (Urbano)	0,327*** (0,0024)	0,785*** (0,0052)	2,126*** (0,0150)	4,998*** (0,0343)	9,734*** (0,072)
Grupo 2 (Rural)	0,472*** (0,0054)	1,100*** (0,0102)	2,884*** (0,0252)	6,808*** (0,0511)	13,290*** (0,111)
Diferença	-0,145*** (0,0060)	-0,315*** (0,0130)	-0,758*** (0,0314)	-1,811*** (0,0578)	-3,551*** (0,136)
Explicada (Efeito Composição)	-0,2601*** (0,0143)	-0,5637*** (0,0176)	-1,5347*** (0,0519)	-2,600*** (0,102)	-3,0309*** (0,235)
Não Explicada (Efeito Coeficiente)	0,185*** (0,0106)	0,392*** (0,0161)	1,058*** (0,0521)	1,341*** (0,0955)	0,7621** (0,216)
Constante	0,172** (0,0485)	0,318*** (0,0631)	0,668*** (0,138)	-0,0957 (0,282)	-3,920*** (0,506)
Observações	123.957	123.957	123.957	123.957	123.957

Nota: Erros padrão obtidos por procedimento *bootstrap*.

Obs: \* p<0,05; \*\* p<0,01; \*\*\* p<0,001.

Fonte: Dados da pesquisa.

A partir dos resultados da Tabela 12, é possível notar que o diferencial entre os casos de DRSI nos grupos urbano e rural é negativo, ou seja, quanto mais negativo ficam essas diferenças, mais desfavorável é a situação da zona rural em relação aos casos de DRSI. Esse diferencial é significativo e crescente (mais negativo), aumentando nos quantis superiores da distribuição, sendo mais acentuado no quantil 90. Portanto, pode-se afirmar que a zona rural tem características que são responsáveis por acentuar as diferenças nos casos de DRSI, comparado a zona urbana. Adicionalmente, é possível explorar de onde essas diferenças são provenientes, ou seja, se estão relacionadas as variáveis observadas (efeito composição) ou a fatores não observados (efeito coeficiente).

<sup>26</sup> É importante destacar que os resultados da Tabela 12 foram obtidos considerando a hipótese de dependência linear entre as covariadas e a RIF, ou seja, não foi utilizada a reponderação da RIF para se obter os coeficientes. Entretanto, no Apêndice B, são apresentados os resultados utilizando tal método, a partir de um modelo *logit* e verifica-se que os resultados são muito semelhantes.

Analisando os resultados da Tabela 12, constatou-se que o efeito composição é, de fato, a fonte responsável por uma diferença negativa entre os grupos. Considerando que o efeito composição mostra a diferença dos casos de DRSI em crianças da zona urbana perante aos casos da zona rural, devido a características específicas, tais resultados confirmam que grande parte da diferença dos casos dessas doenças, observados na zona rural são explicados pelas suas dotações (desiguais) perante o meio urbano. Além disso, essa desigualdade nas dotações parece ser mais problemática para os quantis superiores dos casos de DRSI, haja vista uma maior diferença entre os meios urbano e rural nesses quantis. Isso evidencia que, quanto maior for os casos no público analisado, mais relevante se torna as características específicas para atenuar os problemas de saúde causados por essas doenças na saúde infantil.

Os resultados para o efeito coeficiente também foram significativos. Isso mostra que existe um diferencial nos casos de DRSI em crianças apenas pelo fato dessas residirem na zona rural. A importância desse efeito vai aumentando ao longo dos quantis e, após o quantil 75, começa a apresentar uma redução. Além disso, o sinal positivo neste efeito mostra que se as zonas urbana e rural estivessem sob as mesmas condições, isto é, com características iguais, o retorno dos coeficientes seria maior para a zona urbana, o que tornaria a diferença entre os grupos menos negativa, indicando um aumento nos casos no meio urbano. Em outras palavras, a diferença entre os grupos se reduz não pela melhora das condições de saneamento no meio rural, mas sim pela piora dessas condições no meio urbano.

Para verificar quais são as principais características que provocam essas diferenças, a Tabela 13 apresenta as parcelas de contribuição nos efeitos coeficiente e composição de todas as variáveis para os quantis analisados.

Tabela 13. Efeito composição e coeficiente das características municipais sobre os casos de DRSI por mil habitantes em crianças de até 5 anos, quantis selecionados.

	q10	q25	q50	q75	q90
<b>Efeito composição (Parte I)</b>					
Condição de água (insegura)	0,0152*** (0,0027)	0,0419*** (0,0054)	0,0915*** (0,0124)	0,246*** (0,0290)	0,550*** (0,0760)
Condição de esgoto (céu aberto)	0,0186*** (0,0018)	0,0291*** (0,0025)	0,0470*** (0,0060)	0,0272** (0,0080)	0,0581** (0,0221)
Condição de esgoto (fossa)	-0,143*** (0,0096)	-0,289*** (0,0152)	-0,735*** (0,0314)	-1,205*** (0,0783)	-1,428*** (0,167)
DRSI <sub>t-1</sub>	0,00005 (0,0001)	-0,00032 (0,00031)	0,00078 (0,00065)	-0,00246 (0,00163)	0,159 (0,230)
Plano de saúde	-0,0638*** (0,0057)	-0,140*** (0,0099)	-0,323*** (0,0169)	-0,484*** (0,0267)	-0,632*** (0,063)
Taxa de analfabetismo	-0,0656*** (0,0052)	-0,155** (0,086)	-0,426*** (0,0179)	-0,631*** (0,0476)	-0,558*** (0,101)
Moradias	-0,0215*** (0,00264)	-0,0504*** (0,00394)	-0,190*** (0,0105)	-0,551*** (0,0274)	-1,179*** (0,0843)
<b>Efeito coeficiente (Parte II)</b>					
Condição de água (insegura)	0,0041* (0,0017)	0,0094*** (0,0026)	0,0300*** (0,0069)	0,0502** (0,0173)	0,0691 (0,0492)
Condição de esgoto (céu aberto)	-0,0662 (0,0636)	-0,0577 (0,0314)	0,1780* (0,0723)	0,810*** (0,134)	2,474*** (0,301)
Condição de esgoto (fossa)	-0,00009 (0,0056)	-0,0122 (0,0075)	0,0390* (0,0177)	0,258*** (0,0386)	0,984*** (0,0891)
DRSI <sub>t-1</sub>	-0,0002 (0,0006)	0,0012 (0,0011)	0,0031 (0,0020)	0,0062 (0,0053)	0,0183 (0,0114)
Plano de saúde	0,0699** (0,0107)	0,1400*** (0,0175)	0,295*** (0,0279)	0,421*** (0,0710)	0,709*** (0,126)
Taxa de analfabetismo	0,0628*** (0,0114)	0,0631** (0,0175)	0,0672 (0,0398)	0,3770*** (0,104)	1,289*** (0,262)
Moradias	-0,0572*** (0,0048)	-0,0684*** (0,00608)	-0,219*** (0,0162)	-0,479*** (0,0464)	-0,8430*** (0,117)

Obs: \* p<0,05; \*\* p<0,01; \*\*\* p<0,001. Erros padrão obtidos pelo procedimento *bootstrap*.

