OS IMPACTOS DO SANEAMENTO BÁSICO SOBRE A EDUCAÇÃO: USANDO A PRIVATIZAÇÃO COMO VARIÁVEL INSTRUMENTAL

Juliana Souza Scriptore
FEA-USP & UNIFAL-MG
(juliana.scriptore@gmail.com).

Carlos Roberto Azzoni
FEA-USP
(cazzoni@usp.br)

Naércio Aquino Menezes Filho
Insper e FEA-USP
(NaercioAMF@insper.edu.br)

Resumo

Este artigo investiga o efeito do saneamento sobre educação através do método de variáveis instrumentais, usando o número de empresas que foram privatizadas há mais de 5 anos no município como instrumento para as condições de saneamento nos domicílios. São combinadas as bases de dados do Censo Demográfico (IBGE), Censo Escolar (INEP) e Sistema Nacional de Informações Sobre Saneamento (SNIS) para os anos 2000 e 2010, fazendo uso de econometria espacial e não-espacial para estimação desse efeito. Os resultados indicam que o efeito do aumento de uma unidade percentual no acesso a saneamento está associado a um aumento de 0,11 pontos percentuais na taxa de frequência escolar, a uma queda de 0,31 pontos percentuais na taxa de distorção idade-série e a também a uma queda de 0,12 pontos percentuais na taxa de abandono escolar.

Palavras-chave: Dados em Painel, Econometria Espacial, Saneamento Básico, Educação.
JEL: O18; I18; C21

Abstract

This study the effect of sanitation on education, in other words, the educational consequences for children whose health has been affected by Diseases Related to Inadequate Environmental Sanitation (DRIES). In this sense, infrastructure conditions of dwellings, expressed by the variable of access to basic sanitation, for impacting the health of individuals, is one of the determinants of educational indicators. We applied data at municipality level, from the Population Census (IBGE), the School Census (INEP), the National Information System on Sanitation (SNIS), among others, for the years 2000 and 2010, with the goal of evaluating the effect of sanitation on education made use of spatial and non-spatial econometrics techniques for estimating the effects. The results of the first step indicate that the effect of one percentage increase in access to sanitation is associated with an increase of 0.11 percentage points in school attendance rate, a decrease of 0.31 percentage points in age-grade distortion rate, and also a decrease of 0.12 percentage points in the school dropout rate.

Key words: Panel Data, Spatial Econometrics, Sanitation, Education.
JEL: O18; I18; C21

1 Introdução

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

As precárias condições de saneamento básico propiciam a transmissão de bactérias, vírus e parasitas, que estão presentes nas fezes, urina ou vômito do doente ou portador, causadores de diversas doenças infectocontagiosas. A diarreia é a mais conhecida dentre elas, pelos seus efeitos devastadores sobre a taxa de mortalidade infantil em diversos países do mundo. No entanto, existem diversas outras que também merecem preocupação quanto aos danos causados à população, em especial às crianças: esquistossomose, hepatite A, febre amarela, leishmaniose, malária, febre tifoide, entre outras.

A grande importância em diminuir as taxas de incidência de tais doenças levou a Fundação Nacional da Saúde (Funasa, 2004) a classifica-las em cinco grupos. O primeiro é o de transmissão feco-oral, o segundo é composto por doenças transmitidas por inseto-vetor, o terceiro contém as doenças transmitidas via contato com a água, o quarto diz respeito às doenças relacionadas a higiene e, o último, denomina-se o grupo dos geo-helmintos e teníases. O conjunto de doenças de todos esses grupos foi denominado de Doenças Relacionadas ao Saneamento Ambiental Inadequado (DRSAI) baseado na proposta de classificação das doenças de Cairncross e Feachem (1990; 1993) e Mara e Feachem (1999).

O objetivo desse estudo, a partir do reconhecimento milenar desta associação entre saneamento e saúde, é investigar as consequências educacionais sobre crianças que apresentam estado de saúde debilitado por terem contraído as DRSAI. A hipótese estabelecida é que os sintomas de tais doenças lhes proporcionam um estado de saúde insuficiente para alcançar bons resultados escolares. Segundo Prado *et al.* (2001), as infecções parasitárias estão entre as mais disseminadas desordens que afetam crianças em idade escolar que vivem em áreas pobres de grandes centros urbanos. Os efeitos patológicos diretos destes parasitas são diarreia crônica, má absorção dos alimentos e anemia ferropriva, que é decorrente da privação e deficiência de ferro dentro do

organismo (KUNZ et al., 2008). As consequências dessas desordens tendem a comprometer o comportamento e a função cognitiva das crianças, principalmente quanto à capacidade de atenção e concentração, que afetam o rendimento escolar e dificultam o aprendizado (ARAÚJO et al, 2009 e KUNZ et al., 2008).

Em outras palavras, quando a incidência das DRSAI não leva a população infantil a óbito, quais são os impactos gerados sobre a acumulação de capital humano dos sobreviventes em termos de perda de rendimento escolar?

O objetivo dessa pergunta é ressaltar que, além de destruir capital humano, os efeitos de adversas condições de saneamento podem impactar o perfil de morbidade de uma região no longo prazo, com graves consequências para a aquisição de anos de escolaridade e, dessa forma, comprometimento de ganhos salariais e de produtividade no futuro. Tais efeitos deletérios da falta dessa infraestrutura essencial propagam as sementes do subdesenvolvimento social e educacional, pois não haverá políticas de educação que sejam capazes de alterar tais resultados escolares se não acompanhadas de políticas de infraestrutura.

As condições de infraestrutura dos domicílios, expressas pela variável de acesso a saneamento básico, podem ser mais um dos determinantes dos indicadores educacionais? Em outras palavras, o maior acesso aos serviços de saneamento básico nos domicílios impacta os indicadores educacionais de frequência, distorção idade-série e abandono escolar da população infantil? A hipótese sustentada nesse trabalho é que o efeito destacado acima se dá pelo estado de saúde debilitado dos indivíduos.

Segundo Relatório de Desenvolvimento Humano (2006), infecções parasitárias transmitidas pela água ou pelas más condições de saneamento provocam 443 milhões de faltas escolares por ano no mundo. Miguel e Kremer (2004) também encontraram um efeito associado ao motivo doença para explicar a evasão escolar. Nesse sentido, a pergunta que se pretende responder é se, de fato, as precárias condições de saneamento básico se traduzem em maior incidência de DRSAI. Além disso, esse impacto pode ser ampliado pelos efeitos de transbordamento espacial relacionados à variável de saúde dos vizinhos?

A justificativa da importância de tais perguntas pode ser vista sob três dimensões. O maior acesso aos serviços essenciais de saneamento, além de gerar impactos sobre a saúde, meio ambiente, turismo, entre outros, ocasiona uma nova externalidade de longo prazo por impactar indicadores educacionais de performance. Na área de educação, as condições precárias de saneamento onde a escola está localizada e onde os alunos residem

importam para o rendimento escolar dos mesmos, além do *background* familiar e do efeito escola. Na área da saúde, as consequências dos precários ou ausentes serviços de saneamento básico são maiores do que a literatura dessa área documenta, pois vão além dos impactos em termos de maiores taxas de mortalidade infantil, ou seja, afetam o perfil de morbidez das crianças sobreviventes gerando deletérios efeitos socioeconômicos de longo prazo para o país.

Além dessa introdução, na seção 2 foi apresentado o referencial teórico acerca do impacto de saneamento em educação. A terceira apresenta e descreve os dados municipais, obtidos para os anos 2000 e 2010, que são utilizados na investigação desse efeito. Ademais, apresenta as estatísticas descritivas de todos eles e as fontes nas quais tais dados foram obtidos. Na quarta seção são descritas as metodologias que foram utilizadas. Além da análise econométrica convencional para dados em painel com presença de endogeneidade, adotou-se também metodologia para o controle das características espaciais presentes nesse tipo de dados. A partir do referencial teórico dos determinantes educacionais, buscou-se identificar se as características domiciliares onde os indivíduos residem são importantes para explicar variáveis de frequência escolar, distorção idade-série e abandono escolar. Na mesma seção, os resultados (não espaciais e espaciais), obtidos com a aplicação de cada uma dessas metodologias, são apresentados.

Algumas questões específicas também foram investigadas tais como a sensibilidade de tais efeitos para a população mais jovem. Quanto mais novas as crianças que, em geral, são os organismos mais suscetíveis às DRSAI, maior o impacto na saúde e maior o efeito sobre piores indicadores educacionais? Por fim, a última seção conclui o trabalho.

2 Saneamento básico e educação

Antes de definir os canais de transmissão pelos quais o acesso aos serviços de saneamento básico pode afetar indicadores educacionais, é preciso descrever quais são os principais determinantes dos indicadores de educação considerados pela literatura. Diversos estudos tratam da relação positiva entre anos de escolaridade e salário dos indivíduos ou sucesso no mercado de trabalho (Card, 2001; Kassouf, 2001). Além disso, a educação gera diferenças no nível de produtividade das pessoas, que perduram durante todo seu ciclo de vida (Menezes-Filho, 2001). Nesse sentido, a distribuição da educação também explica boa parte da desigualdade da renda do país: Barros e Mendonça (1996)

argumentam que esta diminuiria em até 50% se a desigualdade educacional fosse eliminada. Os investimentos em educação, que geram notas positivas em testes de proficiência e aumentam a probabilidade de os alunos prosseguirem com os estudos, ocasionam crescimento econômico para um país, dada a importância do capital humano (BISHOP, 1989; TOPEL, 1999; HANUSHEK e KIMKO, 2000).

A importância da educação, tanto para o crescimento econômico de longo prazo quanto para obtenção de resultados individuais em termos de rendimentos salariais no futuro, propulsionou a investigação das várias dimensões que afetam os indicadores de rendimento escolar. Uma delas é referente ao *efeito-escola*, que se materializa nas características atribuídas ao ambiente físico em que ocorre o aprendizado e àquelas atribuídas aos profissionais que conduzem esse processo, os professores. Como ressaltado por Duflo (2001) e Menezes-Filho (2007), as variáveis relativas à infraestrutura da escola são importantes para explicar o aumento da taxa de escolaridade, seja por meio de construção de novas escolas ou pela melhora na infraestrutura física das mesmas. Nesse sentido, melhores instalações no interior das escolas, um quadro de pessoal formado por professores qualificados e instrumentos de gestão eficientes, tais como a presença de diretoria na escola e funcionamento adequado dos trâmites administrativos, são fatores relevantes dentro dessa dimensão. Dessa forma, estudos nacionais e internacionais investigaram formas de melhorar os resultados escolares por meio desse tipo de investimento[1].

A outra dimensão ressaltada pela literatura refere-se ao que se denomina de *background familiar*. Os estudos de Barros *et al* (2001), Vasconcellos (2003) e Menezes-Filho (2001) indicaram a importância das características familiares (principalmente a escolaridade dos pais) e dos fatores socioeconômicos dos alunos nos níveis de escolaridade dos mesmos. Tais autores apontam a família como principal determinante dos resultados educacionais dos filhos. Em relação a essas duas dimensões, Menezes-Filho (2007) destaca que as características relacionadas à escola têm efeitos importantes na explicação do desempenho escolar, porém reduzidos quando comparados às características familiares e socioeconômicas dos alunos.

[1] Para resultados da literatura internacional, ver Coleman (1966); Hanushek (1986); Card e Krueger (1992), Heckman, Layne-Farrar e Todd (1996); Rivkin, Hanushek e Kain (2005). Para o caso brasileiro, ver Ferrão et al (2001), Albernaz, Ferreira e Franco (2002); Macedo (2004).

No entanto, seja pela ação e influência dos pais ou pelo ambiente físico em que a escola funciona, o primeiro passo para elevar o nível médio de escolaridade de um país é elevar a frequência escolar e manter a criança na escola, garantindo-lhe o avanço de seus níveis educacionais (PONTILI e KASSOUF, 2007). Uma das explicações para a baixa frequência escolar, com consequente impacto em outros indicadores de rendimento escolar tais como abandono e distorção idade-série, é a saúde deficiente. De acordo com Banerjee e Duflo (2011), dentre as prováveis explicações para evasão escolar infantil em vários países do mundo, além da falta de exigência dos pais em obrigar as crianças a continuarem os estudos, está a falta de estímulo das mesmas em frequentarem as escolas. Essa falta de estímulo pode estar associada a um estado de saúde comprometido.

Miguel e Kremer (2004) encontraram um efeito associado ao motivo doença para explicar a evasão escolar. O estudo de tais autores indicou efeito positivo da medicação contra *helmintos* ("vermes") intestinais, alojados em crianças quenianas, sobre a frequência escolar. Os autores ressaltam a importância da adequada provisão de bens públicos de saneamento, uma vez que a transmissão de tais agentes patogênicos ocorre pela matéria fecal não tratada ou não disposta corretamente pelos mesmos. Ou seja, oferecer tais serviços é uma forma de impactar positivamente o desempenho das crianças na escola.

O estudo de Neri et al. (2008), ao fazer análises de correlações brutas entre variáveis de infraestrutura e desempenho escolar em 2001 e 2003, encontrou que o acesso à água apresenta correlação positiva com a proficiência escolar e negativa com o índice de reprovação. Neste último caso, segundo os autores, melhor infraestrutura, tanto na casa como na escola, deve melhorar o rendimento dos estudantes, reduzindo a repetência.

A literatura da área médica documenta as possíveis consequências de um estado de saúde debilitado decorrente das parasitoses intestinais e outras doenças que estão relacionadas à falta ou inadequada provisão dos serviços de saneamento básico (DRSAI). Segundo Prado *et al.* (2001), as infecções parasitárias estão entre as mais disseminadas desordens que afetam crianças em idade escolar que vivem em áreas pobres de grandes centros urbanos. Os efeitos patológicos diretos destes parasitas são diarreia crônica, má absorção dos alimentos e anemia ferropriva, que é decorrente da privação e deficiência de ferro dentro do organismo (KUNZ et al., 2008). Segundo ARAUJO *et al.* (2009), a presença de anemia associada a enteroparasitoses é resultante da subnutrição (*Ascaris lumbricóides*), da ação hematofágica (*Ancylostoma sp.*) e da ulceração das mucosas intestinais (*Entamoeba histolytica*). O indivíduo sofre constantes perdas sanguíneas, além

do agravamento do quadro patológico, a depender da carga parasitária, da idade, do estado nutricional e fisiológico do organismo, bem como da associação com outras espécies parasitárias patogênicas.