Fonte: Dados da pesquisa.

Analisando os resultados do efeito composição (parte I) na Tabela 13, constatou-se que todas as variáveis foram significativas em todos os quantis da distribuição, com exceção da variável DRSI defasada. Tal fato promove evidências que o aumento de casos de DRSI em crianças no meio rural é provocado pela desigualdade de acesso aos serviços de saneamento, tendo em vista que há mais casos dessas doenças na zona rural porque, em geral, são nessas regiões em que a população infantil está mais exposta a condições inseguras de água e esgoto a céu aberto, além de se encontrar uma população com menor nível de escolaridade, conforme demonstrado na análise descritiva deste trabalho. Além disso, o hiato dos casos de DRSI foi reduzido em todos os quantis devido as características referentes a condição de esgoto (fossa), plano de saúde e moradias inadequadas.

Embora já fosse esperado que a condição de esgoto a céu aberto não pudesse reduzir os diferenciais de casos de DRSI, é importante destacar o resultado ligado à condição de água

insegura. Conforme já destacado, tradicionalmente, em regiões rurais, o abastecimento alternativo é incentivado, inclusive, pela iniciativa pública devido à dificuldade de sustentabilidade financeira em se executar uma rede de abastecimento nessas regiões. A partir desses resultados, fica evidente que políticas públicas que fomentam o abastecimento de água via caminhões pipa, cisternas, entre outras soluções, embora seja relevante para levar água para áreas rurais, ainda estão longe de serem ações que forneçam uma fonte segura de água para essa população.

Ainda em relação aos resultados da parte I da Tabela 13, duas observações merecem atenção. A primeira é a possibilidade da condição de esgoto (fossa) reduzir os diferenciais de casos de DRSI. Esse resultado pode estar relacionado às condições de esgoto a céu aberto das zonas rurais. Embora o percentual de domicílios com fossa na zona rural seja superior, o fato de poder reduzir o esgotamento a céu aberto com esse tipo de dotação (fossa) contribuiria para que o hiato entre os meios urbano e rural diminuísse, apesar de não reduzir a desigualdade haja vista que a redução da desigualdade de acesso a tais serviços se dá por meio da ampliação da rede pública coletora de esgoto, algo ainda bastante distante para o meio rural.

Uma forma de atenuar as desigualdades no acesso a uma rede coletora de esgoto seria o investimento em fossas biodigestoras. Conforme Costa e Guilhoto (2014) destacam, existem diversos benefícios em se implementar tais fossas que são ecologicamente corretas. Além de reduzir o impacto ambiental e danos a saúde, a geração de renda e emprego poderia ser potencializada através de apoio a esses tipos de projetos.

Além disso, considerando um cenário mais precário, como é o caso das condições de saneamento avaliadas por Ausburg e Rodriguez-Lesmes (2018), na Índia, o resultado obtido para a condição de esgoto (fossa) torna-se semelhante. Regiões em que há ausência de qualquer tipo de coleta dos resíduos domiciliares, a fossa, ainda que rudimentar, pode ser uma alternativa barata e viável capaz de reduzir, ao menos em parte, os casos de DRSI para a população infantil dessas localidades. No Brasil, conforme Saiani e Oliveira (2018) discutem, ainda existem muitas regiões em que as desigualdades de acesso aos serviços de saneamento são grandes e, em função disso, é razoável apontar que a construção de fossas ainda seja relevante para diversas áreas rurais brasileiras no combate as DRSI.

A segunda observação se refere ao fato das moradias, ditas inadequadas, reduzirem o hiato entre os casos de DRSI. Dentro dessa classificação, existe uma hierarquia dos tipos de casas, que vai de “menos ruins” até uma condição totalmente precária. Casas construídas com material aproveitado, como papelão, lona e madeiras são consideradas os piores tipos de moradias, pois não garantem a devida salubridade ao ambiente. Esse tipo de moradia é mais

comum nas regiões periféricas e nas favelas das zonas urbanas do que no meio rural. Por outro lado, nas áreas rurais, as casas feitas com taipa, pilão e barro formam um tipo de moradia que fornece condições salubres ruins, porém melhores que os demais tipos. Assim, por ser um tipo de construção comum nas áreas rurais, é possível que a não divisão dessas categorias de moradias possa ter influenciado os resultados encontrados.

Na parte II da Tabela 13, é apresentada a decomposição dos efeitos não explicados. Diferente do que foi observado na parte I, há uma grande oscilação de sinais e significância estatística entre os quantis, indicando que existe uma heterogeneidade no efeito coeficiente. De uma maneira geral, percebe-se que, para os quantis inferiores até o terceiro quartil, o retorno das condições de água insegura é positivo e significativo, indicando que as diferenças entre os casos de DRSI em crianças, entre o meio urbano e rural, reduziria. Entretanto, essa redução se dá pelo fato do aumento dos casos na zona urbana que, em condições iguais à zona rural, o retorno positivo desse coeficiente desfavorece (mais prejudicial) a zona urbana.

Os resultados para as variáveis de esgotamento sanitário também seguem a mesma lógica anterior, porém, sua significância é observada a partir da mediana da distribuição. Por ser considerado o cenário desfavorável, as condições de esgoto a céu aberto foi a que apresentou o maior retorno para as diferenças de casos de DRSI em crianças. Novamente, neste caso, o maior retorno indica um aumento de casos na zona urbana quando esta parte das mesmas condições que o meio rural. Assim o *gap* é reduzido pela aproximação dos casos de DRSI e não pela redução em ambos os grupos.

Outro resultado relevante para se destacar é o efeito retorno dos planos de saúde sobre os casos de DRSI em crianças. Segundo os resultados da parte II da Tabela 13, o retorno dessa variável é maior (aumenta os casos) para o meio urbano. Apesar de chamar atenção, esse resultado parece ser lógico se for considerado o panorama de saúde no qual as áreas urbanas e rurais se encontram. O Brasil é um país de desigualdades, inclusive no acesso à saúde. Se ambas as regiões estiverem com as mesmas condições, isto é, sob as mesmas características, ao acessar os serviços de saúde via planos particulares, mais diagnósticos de DRSI seriam constatados para as áreas urbanas, tendo em vista a maior população de crianças para essa região que, se expostas às características existentes no meio rural, adoeceriam mais rapidamente em função da dinâmica social existente no meio urbano.

## 6. CONSIDERAÇÕES FINAIS

No Brasil, a falta de acesso aos serviços de saneamento básico acarreta prejuízos à saúde pública, ao meio ambiente e reduz o bem-estar social. O abastecimento da população com água tratada, os serviços de coleta e tratamento de esgoto apresentam grandes desafios para serem superados, rumo à universalização dos serviços. Enquanto não se atinge a universalização, o déficit desses serviços continua afetando a saúde dos mais frágeis, sobretudo as crianças menores de 5 anos de vida, tornando-se imprescindível compreender como que a prestação desses serviços pode influenciar o surgimento de DRSI nesse público. Foi sob essa perspectiva que o presente estudo teve como objetivo avaliar o impacto da eficiência de custo dos serviços de saneamento básico na saúde infantil e, em segundo momento, os efeitos da desigualdade de acesso, entre as áreas urbanas e rurais, nas internações e casos de DRSI em crianças de 0 e 5 anos de idade nos municípios brasileiros.

A partir da construção de uma fronteira estocástica de eficiência e de uma abordagem espacial, foi possível identificar que o *status* de saúde infantil das regiões onde a prestação de serviços de água e esgoto são eficientes, em termos de custo, são melhores do que naquelas localidades onde há menor eficiência. Foi verificada a existência de efeitos de transbordamentos para a eficiência variável, ou seja, o prestador eficiente além de contribuir para reduzir as internações de crianças por DRSI no próprio município, também afeta os vizinhos, diminuindo o número de hospitalizações dessas regiões. Isso revela que as normas, o ambiente regulatório e a gestão da firma, componentes da eficiência variável, são relevantes para combater o aumento das hospitalizações por DRSI em crianças.

Esse resultado deve ser observado pelos agentes reguladores e formuladores das normativas do saneamento, pois, não basta que sejam determinados indicadores de desempenho e outras metas para as firmas que operam nesse setor. É necessário que a regulação forneça condições para os operadores atingirem tais metas em vistas de proporcionarem externalidades positivas para a sociedade, tal como a redução das DRSI. A implementação de tarifas sociais, já prevista na Lei Federal 11.445/2007, é uma das formas de se fazer chegar o saneamento para quem mais precisa, isto é, aquelas famílias com maior vulnerabilidade socioeconômica. É papel da regulação cobrar dos operadores que essa medida saia do papel e comece a ser praticada com mais frequência.

O contexto histórico do setor de saneamento brasileiro é marcado por grandes desigualdades de acesso a esses serviços, privilegiando algumas localidades e gerando déficits entre regiões brasileiras para a maioria dos municípios. Foi observando esse contexto, que

também foi objetivo deste estudo verificar o efeito da desigualdade de acesso ao saneamento, entre os meios urbano e rural, por meio das diferenças de dotação e retorno, sobre os casos de DRSI em crianças com idade inferior a 5 anos, no âmbito municipal.

As evidências encontradas apontaram que existem efeitos diferentes para os quantis da distribuição dos casos de DRSI em crianças. Tanto para a zona urbana quanto para o meio rural, foram observados efeitos maiores das variáveis explicativas nos quantis superiores, ou seja, quanto pior a situação dos municípios, em termos de casos por mil habitantes, mais influente se torna o meio no qual as crianças estão inseridas. Esse resultado deixa claro que não basta melhorias nas condições de água e esgotamento. É necessário que, ao fornecer melhores serviços de água e esgoto, também exista boas políticas que combatam o analfabetismo, garanta acesso aos serviços básicos de saúde e ainda proporcione infraestrutura adequada para a população. Um cenário como esse ainda é uma realidade muito distante para o meio rural e extremamente complexo se for considerada as regiões periféricas das zonas urbanas.

Os resultados da decomposição quantílica apontaram que as diferenças de casos de DRSI em crianças podem ser atribuídas as características das zonas rurais. O percentual de condições de água insegura, esgoto a céu aberto e baixos níveis de escolaridade no meio rural são responsáveis pelo aumento dos casos por mil habitantes nessa região. Entretanto, se as dotações referentes ao esgotamento por fossa, plano de saúde e moradias aumentarem, o hiato entre os casos pode reduzir o efeito composição, isto é, diminuir a diferença de casos entre as zonas urbana e rural.

Considerando essas características como potenciais contribuidoras para o estado de saúde infantil no meio rural, o papel do governo é atuar, na medida do possível, em cada área visando proporcionar a melhoria dessas condições. Projetos populares que tenham como objetivo a criação de fossas ecologicamente corretas poderiam ser apoiados via recursos para popularizar esse tipo de saneamento no meio rural. Nessa mesma perspectiva, os projetos de moradias sustentável também deveriam ser incluídos. Uma complexidade estaria ligada aos planos de saúde. Em um setor que depende, essencialmente, de lucros para se sustentar, a atuação do poder público é muito limitada.

A discussão sobre saúde infantil e saneamento, apesar de ser amplamente conhecida, ainda está longe de terminar. Embora o presente trabalho tenha contribuído para a temática ao mostrar os efeitos de transbordamentos da eficiência do saneamento sobre a saúde infantil, evidenciando também as principais características que contribuem para reduzir/aumentar as

diferenças de casos de DRSI entre o meio urbano e rural, ainda há espaço para aprofundar essas e outras discussões.

A discussão apresentada sobre eficiência no saneamento e saúde infantil se fez importante, pois trouxe luz para uma temática que sempre era abordada na literatura apenas de forma secundária. Entretanto, após a constatação da influência da eficiência sobre a saúde infantil, é necessário que pesquisas futuras avancem sobre qual seria o melhor o modelo para se construir a fronteira de eficiência, incluindo outras variáveis como aquelas ligadas a qualidade e continuidade dos serviços. Para tanto, a base de dados sobre saneamento no Brasil precisa melhorar, como já vem ocorrendo, ao longo do tempo e, em um dado momento, será possível conseguir informações confiáveis e periódicas para a maioria dos municípios, proporcionando a elaboração de painéis espaciais que sejam capazes de acompanhar o efeito que os transbordamentos espaciais podem ter ao longo de cada ano.

A decomposição quantílica aplicada conseguiu detalhar diversos efeitos que as variáveis explicativas têm sobre as diferenças de casos de DRSI para as zonas urbana e rural ao longo de sua distribuição. Essa decomposição é algo novo para os estudos que tratam sobre a desigualdade de acesso aos serviços de saneamento e, portanto, tem caráter exploratório. Assim como a eficiência possui efeitos espaciais sobre a saúde infantil, é provável que existam efeitos espaciais distintos (ou não) nos meios urbano e rural, o que pode refinar os resultados para as diferenças encontradas e direcionar política públicas específicas para cada região a partir da observação dos efeitos espaciais de cada variável explicativa. Desse modo, pesquisas futuras que realizem uma análise quantílica espacial podem corroborar ou refutar os resultados deste estudo, contribuindo para a literatura de microeconomia do desenvolvimento.