As consequências tendem a comprometer o comportamento e a função cognitiva das crianças, principalmente quanto à capacidade de atenção e concentração, que afetam o rendimento escolar e dificultam o aprendizado (ARAÚJO et al, 2009 e KUNZ et al., 2008). Outra consequência destacada por Brito *et al* (2003) é que a anemia tem efeitos negativos sobre o crescimento ponderal e estatural na infância e no período escolar (fases em que a necessidade nutricional é maior). Logo, o desenvolvimento físico e intelectual da população pediátrica é prejudicado (GURGEL et al., 2005; LUDWIG et al., 1999) e, nos adultos, o principal impacto da anemia é a diminuição da capacidade produtiva, que se reflete na menor capacidade de trabalho.

Quando os fatores acima citados estão associados às condições precárias de saneamento básico, em que a ocorrência de verminoses se torna comum, crianças e adolescentes ficam expostos a uma situação de desequilíbrio crônico (DOMENE, 2004). Outro agravante de tal desequilíbrio crônico deve-se à perda dos maiores retornos do investimento em educação no início da infância. Segundo Heckman (2005), todo tipo de ação voltada para a primeira infância do indivíduo constitui o melhor investimento social existente e quanto mais baixa for a idade, mais alto é o retorno do investimento.

Carneiro, Cunha e Heckman (2003) encontraram que nos Estados Unidos o retorno de investimento em educação é muito maior para idades mais precoces. Para o Brasil, Curi e Menezes-Filho (2009) apontam que a educação pré-primária, voltada à população de zero a seis anos de idade, tem relação positiva e significante com a conclusão dos ciclos escolares, com a escolaridade média e, consequentemente, com a renda do trabalho principal, indicando efeitos socioeconômicos positivos no futuro. Schady (2006) reforça tais conclusões, ao destacar que os prejuízos causados pelo desenvolvimento inadequado nas idades iniciais das crianças afetam não somente o desempenho escolar, mas também emprego, salário, criminalidade e medidas de integração social de adultos.

A última dimensão dos determinantes educacionais que será tratada nessa seção diz respeito aos possíveis efeitos espaciais presentes nas variáveis de educação. Ou seja, os municípios que apresentam altas taxas de rendimento escolar são capazes de gerar algum tipo de *spillover* espacial para os municípios vizinhos? Sabe-se que a literatura dessa área considera inúmeras variáveis não observadas presentes no termo de erro, tais

como habilidade dos alunos e dos professores e características de gestão educacional específicas das escolas e municípios. Caso exista autocorrelação das variáveis não observadas no espaço, esse efeito pode ser controlado ao considerar a estrutura espacial dos dados, de forma a fornecer estimativas mais consistentes dos parâmetros. Por outro lado, ignorar efeitos espaciais significativos pode proporcionar estimativas inconsistentes e viesadas dos parâmetros. Portanto, deve-se avaliar, assim como no estudo de Baylis et al. (2011), se a introdução de tais efeitos altera significativamente os resultados obtidos por meio da análise econométrica convencional.

Os efeitos da interação no espaço entre alunos, escolas e professores tem sido tema de estudo de alguns pesquisadores, principalmente estudos acerca da relação entre educação (e/ou sua distribuição) e renda (e/ou desigualdade de renda) em um país ou região. Rodriguez-Pose e Tselios (2008) apontam para uma relação positiva e robusta entre a desigualdade educacional e a desigualdade de renda entre as regiões da União Europeia. Umar *et al* (2014) encontraram efeitos positivos da distribuição equitativa da educação sobre o nível de renda regional nigeriano.

Na dimensão dos determinantes dos indicadores educacionais referentes ao *efeito-escola*, Milimet e Rangaprasad (2007) mostraram que os administradores de escolas públicas em Illinois, nos Estados Unidos, levam em consideração nas suas decisões acerca dos insumos escolares destinados a melhorar a qualidade educacional a escolha pelos mesmos insumos feita pelos administradores das escolas públicas dos distritos vizinhos. Ou seja, as decisões das escolas públicas relativas à razão aluno-professor, média salarial dos professores, gastos por aluno e tamanho da turma influenciam estrategicamente a escolha por tais insumos das escolas públicas vizinhas que se encontram no mesmo município. Plenzler (2004) também destaca a importância da escala espacial na análise da relação entre desempenho do aluno e recursos educacionais. Naidoo *et al* (2014), ao utilizar o método de regressão ponderada geograficamente, encontraram forte relação entre desempenho escolar e características socioeconômicas da localidade onde a escola está localizada na África do Sul. Dentro de tais características, ressaltam aquelas relativas às características dos pais e dos domicílios em que as famílias residem.

Nesse contexto, Borjas (1995) encontrou que o nível educacional e de rendimentos alcançados pelos jovens dos Estados Unidos eram impactados pela renda dos pais e também pela renda média dos vizinhos de mesma etnia da geração de seus pais. Os exemplos de interações espaciais demonstradas em estudos acima apontam a relevância

de se levar em conta os efeitos de um espaço geográfico mais abrangente do que o municipal para a determinação de políticas educacionais.

3 Dados

Os dados municipais utilizados nesse estudo são provenientes de diversas fontes de dados e compreendem os anos de 2000 e 2010. No modelo a ser estimado, que será apresentado na próxima seção, as variáveis dependentes são provenientes do Censo Demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE)[2] e do Censo Escolar do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). A variável explicativa de interesse, referente a saneamento, provém também do Censo Demográfico (IBGE). A variável instrumental, utilizada para identificar o impacto da falta de saneamento nos indicadores educacionais, foi construída a partir de informações obtidas pela ABCON (Associação Brasileira das Concessionárias Privadas de Serviços Públicos de Água e Esgoto)[3]. O SNIS (Sistema Nacional de Informações sobre Saneamento) foi consultado para confirmação das informações da ABCON, bem como para identificar outros dados dos municípios atendidos por cada empresa. O Quadro 3.1 apresenta tais variáveis, a descrição de cada uma delas, a unidade e a fonte dos dados.

[2] Dados obtidos via Atlas do Desenvolvimento Humano do Programa das Nações Unidas para o Desenvolvimento (PNUD).

[3] A ABCON congrega empresas privadas prestadoras de serviços públicos de água e saneamento básico bem como diversas informações acerca do tipo de contrato de concessão, data em que o mesmo passou a vigorar, município em que a empresa privada oferece os serviços de saneamento básico, população atendida, investimentos previstos e realizados entre outros. A partir de tais informações, foi possível determinar a partir de que ano a empresa privada de saneamento começou a atuar nos municípios em questão de forma a construir a variável instrumental.

Variáveis	Descrição	Unidade	Fonte
Variáveis Dependentes			
Frequência Escolar (**t_freq_6a14**)	População de 6 a 14 anos frequentando a escola em qualquer nível ou série	%	Censo Dem.
Abandono escolar (**abandono_fund**)	Alunos que deixaram de frequentar a escola, e tiveram a matrícula cancelada, nas duas fases do Ensino Fundamental (1ª a 8ª série ou 1º ao 9º ano)	%	Censo Escolar
Abandono escolar fase 1 (**abandono_fund_1**)	Alunos que deixaram de frequentar a escola, e tiveram a matrícula cancelada, na primeira fase do Ensino Fundamental (1ª a 4ª série ou 1º ao 5º ano)	%	Censo Escolar
Abandono escolar fase 2 (**abandono_fund_2**)	Alunos que deixaram de frequentar a escola, e tiveram a matrícula cancelada, na segunda fase do Ensino Fundamental (5ª a 8ª série ou 6º ao 9º ano)	%	Censo Escolar
Distorção Idade-Série (**tdi_fund**)	Discentes cuja idade está acima da adequada para a série que frequenta nas duas fases do Ensino Fundamental (1ª a 8ª série ou 1º ao 9º ano)	%	Censo Escolar
Distorção Idade-Série fase 1 (**tdi_fund_1**)	Discentes cuja idade está acima da adequada para a série que frequenta na primeira fase do Ensino Fundamental (1ª a 4ª série ou 1º ao 5º ano)	%	Censo Escolar
Distorção Idade-Série fase 2 (**tdi_fund_2**)	Discentes cuja idade está acima da adequada para a série que frequenta na segunda fase do Ensino Fundamental (5ª a 8ª série ou 6º ao 9º ano)	%	Censo Escolar
Variável de Saneamento			
Banheiro e água (**t_banagua**)	População que vive em domicílios com banheiro e água encanada	%	Censo Dem.
Variável Instrumental			
Privatização saneamento (**priv5anos**)	Indicadora se a empresa de saneamento atuante no município é privada há pelo menos cinco anos	0 ou 1	Abcon/SNIS

Quadro 3.1: Variáveis dependentes, variável instrumental e de saneamento.
Fonte: Elaboração própria a partir de dados da pesquisa.

As variáveis explicativas restantes utilizadas nas estimações foram obtidas das mesmas fontes de dados do Quadro 3.1, além de algumas novas tais como o IPEA (Instituto de Pesquisa Econômica Aplicada), Finanças do Brasil (FINBRA), da Secretaria do Tesouro (STN), e Departamento de Informática do Sistema Único de Saúde (DataSUS). Tais variáveis foram organizadas em três vetores, denominados características socioeconômicas (1), características da escola e professores (2) e saúde e educação (3). De forma semelhante ao Quadro 3.1, encontram-se, no Quadro 3.2, a apresentação de tais variáveis, a descrição de cada uma delas, a unidade e também a fonte dos dados de onde são provenientes.

Variáveis	Descrição	Unidade	Fonte
Características Socioeconômicas			
Escolaridade pop. adulta (**t_analf25m**)	População de 25 anos ou mais analfabeta	%	Censo Dem.
Etnia (**pop_negra_parda**)	População negra ou parda	%	Censo Dem.
Renda per capita (**renda**)	Renda por habitante média	R$/hab	Censo Dem.
Índice de gini (**gini**)	Grau de desigualdade existente na distribuição de indivíduos segundo a renda domiciliar per capita	0-1	Pnud
Distribuição etária população (**pop15a29 e pop55mais)**	População de 15 a 29 anos e população de 55 anos ou mais	%	Censo Dem.
População total (**poptot**)	População total (habitantes)	Hab	Censo Dem.
Bolsa Famíllia (**bolsafam**)	Número de famílias beneficiárias do Programa Bolsa Família em relação ao número total de famílias	%	IPEA
Taxa de urbanização (**urb**)	População urbana em relação à população total	%	Censo Dem.
Características da escola e professores			
Alunos_turma (**alunos_turma_fund**)	Número Médio de Alunos por Turma no ensino fundamental	Al/turma	Censo Escolar
Docente_Superior (**doc_sup_fund**)	Percentual de Docentes com Curso Superior no ensino fundamental	%	Censo Escolar
Média_horas_aula (**horas_aula_fund**)	Número Médio de Horas-Aula Diária no ensino fundamental	Hs/aula	Censo Escolar
Infra Tecnológica (**iiet**)	Indice de infraestrutura escolar tecnológica	0-1	Censo Escolar
Infra Básica (**iiba**)	Indice de infraestrutura escolar básica	0-1	Censo Escolar
Escolas públicas (**p_pub_esc**)	Porcentagem de escolas públicas	%	Censo Escolar
Diretoria (**p_diretoria**)	Porcentagem de escolas com existência de diretoria	%	Censo Escolar
Saúde e educação			
Desp_saude (**desp_saude**)	Despesa com saúde por habitante	R$/hab	Finbra
Desp_ed_cult (**desp_educ_cult_hab**)	Despesa com educação e cultura por habitante	R$/hab	Finbra
Imunização (**doses_calculos_cv**)	Doses de vacinas aplicadas	Doses	DataSUS

Quadro 3.2: Variáveis independentes.
Fonte: Elaboração própria a partir de dados da pesquisa.

O primeiro vetor de observáveis engloba variáveis demográficas, de renda, urbanização e escolaridade da população adulta. O segundo refere-se às características dos professores e da escola. O último apresenta informações de despesa municipal com saúde, educação e cultura, além da variável relativa às doses de vacinas aplicadas na população de crianças (imunização). Essas variáveis, mais a variável de saneamento (denominada variável de interesse), compõe o vetor de variáveis explicativas das estimações com dados municipais.

3.1 Estatísticas descritivas dos dados municipais

A partir da compatibilização dessas diversas fontes de dados por município e para os anos 2000 e 2010 foram calculadas as estatísticas descritivas que serão apresentadas nesta seção. A tabela 3.3 exibe informações sobre as principais variáveis de educação utilizadas nesse estudo, para os anos 2000 e 2010. O primeiro e o segundo ciclo correspondem, respectivamente, da primeira à quarta série ou do primeiro ao quinto ano, e da quinta à oitava série, ou do quinto ao nono ano. Além disso, foram apresentadas as

mesmas estatísticas descritivas para as variáveis de acesso aos serviços de saneamento básico e da variável que será utilizada como instrumento. Todas as variáveis apresentaram evolução positiva ao longo década, com destaque para a redução da taxa de abandono escolar, particularmente no primeiro ciclo. Contudo, o avanço na universalização do acesso não pode ter eliminado as deficiências do sistema educacional, que se refletem na baixa qualidade do ensino e na deficitária aprendizagem dos alunos.

Tabela 3.1: Estatísticas Descritivas – dependentes, saneamento e instrumental

Variáveis	Obs		Média		Desvio Padrão		Minimo		Máximo	
Períodos	**2000**	**2010**	**2000**	**2010**	**2000**	**2010**	**2000**	**2010**	**2000**	**2010**
Variáveis Dependentes										
t_freq6a14	4369	4369	92,85	97,22	6,12	2,54	43,68	51,77	100	100
abandono_fund	4314	4369	11,08	2,84	7,14	2,93	0	0	53,40	34,10
abandono_fund_1	4313	4369	8,35	1,31	7,93	2,16	0	0	59,90	25,40
abandono_fund_2	4312	4369	14,10	4,67	8,12	4,51	0	0	100	45,10
tdi_fund	4314	4275	38,39	23,17	18,23	11,63	6,00	1,30	87,00	66,10
tdi_fund_1	4314	4267	29,92	16,93	19,70	11,23	0,90	0,40	86,70	64,40
tdi_fund_2	4313	4275	48,86	30,35	20,07	13,70	8,70	1,40	96,60	79,40
Variável de Saneamento										
t_banagua	4369	4369	71,62	86,73	27,18	18,01	1,26	6,68	100	100
Variável instrumental										
priv5anos	4369	4369	0,002	0,04	0,05	0,20	0	0	1	1

Fonte: Elaboração própria a partir de dados da pesquisa.

A Tabela 3.2 apresenta as estatísticas descritivas para as variáveis independentes dos vetores de características socioeconômicas, da escola e professores e da saúde e educação. Os dados da tabela abaixo mostram que significativas mudanças demográficas ocorreram no período: queda da população cuja faixa etária demanda atendimento da educação básica e superior (de 0 a 14 anos e de 15 a 29 anos) e, por outro lado, envelhecimento populacional, caracterizado pelo aumento da quantidade de pessoas com 55 anos ou mais. Além disso, o país tornou-se mais urbano, o valor da renda real aumentou e a mesma distribui-se de forma menos desigual (menor coeficiente de Gini). As iniciativas de transferência direta de renda às camadas da população abaixo da linha de pobreza, o Programa Bolsa Família (PBF), podem ter contribuído para esse cenário.