## REFERÊNCIAS

- AIGNER, D.; LOVELL, C. K.; SCHMIDT, P. Formulation and estimation of stochastic frontier production function models. **Journal of Econometrics**, v. 6, n. 1, p. 21-37, 1977.
- ALMEIDA, E. **Econometria espacial aplicada**. São Paulo: Ed alínea, 2012.
- ALMOND, D., AND B. A. MAZUMDER. Health capital and the prenatal environment: The effect of Ramadan observance during pregnancy. **American Economic Journal: Applied Economics**, v. 3, n. 4, p. 56–85, 2011.
- ALMOND, D., J. CURRIE, AND M. HERRMANN. From infant to mother: Early disease environment and future maternal health. **Labour Economics**, v. 19, n. 4, p. 475–83, 2012.
- ANSELIN, L. **Spatial Econometrics: Methods and Models**. Boston: Kluwer Academic, 1988.
- \_\_\_\_\_. Spatial externalities, spatial multipliers, and spatial econometrics. **International Regional Science Review**, v. 26, n. 2, p. 153-166, 2003.
- ARRUDA, N. M.; MAIA, A. G.; ALVES, L. C. Desigualdade socioeconômicas no acesso à saúde entre as áreas urbanas e rurais do Brasil: uma decomposição. **Cad. Saúde Pública**, v. 34, n.6, p. 3-14, 2018
- AUSBURG, B.; RODRIGUEZ-LESMES, P. A. Sanitation and child health in India. **World Development**, v. 107, p. 22-39, 2018.
- BARRETO, M.; GENSER, B.; STRINA, A.; TEIXEIRA, M.; ASSIS, A.; REGO, R.; TELES, C.; PRADO, M.; MATOS, S.; ALCÂNTARA-NEVES, N.; CAIRNCROSS, S. Impact of a citywide sanitation program in Northeast Brazil on intestinal parasites infection in Young children. **Environ. Health Persp.**, v. 118, n. 11, p. 1637–1642. 2010.
- BLINDER, A. S. Wage discrimination: reduced form and structural estimates. **J. Hum. Resources**, v. 8, p. 436-455, 1973.
- BORJA, P. C. Política pública de saneamento básico: uma análise da recente experiência brasileira. **Saúde e Sociedade**, v. 23, n. 2, p. 432-447, 2014.
- BRASIL. **Lei n. 11.445, de 05 de janeiro de 2007**. Estabelece diretrizes nacionais para o saneamento básico. 2007. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/ato2007-2010/2007/lei/l11445.htm](http://www.planalto.gov.br/ccivil_03/ato2007-2010/2007/lei/l11445.htm)>. Acesso em 06 jan. 2020.
- CESAR, J. A.; GONÇALVES, T. S.; NEUMANN, N. A.; FILHO, J. A.; DIZIEKANIAK, A. Saúde infantil em áreas pobres das regiões Norte e Nordeste do Brasil: comparando indicadores básicos em áreas atendidas pela pastoral da criança. **Cad. Saúde Pública**, v. 21, n. 6, p. 1845-1855, 2005.
- CHEN, Y. Y.; SCHMIDT, P.; WANG, H.-J. Consistent Estimation of the Fixed Effects Stochastic Frontier Model, **Journal of Econometrics**, v. 18, n. 2, 65-76, 2014.

COELLI, T. J.; RAO, D. S. P.; O'DONNELL, C. J.; BATTESE, G. E. **An introduction to efficiency and productivity analysis**. Springer Science & Business Media, 2005.

COLOMBI, R.; MARTINI, G.; VITTADINI, G. **A Stochastic Frontier Model with Short-Run and Long-Run Inefficiency Random Effects**. Department of Economics and Technology Management, University of Bergamo, Working Paper Series, 2011.

CURRIE, J. **Child health in developed countries**. In A. J. Culyer & J. P. Newhouse (Eds.), *Handbook of health economics*, 1B (pp. 1053–1090). Amsterdam: Elsevier. 2000.

DA HORA, A. L.; SHIMODA, E.; DA HORA, H. M.; COSTA, H. G. Análise da eficiência dos serviços de saneamento básico nos municípios do estado do Rio de Janeiro. **Pesquisa Operacional para o Desenvolvimento**, v. 7, n. 1, p. 55-81. 2015.

DA SILVA e SOUZA, G.; FARIA, R. C.; MOREIRA, T. B. Estimating the relative efficiency of brazilian publicly and privately owned water utilities: a stochastic cost frontier approach. **Journal of the American water resources association**, v. 43, n. 5, p.1237-1245, 2007.

FERREIRA, M. A. M. **Eficiência Técnica e de Escala de cooperativas e sociedade de capital na indústria de laticínios do Brasil**. Tese (Doutorado em Economia Aplicada). Viçosa, Minas Gerais, UFV/DER, 2005.

FERREIRA, P. S.; MOTTA, P.C.; SOUZA, T.C.; SILVA, T.P.; OLIVEIRA, J.F.; SANTOS, A.S.P. Avaliação preliminar dos efeitos da ineficiência dos serviços de saneamento na saúde pública brasileira. **Revista internacional de Ciências**, v. 6, n. 2, p. 214-229, 2016.

FERRO, G.; LENTINI, E.; MERCADIER, A. C.; ROMERO, C. A. E. Efficiency in Brazil's water and sanitation sector and its relationship with regional provision, property and the independence of operators. **Utilities Policy**, v. 28, n. 2, p. 42-51, 2014.

FERRO, G.; MERCADIER, A. C. Technical efficiency in Chile's water and sanitation providers. **Utilities Policy**, v. 46, p. 97-109, 2016.

FERRO, G.; ROMERO, C. A.; CASTIGLIONE, I. Efficiency in saving infant lives: the influence of water and sanitation coverage. **Working paper hal-00612956**, 2011. Disponível em: <<http://hal.archives-ouvertes.fr/hal-00612956>>. Acesso em 12 jan. 2020.

FEWTRELL, Lorna et al. **Water, sanitation and hygiene: quantifying the health impact at national and local levels in countries with incomplete water supply and sanitation coverage**. Environmental Burden of Diseases Series. 2007.

FIRPO, S.; FORTIN, N.M.; LEMIEUX, T. Unconditional Quantile Regressions. **Econometrica**, v. 77, n. 3, p. 953-973, 2009.

\_\_\_\_\_. Decomposing wage distributions using recentred influence function regressions. **Econometrics**, v. 6, n. 2, p. 28-68, 2018.