Houve aumento real de recursos financeiros destinados à educação e à saúde, evidência que pode ser constatada pelo aumento real do valor das despesas nessas áreas. Ademais, a redução da população na faixa escolar significou mais recursos para o atendimento de crianças, jovens e adultos. Registrou-se aumento da porcentagem de

docentes com nível superior, do índice de infraestrutura escolar tecnológica e básica, da porcentagem de escolas que têm diretoria e ligeiro aumento do número médio de horas-aula diárias no ensino fundamental. Por outro lado, a razão alunos por turma diminuiu, juntamente com a porcentagem de escolas públicas. Como consequência dessa evolução, além de outros fatores específicos relacionados às políticas públicas específicas, os indicadores educacionais apresentaram evolução positiva.

Tabela 3.2: Estatísticas Descritivas das variáveis independentes

Variáveis	Obs		Média		Desvio Padrão		Minimo		Máximo	
Períodos	**2000**	**2010**	**2000**	**2010**	**2000**	**2010**	**2000**	**2010**	**2000**	**2010**
Características socioeconômicas										
t_analf25m	4369	4369	22,13	16,67	13,15	10,72	0,95	1,10	69,38	54,57
pop_negra_parda	4312	4369	41,20	46,93	25,10	24,23	0	0,25	94,81	93,43
rendapc	4369	4369	391,69	556,40	184,25	234,37	89,8	141,88	1759,76	2043,74
gini	4369	4369	0,54	0,48	0,07	0,07	0,30	0,28	0,87	0,80
pop55mais	4315	4369	13,09	17,01	3,27	4,30	2,70	4,30	29,60	38,30
pop15a29	4315	4369	26,85	25,54	2,27	2,61	16,20	14,90	34,90	43,80
pop0a14	4315	4369	30,49	24,36	5,15	4,76	17,30	7,30	52,80	45,60
bolsafam	4369	4369	0	19,51	0	13,77	0	0,08	0	100
poptot	4315	4369	32695	36077	204357	220636	795	805	10400000	11300000
t_urb	4315	4369	61,46	66,45	23,49	22,09	0	4,18	100	100
Características da escola e professores										
alunos_turma_fund	4314	4276	31,73	21,77	9,06	3,84	12,50	6,50	276,30	37,70
doc_sup_fund	3978	4188	42,49	75,97	27,98	21,94	0	0	100	100
horas_aula_fund	4314	4275	4,31	4,43	0,32	0,50	3,60	3,70	6,80	9,20
iiet	4315	4369	0,09	0,21	0,10	0,13	0	0,0001	0,67	0,73
iiba	4315	4369	0,47	0,50	0,19	0,17	0,03	0,02	1	0,98
p_pub_esc	4315	4369	91,54	89,13	11,11	12,50	32,86	25,70	100	100
p_diretoria	4315	4369	32,79	53,38	26,56	26,68	0	0	100	100
Saúde e educação										
desp_educ_cult_hab	4203	4323	187,20	234,44	329,12	95,48	0	0	17754,50	1383,97
desp_saude_hab	4155	4196	97,346	475,788	136,48	256,49	0	0	6096,55	4549,625
doses_calculos_cv	4316	4369	12720	13512,8	75568	77356,8	275	259	3967753	3986617

Fonte: Elaboração própria a partir de dados da pesquisa.

Os dados municipais descritos acima foram utilizados com análise econométrica convencional: estimação em dois estágios de um modelo de efeitos fixos para dados em painel com variável instrumental. A abordagem espacial utilizou os mesmos dados, porém perdeu algumas informações, pelo fato de que o painel espacial utilizado para estimação do mesmo modelo deveria ser balanceado, sem *missings* e sem "ilhas"[4]. A Tabela 3.3 mostra a evolução das variáveis dependentes utilizadas nas estimações econométricas espaciais. Os resultados apontam para a mesma tendência de evolução dos dados das seções anteriores. Os valores para a variável de frequência escolar de 6 a 14

[4] As "ilhas" são definidas como localidades que não tem nenhum vizinho.

anos, de acesso aos serviços de saneamento básico adequado e o instrumento relativo às empresas que privatizaram os serviços de saneamento apresentaram aumento entre os anos de 2000 e 2010. Os valores relativos à taxa de distorção idade-série e ao abandono escolar mostraram queda no mesmo período.

Tabela 3.3: Estatísticas Descritivas – variáveis dependentes e outras.

Variáveis	Obs	Média		Desvio Padrão		Mínimo		Máximo	
		2000	2010	2000	2010	2000	2010	2000	2010
t_freq_6a14	3612	92,18	96,66	5,96	2,94	43,68	51,77	100	100
abandono_fund	3612	10,65	2,81	7,03	2,94	0	0	53,40	34,10
tdi_fund	3612	37,87	23,01	18,34	11,60	6,00	1,30	82,70	66,10
t_banagua	3612	72,07	86,83	27,15	17,75	1,56	6,68	100	100
priv5anos	3612	0,002	0,04	0,05	0,21	0	0	1	1

Fonte: Elaboração própria a partir de dados da pesquisa.

Os valores das estatísticas descritivas das variáveis explicativas utilizadas no painel espacial, demonstrados na Tabela 3.4, também estão de acordo com os resultados anteriores.

Tabela 3.4: Estatísticas Descritivas – variáveis independentes.

Variáveis	Obs	Média		Desvio Padrão		Mínimo		Máximo	
		2000	2010	2000	2010	2000	2010	2000	2010
t_analf25m	3612	22,19	16,74	13,58	11,10	0,95	1,10	69,38	54,57
pop_negra_parda	3612	3.469	4.869	20.854	28.377	0	2	811.090	1.113.998
renda per capita	3612	396,81	559,55	190,83	240,90	98,43	141,88	1759,76	2043,74
gini	3612	0,54	0,48	0,069	0,066	0,30	0,28	0,87	0,80
pop55mais	3612	13,20	17,05	3,31	4,24	2,70	4,8	29,60	34,00
pop15a29	3612	26,82	25,60	2,30	2,58	16,20	14,9	34,90	43,80
bolsafam	3612	0	19,15	0	13,31	0	0,08	0	100
poptot	3612	35.952	40.002	222.300	241.247	795	805	10.400.000	11.300.000
t_urb	3612	61,75	67,14	23,83	22,21	0	4,18	100	100
alunos_turma_fund	3612	31,73	21,89	8,94	3,82	12,50	7,80	276,30	37,70
doc_sup_fund	3612	44,88	76,38	27,61	21,68	0	0	100	100
horas_aula_fund	3612	4,32	4,44	0,33	0,50	3,60	3,70	6,80	9,20
iiet	3612	0,10	0,21	0,10	0,13	0	0,0001	0,67	0,73
iiba	3612	0,48	0,50	0,20	0,18	0,025	0,022	1,00	0,98
p_pub_esc	3612	91,40	88,87	11,46	12,92	32,86	25,70	100	100
p_diretoria	3612	33,28	52,95	26,83	26,60	0	0	100	100
desp_educ_cult_hab	3612	189,11	234,72	353,05	93,35	1,64	0	17.755	1.384
desp_saude_hab	3612	88,51	181,43	51,38	79,49	10,26	42,96	613,34	849,06
doses_calculos_cv	3612	13.911	14.906	82.153	84.522	275	259	3.967.753	3.986.617

Fonte: Elaboração própria a partir de dados da pesquisa.

De acordo com os dados apresentados acima, novamente tem-se um retrato de um país com mais pessoas idosas, menor população adulta em termos percentuais, mais

urbano, com maior renda per capita distribuída de forma menos concentrada ao longo do período e com mais beneficiários do programa bolsa família. No que se refere ao vetor de características relacionadas à educação, houve aumento do número de docentes com ensino superior, da quantidade média de horas por aula, maior porcentagem de escolas que têm diretoria e maiores índices de infraestrutura escolar básica e tecnológica. Por outro lado, o número de escolas públicas diminuiu assim como a razão do número de alunos por turma. Em relação às variáveis de despesa todas registraram maiores recursos destinados à saúde, educação e cultura. Além disso, a cobertura das taxas de imunização também aumentou ao longo desse período.

4 Metodologia Econométrica – Variáveis Instrumentais

O modelo geral para dados em painel é representado da seguinte forma:

$$(4.1) \quad EDUC_{it} = \beta_{0it} + \beta_{1it}SAN_{1it} + \ ... + \beta_{kit}X_{kit} + a_i + \varepsilon_{it}$$

O subscrito *i* indica município (*i = 1, ..., 5.565*), o subscrito *t* denota o ano (*t = 2.000 e 2.010*); $EDUC_{it}$ são os indicadores de educação (taxa de frequência escolar, de distorção idade-série e de abandono escolar), X_{kit} é o valor da *k-ésima* variável explicativa para o município *i* no instante *t* (*k = 1, ..., 19*); os β_{ki} são os parâmetros a serem estimados (β_0 é o intercepto da equação); o termo a_i capta todos os fatores não observados dos municípios, constantes no tempo, que afetam $EDUC_{it}$; ε_{it} é o termo de erro para o *i-ésimo* município em *t*.

Para que esse modelo seja estimado consistentemente é necessário estabelecer hipóteses acerca do termo a_i. Neste estudo, considera-se que a estimação dos indicadores educacionais é condicional aos efeitos específicos dos municípios, ou seja, os a_i são tratados como parâmetros a serem estimados. Além disso, permite-se que esse termo seja correlacionado com as variáveis explicativas: $E(a_iX_{it}) \neq 0$. É razoável supor que existam características não-observáveis dos municípios que também influenciem os indicadores educacionais e que podem estar correlacionadas com as variáveis explicativas, ou seja, características comunitárias (entre elas, o acesso aos serviços de saneamento), características socioeconômicas das famílias e os insumos escolares. A

estimação do modelo de efeitos fixos para dados em painel necessita de três hipóteses. A primeira delas é a necessidade de exogeneidade forte ou estrita:

(4.2) $E(\ddot{X}_{it}\ddot{\varepsilon}_{it}) = 0$, em que $\ddot{X}_{it} = X_{it} - \overline{X}_i$ e $\ddot{\varepsilon}_{it} = \varepsilon_{it} - \overline{\varepsilon}_i$.

Dado que a implementação desse método subtrai a média ao longo do tempo para cada muniípio i, não é possível que as variáveis explicativas sejam correlacionadas com o termo de erro contemporaneamente e nem em qualquer outro período do tempo. A outra hipótese necessária é a inexistência de perfeita multicolinearidade, ou seja,

(4.3) $Posto\ [E(X_{it}X_{it})] = \mathrm{K}$

A terceira hipótese refere-se à presença de erros homocedásticos e não autocorrelacionados, conforme enunciado abaixo

(4.4) $E\left(\ddot{\varepsilon}_{it}\ddot{\varepsilon}_{it}/\ddot{X}_{it}\, a_i\right) = \sigma_u^2 I.$

No entanto, caso se assuma que os efeitos individuais são aleatórios, de forma que se entenda que o termo específico não observado (a_i) seja não correlacionado com X_{it}, a variância desse termo poderá ser estimada de acordo com o modelo de efeitos aleatórios. Esse modelo também necessita da hipótese de exogeneidade forte, $E(\varepsilon_{it}/X_{it}\,, a_i) = 0$, e da hipótese de não correlação de a_i para cada valor de X_{it}, $E(a_i/X_{it}) = 0$. Para a escolha entre o modelo de efeitos fixos ou aleatórios, foi utilizado o teste de especificação de Hausman (1978), que é utilizado para comparar os coeficientes estimados no modelo com efeitos fixos com aqueles estimados no modelo com efeitos aleatórios. Sob a hipótese nula deste teste, se os efeitos específicos não estão correlacionados com as variáveis explicativas, o estimador de efeitos aleatórios é consistente e eficiente, enquanto que o estimador de efeitos fixos produz estimativas consistentes, mas não eficientes.

Por outro lado, caso tais efeitos sejam correlacionados com as variáveis explicativas (ou seja, se a hipótese nula é rejeitada), o estimador de efeitos fixos será consistente, enquanto que o de efeitos aleatórios será inconsistente (o mais grave problema dentre as características desejáveis para um estimador). O teste de Hausman

apontou o modelo de efeitos fixos como o mais adequado para todas as especificações. O modelo de efeitos fixos controla as características não observáveis de cada município e constantes no tempo e, além disso, pode minorar os inúmeros problemas que cabem no termo de endogeneidade nas estimações com dados prospectivos (Franco e Menezes-Filho, 2009).

Mesmo considerando as vantagens de se aplicar a metodologia de efeitos fixos para dados em painel, a variável de saneamento é endógena nesse modelo: $Cov(SAN_{it}, \varepsilon_{it}) \neq 0$. O teste de Durbin-Wu-Hausman de endogeneidade dos regressores para o modelo de efeitos fixos rejeitou a hipótese nula de que a variável dependente é exógena. Além desse teste detectar a endogeneidade, a literatura do setor também traz argumentos que reforçam essa hipótese. A presença de saneamento afeta os indicadores de educação, pois melhores condições sanitárias nos domicílios ou escolas fornecem condições de infraestrutura física mais adequadas que, por sua vez, se refletem em superiores indicadores de desempenho escolar[5]. Por outro lado, maiores níveis de escolaridade levam a melhores condições de saneamento, seja por meio do conhecimento sobre as condições de higiene ou sobre as formas de se evitar e prevenir a incidência de doenças relacionadas a saneamento[6].

De acordo com Wooldridge (2002), há três possíveis fontes para a existência de endogeneidade nos dados: simultaneidade, erro de medida e a omissão de variáveis correlacionadas com pelo menos uma das variáveis explicativas. Neste estudo ocorre tanto o viés de simultaneidade, em que a variável explicativa afeta a dependente e vice-versa, quanto a existência de variáveis omitidas na explicação das variáveis de desempenho educacionais. A solução para resolver o problema de endogeneidade é a utilização do método de Mínimos Quadrados em Dois Estágios (MQ2E ou 2SLS) com variáveis instrumentais (Greene, 2003; Davidson & Mackinnon, 2004).