FLORAX, R.; ANSELIN, L.; REY, S. **Advances in spatial econometrics: Methodology, tools and applications**. Springer Science & Business Media, 2013.

FLORAX, R.; 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.

FONSECA, F. R.; VASCONCELOS, C. H. Análise espacial das Doenças Relacionadas ao Saneamento Ambiental Inadequado no Brasil. **Cad. Saúde Colet**, v. 19 n. 4, p. 448-453. 2017.

GARCIA, S.; THOMAS, A. The structure of municipal water supply costs: application to a panel of French local communities. **Journal of Productivity analysis**, v. 16, n. 1, p. 5-29, 2001.

GROSS, R.; SCHELL, B.; MOLINA, M. C.; LEÃO, M. A.; STRACK, U. The impact of improvement of water supply and sanitation facilities on diarrhea and intestinal parasites: A brazilian experience with children in two low-income urban communities. **Rev. Saúde pública**, v. 23, n. 3, p. 214-220. 1989.

GROSSMAN, M. On the concept of health capital and the demand for health. **Journal of Political economy**, v. 80, n. 2, p. 223-255, 1972.

HUTTON, G.; RODRIGUEZ, U.; WINARA, W.; ANH, N.; PHYRUM K.; CHUAN, L.; BLACKETT, I.; WEITZ, A. Economic efficiency of sanitation interventions in southeast Asia. **Journal of Water, Sanitation and Hygiene for Development**, v. 4, n. 1, p. 23-36, 2014.

INSTITUTO TRATA BRASIL. **Benefícios econômicos e sociais da expansão do saneamento no Brasil. 2018**. Disponível em: <<http://www.tratabrasil.org.br/estudos/estudos-itb/beneficios-economicos-e-sociais>>. Acesso em: 06 jan. 2020.

\_\_\_\_\_. **Principais Estatísticas: Planos de Saneamento**. Disponível em: <<http://www.tratabrasil.org.br/saneamento/principais-estatisticas/no-brasil/planos-de-saneamento>>. Acesso em: 09 jan. 2020.

ISUNJU, J. B.; SCHWARTZ, K.; SCHOUTEN, M. A.; JOHNSON W. P.; VAN DIJK, M. P. Socio-economic aspects of improved sanitation in slums: A review. **Public Health**, v. 125, n. 6, p. 368-376, 2011.

JEULAND, M. A.; FUENTE, D.; OZDEMIR, S.; ALLAIRE, M. C.; WHITTINGTON, D. The long-term dynamics of mortality benefits from improved water and sanitation in less developed countries. **PLoS ONE**, v. 8, n. 10, p. 1-16, 2013.

JUSTINO, D. C. P.; ANDRADE, F. B. Análise espacial das causas de mortalidade infantil no Brasil de 2000 a 2015. *Revista Ciência Plural*, v. 6, n. 3, p. 174-193, 2020.

KOENKER, R.; BASSET, G. Regression Quantiles. **Econometrica: Journal of the Econometric Society**, p. 33-50, 1978.

KUMBHAKAR, S. C.; WANG, H.; HORNCastle, A. P. **A practitioner's guide to stochastic frontier analysis using Stata**. Cambridge University Press, 2015.

KUMBHAKAR, S.; LIEN, G.; HARDAKER, J. Technical efficiency in competing panel data models: a study of Norwegian grain farming. **Journal of Productivity Analysis**, v. 41, n. 2, p. 321-337, 2014.

LOPES, T. G.; MORAES, L. R. S. **Importância da compatibilização entre os tipos de prestação de serviço público de saneamento rural, a matriz tecnológica e o modo de vida camponês**. IN: Engenharia Hidráulica e Sanitária, cap. 4. Atena Editora, Ponta Grossa-PR. 2019.

LU, M. C., TACHE V.; ALEXANDER, G. R.; KOTELCHUCK M. Preventing low birth weight: is prenatal care the answer? **Journal of Maternal and Fetal Neonatal Medicine**, v. 13, p. 362–380, 2003.

MARCONATO, M.; DAL MORO, O. F.; PARRÉ, J. L.; FRAVO, J. Uma análise espacial sobre a saúde nos municípios brasileiros em 2010. **Revista de Economia e Agronegócio**, v. 18, n. 1, p. 1-26, 2020.

MATTOS, E.; PINTO, C.; TEIXEIRA, L.; MELONI, L. Sanitation and health: empirical evidence for Brazilian municipalities. **Brazilian Review of Econometrics**, v. 39, n. 2, p. 269-302, 2019.

MEDEIROS, R. V. V.; COSTA, J. G. A.; CARDOSO, L. C. B. O efeito das UPAs na taxa de internações por condições sensíveis à atenção primária. **Estudos Econômicos**, v. 51, n. 4, p. 677-698, 2021.

MONTGOMERY, M. A.; ELIMELECH, M. Water and Sanitation in Developing Countries: Including Health in the Equation. **Environmental Science & Technology**, January, p. 17-24, 2007.

MURILLO-ZAMORANO, L. R. Economica efficiency and frontier techniques. **Journal of economic surveys**, v. 18, n. 1, p. 33-77, 2004.

OAXACA, R. L.; Male-female wage differentials in urban labor markets. **Int. Econ. Rev.**, v. 14, p. 693-709, 1973.

OHIRA, T. H.; SHIROTA, R. Eficiência econômica: uma aplicação do modelo de fronteira estocástica em empresas de saneamento. **In: Anais do XXXIII Encontro Nacional de Economia**, p. 1-11, 2005.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. The 2022 Revision of World Population Prospect. Disponível em: <<https://population.un.org/wpp/>>. Acesso em 01 jul. 2022.

PEARSON, J.; MCPHEDRAN, K. A literature review of the non-health impacts of sanitation. **Waterlines**, v. 27, n. 1, p. 48-61, 2008.

PONTES, R. P. **Impactos do planejamento e da regulação no desempenho das empresas do setor de saneamento**. Tese (Doutorado em Economia Aplicada). Viçosa, Minas Gerais, UFV/DER, 2019.

REICHMAN, N. E.; CORMAN, H.; NOONAN, K.; DAVE, D. Infant health production functions: what a difference the data make. **Health Economics**, v. 18, n. 7, p. 761-782, 2009.

ROCHA, A. C. T.; ROSSONI, H. A. V.; FARIA, M. T. Determinantes envolvidos no perfil de doenças relacionadas às condições sanitárias inadequadas nos municípios brasileiros: avaliação realizada nas 10 melhores e 10 piores cidades do ranking do saneamento. **ForSci IFMG**, v. 6, n. 2, p. 2-22, 2018.

RODRIGUES, K. C. T.; VENSON, A. H.; CAMARA, M. R. Distribuição espacial do acesso aos serviços de saneamento básico nas microrregiões brasileiras de 2006 a 2013. **Revista Brasileira de Gestão e Desenvolvimento Regional**, v. 15, n. 1, p. 137-151. 2019.

SABBIONI, G. Efficiency in the Brazilian sanitation sector. **Utilities Policy**, v. 16, n. 2, p. 11-20, 2008.

SAIANI, C. S.; OLIVEIRA, W. T. Desigualdades de acesso a serviços públicos de saneamento básico: Evidências de seletividade das políticas e efeitos do crescimento econômico. In: **Anpec – Encontro Nacional da Associação Nacional dos Centros de Pós-Graduação em Economia**, 2018.

SAIANI, C.; AZEVEDO, P. F. Is privatization of sanitation services good for health? **Utilities Policy**, v. 52, n. 1, p. 27-36, 2018.

SAIANI, C.; TONETO JR., R. Déficit de acesso aos serviços de Saneamento Básico no Brasil. **IN: XXXIV Encontro Nacional de Economia**, 2006.

SANTOS, A. M. A.; TEJADA, C. A. O.; EWERLING, F. Os determinantes socioeconômicos do Estado de Saúde das crianças do Brasil rural. **Revista de Economia e Sociologia Rural**, v. 50, n. 3, p. 473-492, 2012.

SCRIPTORE, J. S. **Impactos do saneamento sobre saúde e educação: uma análise espacial**. Tese (Doutorado em Ciências). São Paulo, SP. USP/FEA, 2016.

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

SILVA, K. R.; LEIVAS, P. H. S.; SANTOS, A. M. A.; HALMENSCHLAGER, V. Saneamento básico e saúde no Brasil: uma análise via painel espacial. **IN: XXIII Encontro Regional de Economia ANPEC-SUL, 2020**. Disponível em: <<http://www.anpec.org.br/novosite/br/xxiii-encontro-de-economia-da-regiao-sul--artigos-selecionados>>. Acesso em: 22 fev. 2021.

SILVEIRA, A. B. G. Explorando o déficit em saneamento no Brasil: evidências da disparidade urbano-rural. **Paranoá**, v. 10, n. 1, p. 37-48, 2013.

TEIXEIRA, J. C.; HELLER, L.; BARRETO, M. Giardia duodenalis infection: risk factors for children living in sub-standard settlements in Brazil. **Cad. Saúde Pública**, v. 23, n. 6, p. 1489-1493. 2007.

TEIXEIRA, J. C.; OLIVEIRA, G. S.; VIALI, A. M.; MUNIZ, S. S. Estudo do impacto das deficiências de saneamento básico sobre a saúde pública no Brasil no período de 2001 a

2013. **Engenharia Sanitária e Ambiental**, Rio de Janeiro, v. 19, n. 1, p. 87-96, jan./mar. 2014.

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. **RACEF**, v. 7, n. 2, p. 01-16, 2016.

VAN MINH, H.; HUNG, N. V. Economic aspects of sanitation in developing countries. **Environmental health insights**, v. 5, p. EHI. S8199, 2011.

VENSON, A. H.; JACINTO, P. A.; SBICCA, A. Acesso ao esgotamento sanitário das famílias pobres no Brasil. **Revista de Desenvolvimento Econômico**, v. 44, n. 3, p. 199-229, 2019.

WANG, H.-J.; HO, C. W. Estimating Fixed-Effect Panel Stochastic Frontier Models by Model Transformation. **Journal of Econometrics**, v. 157, 286-296, 2010.

WHITTINGTON, D.; HANEMANN, W. M.; SADOFF, C.; JEULAND, M. The Challenge of improving water and sanitation services in less developed countries. **Foudantions and Trends in Microeconomics**, v. 4, p. 469-609, 2007.

## APÊNDICE A

Para diagnosticar se o viés da amostra causa um atrito grande o suficiente no painel para gerar problemas nos resultados, pode-se utilizar o procedimento de Nijman e Verbeek (1992). Neste teste, segundo os autores, é estimado um modelo em painel com efeitos fixos que consiste em incluir na regressão variáveis *dummies* que visam capturar se a perda de informação é totalmente aleatória ou não. Adaptando a abordagem de Nijman e Verbeek (1992) para o contexto do saneamento, tem-se o seguinte modelo para estimação da fronteira estocástica de custo:

$$\ln CV_{ijt} = \beta_0 + \beta_1 \ln P_{Lijt} + \beta_2 \ln K_{ijt} + \beta_3 \ln Q_{ijt} + \alpha \sum_{i=1}^{27} UF + \phi \sum_{t=2008}^{2018} ANO + \eta_1 s_{i,t-1} + \eta_2 s_{i,t-2} + \varepsilon_{ijt} \quad (a.1)$$

Os termos  $s_{i,t-1}$  e  $s_{i,t-2}$  são variáveis binárias que assume valores iguais 1 quando a operadora divulga suas informações, isto é, aparece na amostra, nos período  $t - 1$  e  $t - 2$ , respectivamente. Caso a ausência das informações seja completamente aleatória, os parâmetros  $\eta_1$  e  $\eta_2$  não serão significativos e, desse modo, não há indícios que apontem que os dados *missing* causem viés de seleção. Por outro lado, se pelo menos um dos parâmetros apresentar significância, então o atrito no painel é grande o suficiente para levar a um viés de seleção e, portanto, uma correção deve ser realizada. A seguir, Tabela A1 apresenta os resultados para  $s_{i,t-1}$  e  $s_{i,t-2}$ .

Tabela A1. Resultado do teste de atrito de Nijman e Verbeek (1992).

	Ln custo variável
$\ln Q_{ijt}$	0,1606*** (0,0047)
$\ln K_{ijt}$	0,4073*** (0,0079)
$\ln P_{Lijt}$	-0,0081* (0,0042)
$\ln CMM$	0,0296*** (0,0031)
Dummy UF	Sim
Dummy Ano	Sim
$S_{i,t-1}$	-0,0718*** (0,0165)
$S_{i,t-2}$	0,2110*** (0,0087)
Constante	11,6901*** (0,6369)
Observações	15784
R <sup>2</sup>	0,078

Obs: \*\*\* p<0,01; \*\* p<0,05; \* p<0,10. Erros padrão entre parênteses.

Fonte: Elaboração própria a partir de dados da pesquisa.

A partir dos resultados observados na Tabela A1 fica constatado que os termos  $S_{i,t-1}$  e  $S_{i,t-2}$  foram significativos e, portanto, a ausência de informação não é aleatória e pode provocar um viés de seleção na amostra das firmas que operam saneamento.