Para entender a construção do instrumento utilizado nesse estudo é necessário investigar a estrutura de provisão dos serviços de saneamento no Brasil. A privatização dos serviços de saneamento foi possível com a Lei de Concessões de 1995. Antes disso, os serviços eram oferecidos somente por empresas públicas, municipais ou estaduais. O governo federal implementou em 1971 o primeiro plano voltado para eliminação do

[5] Os artigos que ressaltam o *efeito-escola* mostram que a infraestrutura física das escolas, ou seja, a adequada provisão dos serviços de saneamento básico, é um fator que pode afetar o rendimento escolar.
[6] Kassouf (1995).

déficit de abastecimento de água e serviços de esgoto, o Plano Nacional de Saneamento, Planasa (MPO e Sepurb, 1995). Segundo Galvão Jr. *et. al.* (2009), o Planasa constituiu marco da participação dos estados na provisão e operação da infraestrutura de saneamento no Brasil, uma vez que foram criadas as Companhias Estaduais de Saneamento Básico (CESBs), que deveriam obter as concessões dos serviços diretamente do poder concedente, as autoridades municipais.

A estrutura de provisão do setor de saneamento é caracterizada pelo predomínio das empresas públicas regionais, refletindo os diversos incentivos concedidos no Planasa para esse tipo de organização. Mas existem outros tipos de provedores que apresentam diferentes formas de atuação geográfica e diversas naturezas jurídicas na forma de gestão. Segundo Seroa da Motta (2006), o setor de saneamento básico apresenta uma complexa estrutura de provisão, em que predominam diversos tipos de prestadores com objetivos diferentes. O SNIS (2015) permite identificar basicamente três grupos de prestadores, agrupados a partir de sua abrangência geográfica, que são: prestadores regionais, microrregionais e locais, cuja natureza jurídica pode ser tanto pública quanto privada[7].

As empresas que atendem a mais de um município, distribuídas por estado, e que geralmente atendem as regiões metropolitanas, são chamadas pelo SNIS de prestadores regionais. Existem empresas regionais públicas e privadas. As do primeiro tipo são as que atendem a maior parte da população brasileira. A Companhia de Saneamento Básico do Estado de São Paulo (Sabesp) e a Companhia de Saneamento de Minas Gerais (Copasa) são dois exemplos de empresas desse grupo. Existe uma única empresa regional privada no país, a Companhia de Saneamento do Tocantins (Saneatins), cuja natureza jurídica é de sociedade de economia mista, porém com administração privada. As empresas locais fornecem serviços de saneamento básico exclusivamente a um único município e também podem ser públicas ou privadas. A Tabela 4.1 apresenta o número de empresas nessas quatro categorias.

[7] Nesse trabalho convencionou-se que as empresas do grupo "privadas" são aquelas classificadas pelo SNIS como *empresa privada* e aquelas do grupo "públicas" são as empresas que exibem as seguintes naturezas jurídicas: *sociedade de economia mista com administração pública*, *empresa pública*, *autarquia* e *administração pública direta*.

Tabela 4.1: Empresas e municípios atendidos por empresas públicas e privadas

		2000		2010	
		Número de empresas	Número de municípios	Número de empresas	Número de municípios
PÚBLICAS	REGIONAL	27	3386	27	3522
	LOCAL	2047	2047	1813	1813
	TOTAL PÚBLICAS	2074	5433	1840	5335
PRIVADAS	REGIONAL	1	102	1	138
	LOCAL	28	28	90	90
	TOTAL PRIVADAS	29	**130/9***	91	**228/186***
PÚBLICAS + PRIVADAS	TOTAL	2103	5565	1931	5565

Fonte: Elaboração própria a partir de dados da Abcon e SNIS (2013)
*municípios que possuem empresas privadas de saneamento há cinco anos.

De acordo com essa Tabela 4.1, no ano 2000, do total de 130 municípios atendidos por empresas privadas de saneamento básico, em apenas nove deles essas empresas atuavam há pelo menos cinco anos. Em 2010, esse número aumentou para 186, de um total de 228 municípios que tinham seus serviços de saneamento providos por empresa privada há qualquer tempo.

O mapa da Figura 4.1 apresenta a distribuição espacial de tais municípios para o ano de 2010. O tempo de exposição refere-se ao número de anos que a empresa privada, que fornece os serviços de saneamento, atua no município. Nota-se que as empresas que ofertam tais serviços nos municípios do Tocantins operam nestes há mais de uma década.

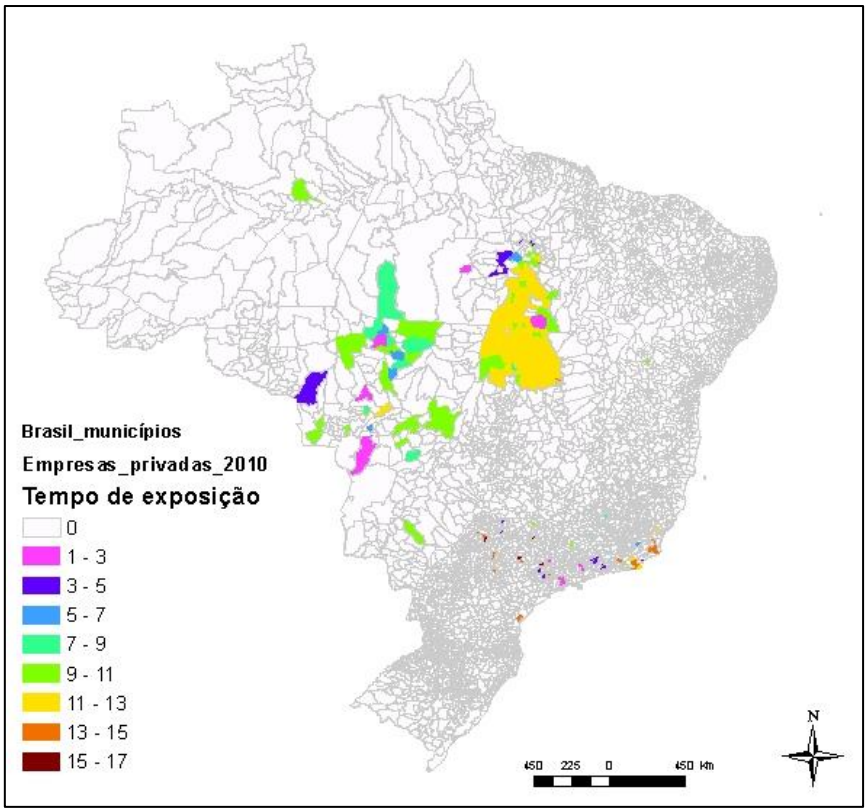

Figura 4.1: Tempo de exposição das empresas privadas nos municípios
Fonte: Elaboração própria a partir de dados da pesquisa.

Portanto, o instrumento utilizado nesse estudo será o número de empresas que foram privatizadas há mais de 5 anos para os dois períodos de tempo. Para que o instrumento seja válido, é importante que duas hipóteses sejam satisfeitas. A primeira delas refere-se à relevância do instrumento: $E(PRIV_{it}, SAN_{it}) \neq 0$. Ou seja, o mesmo deve ser correlacionado com a variável endógena (saneamento). Os resultados a serem apresentados adiante indicam que os municípios com serviços de água e esgoto fornecidos por uma empresa privada há mais de cinco anos exibem melhores indicadores de acesso a saneamento básico. No entanto, o problema de instrumento fraco pode surgir, mesmo quando a correlação entre a variável endógena e o instrumento é significante aos níveis convencionais de 5% e 1%. O teste de Stock-Yogo[8] (2005), que verifica se o instrumento é fraco ou o estimador é fracamente identificado, rejeitou a hipótese nula de que o instrumento é fraco.