Para contornar o problema do viés causado pelo atrito, o primeiro passo é estimar os fatores que se relacionam com a probabilidade das firmas reportarem informações em cada período de tempo. Nesta etapa, o procedimento aplicado foi um modelo *probit* em cada ano da amostra. Os resultados são apresentados na Tabela A2. Como o período base é 2007, optou-se por não utilizar informações anteriores a este ano para determinar a probabilidade das operadoras reportarem informações em 2007, tendo em vista que nos anos anteriores não são contemplados pelo marco regulatório.

Tabela A2. Probabilidade das firmas reportarem informações na base de dados do SNIS.

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
$\ln Q_{ijt}(t-1)$	-0,0733 (0,0688)	-0,1759*** (0,0531)	-0,0725 (0,2258)	-0,0033 (0,5736)	-0,1779 (0,5604)	-0,1447 (0,1956)	-0,1572 (0,2046)	0,0214 (0,3097)	-0,2620 (0,2902)	-0,1041*** (0,0387)	-0,0812** (0,0319)
$\ln K_{ijt}(t-1)$	0,0907 (0,0903)	0,2557*** (0,0677)	0,1309 (0,2826)	0,0598 (0,0806)	-0,1439 (0,6934)	0,1121 (0,2551)	0,1316 (0,2653)	-0,0854 (0,4380)	0,2538 (0,3640)	0,1736*** (0,0501)	0,1478*** (0,0422)
$\ln P_{Lijt}(t-1)$	-0,1918*** (0,0681)	-0,0096 (0,0530)	-0,2861 (0,2166)	-0,2026*** (0,0614)	-0,4039 (0,4982)	0,0345 (0,1765)	0,0242 (0,1794)	0,0561 (0,3485)	-0,0334 (0,2440)	-0,1573*** (0,0368)	-0,1571*** (0,0312)
$\ln CMM(t-1)$	-0,1519*** (0,0554)	-0,1622*** (0,0463)	-0,4682** (0,2043)	-0,2658*** (0,0534)	-0,0864 (0,3727)	0,1013 (0,0823)	0,0864 (0,0864)	0,0154 (0,2660)	0,0851 (0,1216)	-0,2705*** (0,0301)	-0,2538*** (0,0255)
$\ln Q_{ijt}(t-2)$	-	0,1010*** (0,035)	0,1247*** (0,0342)	0,1349*** (0,0334)	-0,1663 (0,0464)	0,1414*** (0,0450)	0,1662*** (0,0434)	0,3216*** (0,0505)	-0,0429 (0,0503)	-0,1649*** (0,0584)	-0,2111*** (0,0612)
$\ln K_{ijt}(t-2)$	-	-0,1068** (0,0497)	-0,1390*** (0,0490)	-0,1713*** (0,0480)	0,1849*** (0,0620)	0,1599*** (0,0606)	0,2201*** (0,0588)	0,4224*** (0,0674)	0,1422** (0,0694)	0,2382*** (0,0778)	0,2950*** (0,0790)
$\ln P_{Lijt}(t-2)$	-	-0,0281*** (0,0431)	-0,3234*** (0,0453)	-0,1534*** (0,0391)	0,0356 (0,0512)	0,1205** (0,0495)	0,0557 (0,0447)	0,0788 (0,0541)	0,0258 (0,0553)	0,1151** (0,0529)	0,0493 (0,0590)
$\ln CMM(t-2)$	-	-0,1638*** (0,0344)	-0,1453*** (0,0346)	-0,0641** (0,0317)	0,0530 (0,0395)	0,0625 (0,0414)	0,0402 (0,0381)	0,1356*** (0,0402)	0,0609* (0,0366)	0,1064*** (0,0318)	0,0893** (0,0394)
Constante	3,0702*** (0,4546)	2,2818*** (0,2941)	2,2217*** (0,2996)	1,4199*** (0,2612)	1,2489*** (0,3003)	0,7512** (0,3300)	0,9499*** (0,3344)	0,8382** (0,3254)	0,8876*** (0,3355)	0,6707** (0,3364)	1,211*** (0,3584)
Observações	1.260	1.359	1.507	1.527	1.627	1.645	1.706	1.714	1.714	1.766	1.811
Pseudo R <sup>2</sup>	0,0599	0,0316	0,0362	0,0179	0,0131	0,0143	0,0157	0,0571	0,0152	0,0312	0,0287

Nota: \*\*\* p<0,01; \*\* p<0,05; \* p<0,10. Erros padrão entre parênteses.

Fonte: Elaboração própria a partir de dados da pesquisa

Como foi possível notar na Tabela A2, as informações reportadas em  $t - 2$  são as que mais afetam a probabilidade das firmas reportarem seus dados tendo em vista que nesse períodos os coeficientes apresentaram mais significância.

Além disso, também se observa que os sinais alternam ao longo dos anos, indicando que não há uma variável específica que negativamente ou positivamente afete a probabilidade de reportar as informações. Esse comportamento é justificável por, pelo menos, duas razões. O primeiro está ligado a decisão arbitrária das firmas observarem seus resultados e, caso não alcancem o que foi planejado, podem optar por informar a variável que obteve os resultados indesejados. O segundo ponto está ligado as decisões políticas que podem permear o setor de saneamento. Em muitos municípios existem os Serviços Autônomos de Água e Esgoto (SAAEs) que, embora sejam autarquias, acabam sendo influenciadas por decisões políticas locais. Em muitos casos, os custos com energia elétrica, funcionários, entre outros custos, são subsidiados pelas prefeituras e, em contrapartida, os SAAEs retardam os ajustes tarifários.

## APÊNDICE B

Tabela B.1 Matriz de correlações entre as RIF's estimadas e as covaridas utilizadas no estudo.

	RIF10	RIF25	RIF50	RIF75	RIF90	Plano de saúde	Condição de esgoto (fossa)	Condição de esgoto (céu aberto)	Tx de analfabetismo	Moradias
<b>RIF10</b>	1,0000									
<b>RIF25</b>	0,4981***	1,000								
<b>RIF50</b>	0,2907***	0,5837***	1,000							
<b>RIF75</b>	0,1661***	0,3335***	0,5714***	1,000						
<b>RIF90</b>	0,0961***	0,1930***	0,3307***	0,5788***	1,0000					
<b>Plano de saúde</b>	-0,0572**	-0,1145**	-0,1326**	-0,1012***	-0,0550**	1,0000				
<b>Condição de esgoto (fossa)</b>	0,0744***	0,1347***	0,1668***	0,1559***	0,0990***	-0,3145**	1,0000			
<b>Condição de esgoto (céu aberto)</b>	0,0105***	-0,0027	-0,0103**	-0,0062	-0,0092***	-0,0979***	-0,2825***	1,0000		
<b>Condição de água (insegura)</b>	0,0713***	0,1203***	0,1428***	0,1286***	0,0917***	-0,2382***	0,5281***	-0,0980***		
<b>Tx de analfabetismo</b>	0,1074***	0,1732***	0,1978***	0,1649***	0,1016***	-0,3289***	0,4505***	-0,1092***	1,0000	
<b>Moradias</b>	-0,0181**	-0,0035	0,0297***	0,0723***	0,0726***	-0,0437**	0,1349***	0,2877***	-0,1658***	1,0000

Nota: \*\*\* p<0,01; \*\* p<0,05; \* p<0,10.

Fonte: Elaboração própria a partir de dados da pesquisa.

Tabela B.2. Efeito da Zona rural em cada quantil de distribuição de casos de DRSI considerando o modelo ponderado.

	q10	q25	q50	q75	q90
Grupo 1	0,327*** (0,0041)	0,786*** (0,0058)	2,131*** (0,0132)	5,009*** (0,0274)	9,775*** (0,0670)
Contrafactual	0,342*** (0,0046)	0,705*** (0,0059)	1,697*** (0,0126)	4,224*** (0,0383)	9,062*** (0,0931)
Grupo 2	0,473*** (0,0063)	1,101*** (0,0098)	2,888*** (0,0228)	6,823*** (0,0479)	13,35*** (0,108)
Diferença	-0,146*** (0,0075)	-0,315*** (0,0115)	-0,758*** (0,0264)	-1,814*** (0,0552)	-3,571*** (0,127)
Explicada (Efeito Composição)	-0,131*** (0,0078)	-0,396*** (0,0115)	-1,192*** (0,0261)	-2,599*** (0,0613)	-4,283*** (0,143)
Não Explicada (Efeito Coeficiente)	0,0150* (0,0062)	0,0813*** (0,00834)	0,434*** (0,0183)	0,785*** (0,0471)	0,712*** (0,115)
Constante	-0,285*** (0,0198)	-0,435*** (0,0261)	-0,847*** (0,0565)	-1,689*** (0,148)	-4,005*** (0,364)
Erro de especificação	0,0190 (0,119)	0,0306 (0,1770)	0,0633 (0,4020)	0,6000 (0,8910)	0,1720 (0,207)
Erro de ponderação	0,0029*** (0,0010)	0,0005 (0,0014)	0,0008* (0,0034)	0,0061** (0,0009)	0,0121 (0,1990)

Nota: Grupo 1 = Zona urbana; Grupo 2 = Zona rural. Erros padrão obtidos por procedimento *bootstrap*.

Obs: \* p<0,05; \*\* p<0,01; \*\*\* p<0,001.

Fonte: Dados da pesquisa.

Tabela B.3. Efeito composição e coeficiente das covariadas nos casos de DRSI com o modelo ponderado.

	q10	q25	q50	q75	q90
<b>Efeito composição (Parte I)</b>					
Condição de água (insegura)	0,0160*** (0,0039)	0,0438*** (0,0060)	0,0973*** (0,0139)	0,260*** (0,0294)	0,597*** (0,0676)
Condição de esgoto (céu aberto)	0,0028*** (0,0009)	0,0045** (0,0015)	0,0072*** (0,0025)	-0,0038 (0,0020)	0,0103* (0,0050)
Condição de esgoto (fossa)	-0,138*** (0,0091)	-0,2810*** (0,0141)	-0,7130*** (0,0323)	-1,163*** (0,0680)	-1,362*** (0,1560)
Plano de saúde	-0,0724*** (0,0050)	-0,1570*** (0,0078)	-0,3620*** (0,0179)	-0,538*** (0,0374)	-0,693*** (0,0850)
Taxa de analfabetismo	0,0700*** (0,0053)	0,1650*** (0,0082)	0,4550*** (0,0190)	0,680*** (0,0399)	0,6110*** (0,0910)
Moradias	-0,0214*** (0,0028)	-0,0504*** (0,0043)	-0,190*** (0,0102)	-0,5540*** (0,0223)	-1,202*** (0,0509)
<b>Efeito coeficiente (Parte II)</b>					
Condição de água (insegura)	0,00069* (0,0029)	0,0069 (0,0038)	0,0230** (0,0083)	0,1260*** (0,0221)	0,270*** (0,0544)
Condição de esgoto (céu aberto)	0,1010*** (0,0107)	0,1880*** (0,0141)	0,5170*** (0,0304)	0,9250*** (0,0808)	1,824*** (0,1990)
Condição de esgoto (fossa)	-0,0739*** (0,0051)	-0,1000*** (0,0067)	-0,3000*** (0,0145)	-0,6380*** (0,0383)	-1,470*** (0,0941)
Plano de saúde	-0,0009 (0,0511)	0,0476*** (0,0068)	0,1280*** (0,0149)	-0,0261 (0,0369)	-0,0391 (0,0909)
Taxa de analfabetismo	0,1270*** (0,0115)	0,2790*** (0,0152)	0,5750*** (0,0331)	0,8620*** (0,0852)	1,354*** (0,210)
Moradias	-0,0274*** (0,0052)	0,00132 (0,0068)	-0,0312* (0,0149)	-0,1400*** (0,0388)	-0,258*** (0,0954)

Obs: \*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$ . Erros padrão obtidos pelo procedimento *bootstrap*.  
Fonte: Dados da pesquisa.

Os resultados da Tabela B.3 se assemelham ao que foi apresentado na Tabela 11 do estudo, indicando poucas variações em termos de significância e direção dos efeitos. O que é relevante destacar é que, na Tabela B.2, é possível verificar a possibilidade de erros de especificação e ponderação do modelo utilizado. Conforme Rios-Avila (2020) descreve, a implementação de um modelo em que a RIF computada seja ponderada por uma matriz de pesos pode auxiliar no ajuste do modelo original (principal). Caso o modelo considerando a matriz de pesos possua significância estatística para o erro de especificação estimado, então implica que o modelo principal possui problemas em sua especificação e, desse modo, a ponderação deve ser utilizada para melhorar o ajuste do modelo. Observando os resultados da Tabela B.2 verifica-se a ausência de significância para os erros de especificação, o que é indicativo que o modelo principal possui um ajuste adequado. Tal fato é reforçado ao ser observar matriz de correlações apresentada, que aponta para a existência de dependência linear.

Além disso, na Tabela B.2 também há o resultado para o erro de ponderação. Caso o modelo ponderado foi utilizado para realizar inferência sobre os efeitos das covariadas, seria necessário observar se a ponderação aplicada, neste caso via modelo *logit*, estava adequada. Segundo Rios-Avila (2020), quando os erros de ponderação tendem a zero e não significativos é possível dizer que não erros de ponderação no modelo especificado. Observando os resultados da Tabela B.2 nota-se que, embora em alguns quantis seja detectada significância estatística, os valores para o erro de ponderação estão muito próximos de zero, o que significa que a ponderação aplicada possui erros pequenos que não devem afetar com grande magnitude os resultados da decomposição apresentada.