[8] O teste de Stock-Yogo (Stock e Yogo, 2005) é calculado tendo por base a estatística F de Cragg e Donald (1993). Sob a hipótese nula o estimador é fracamente identificado no sentido de que o viés verificado é inaceitavelmente grande. Os valores críticos deste teste são obtidos a partir de uma distribuição não padrão.

A outra hipótese necessária para assegurar a validade do instrumento diz respeito à exogeneidade do mesmo: $E(PRIV_{it}, \varepsilon_{it}) = 0$. Ou seja, a decisão de privatizar, que envolve um contrato entre o município e a empresa privada, não deve ser correlacionada com fatores não observados dos indicadores de educação. Apesar dessa hipótese não ser passível de teste, é possível garantir sua validade, pois a literatura mostra que essa decisão é política. Saiani e Azevedo (2014) apontam que as privatizações podem ser entendidas como estratégias políticas. O primeiro fato desse resultado é que os riscos eleitorais percebidos pelos prefeitos, diante da perspectiva de insucesso no próximo pleito, motivariam a decisão pela privatização, que seria adotada para "amarrar as mãos" da próxima administração. Ou seja, os prefeitos privatizariam para diminuir a discricionariedade[9] de um eventual oponente incumbente (mesmo que isso implique em reduzir o seu próprio escopo de atuação). Em relação ao segundo fato, considerando a existência de provisão dos serviços de saneamento básico pelos governos estaduais, o outro fato está relacionado a reduzir (ou não aumentar) o escopo de atuação dos governadores estaduais nos municípios nos quais os prefeitos não pertencem a partidos da coligação de tais governadores (SAIANI e AZEVEDO, 2014). Em outras palavras, esta ação se constitui em estratégia política para diminuir a importância do governo do estado na economia do município e até mesmo do próprio estado.

A partir de tais argumentos e dos testes realizados, que são apresentados na seção de resultados, é possível garantir que o instrumento apresentado é válido como forma de solucionar a endogeneidade por meio do método de Mínimos Quadrados em dois estágios para dados em painel.

A partir de tal instrumento e das variáveis exógenas, que são instrumentos delas mesmas, uma matriz Z é utilizada para realizar o procedimento em dois estágios. No primeiro estágio, estima-se a equação na forma reduzida, ou seja, a variável explicativa endógena é regredida em função da matriz de instrumentos Z, de tal forma que:

(4.5) $San_{it} = Z_{it}\gamma + \vartheta_{it}$

(4.6) $San_{it} = \widehat{San_{it}} + \widehat{\vartheta_{it}} \rightarrow \widehat{\vartheta_{it}} = San_{it} - \widehat{San_{it}}$

[9] Ações discricionárias em diversas áreas de serviços públicos, entre elas, a de saneamento, podem ser favoráveis para garantir votos a políticos que desejam se reeleger. Por exemplo, evitar a interrupção de serviços por falta de pagamento e assegurar empregos no setor são medidas populares que podem render votos (principalmente quando ações como essa recaem sobre a população de baixa renda).

O parâmetro γ precisa ser significante a fim de que seja válida a relevância do instrumento. A variável endógena tem duas partes: uma que não é correlacionada com o erro ε_{it} e a segunda que é correlacionada com o mesmo. No segundo estágio, estima-se o modelo para a variável dependente em função dos valores previstos do primeiro estágio, ou seja, utiliza-se $\widehat{San_{it}}$ como instrumento para San_{it}.

(4.7) $$EDUC_{it} = \widehat{San_{it}} + \widehat{\vartheta_{it}} + a_i + \varepsilon_{it}$$

A intuição do método é que o instrumento afete a variável de resposta somente por meio da correlação que apresenta com a variável endógena. Em outras palavras, o instrumento impacta a variável endógena via parte que não é correlacionada com erro. A estimação não realizou o segundo estágio manualmente, e sim de uma única vez, de forma a evitar que os erros-padrão e as estatísticas de teste obtidos estivessem invalidados[10].

4.1.1 Análise Exploratória dos Dados Espaciais (AEDE)

Anteriormente à análise empírica dos dados espaciais, é necessário verificar se existem evidências de dependência espacial no modelo. No contexto dessa seção, é preciso investigar se os resíduos do modelo convencional estimado em dois estágios com efeitos fixos e variável instrumental mostram-se autocorrelacionados espacialmente. Ou seja, deve-se aplicar o teste *I* de Moran para os resíduos das unidades de corte transversal para cada período (ALMEIDA, 2012). A matriz de ponderação espacial (w_{ij}^*) contígua e normalizada na linha e o coeficiente de autocorrelação espacial *I* de Moran são definidos a seguir:

(4.8) $$w_{ij}^* = \frac{w_{ij}}{\sum_j w_{ij}}$$

(4.9) $$w_{ij}^* = \begin{cases} 0, se\ não\ há\ fronteira\ entre\ i\ e\ j \\ 1, se\ há\ fronteira\ entre\ i\ e\ j. \end{cases}$$

(4.10) $$I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_{i=1}^{n} z_i^2}$$

Na equação 4.10, *n* é o número de regiões, *z* denota os valores das variáveis de interesse padronizadas e *Wz* representa os valores médios das variáveis de interesse padronizadas dos vizinhos. As variáveis de interesse são os resíduos da regressão do

[10] As estimações foram realizadas pelo *software Stata* por meio do comando "xtivreg".

modelo não espacial e as variáveis dependentes de educação. Dado que a matriz de pesos espaciais é normalizada na linha, tem-se que $n = S_0$. Duas formas de inferência foram adotadas: a primeira delas ("Res_1" e "Lag") foi baseada em um processo que assume 99 permutações aleatórias de Monte Carlo; a segunda ("Res_2") assume que a variável padronizada de interesse *Z(I)* segue uma distribuição normal com média zero e variância unitária.

Na Tabela 4.2 encontram-se, para cada ano, as estatísticas *I* de Moran calculadas com os resíduos, segundo os dois tipos de inferência, e com a variável dependente segundo inferência via permutações aleatórias.

Tabela 4.2: Estatísticas I de Moran para os resíduos e defasagem espacial.

I de Moran		Frequência Escolar	Distorção Idade-Série	Abandono Escolar
2000 (Res_1)	Estatística	0,0037	0,0432	0,054
	P-Valor	0,37	0,001	0,001
2000 (Res_2)	Estatística	0,00374932	0,043228031	0,05398291
	P-Valor	0,717	0,0001	0,000001
2010 (Res_1)	Estatística	0,0037	0,0432	0,054
	P-Valor	0,37	0,001	0,001
2010 (Res_2)	Estatística	0,00374937	0,043228032	0,05398290
	P-Valor	0,717	0,0001	0,000001
2000 (Lag)	Estatística	0,0514	0,365	0,0336
	P-Valor	0,001	0,001	0,001
2010 (Lag)	Estatística	0,0166	0,0722	0,2196
	P-Valor	0,07	0,001	0,001

Fonte: Elaboração própria a partir de dados da pesquisa.

Os dados da Tabela 4.2 mostram que a hipótese nula de aleatoriedade espacial com os resíduos do modelo não espacial foi rejeitada com segurança para as variáveis de distorção idade-série e abandono escolar pelos dois tipos de testes. No entanto, as três variáveis apresentaram evidências de presença de autocorrelação espacial na forma de defasagem espacial da variável dependente. Tais conclusões valem para os dois períodos de tempo em análise. Portanto, pelos testes acima apresentados, os efeitos espaciais precisam ser controlados nas estimações econométricas para garantir que os resultados não estejam sujeitos a viés ou inconsistência.

4.1.2 Metodologia espacial

Apesar de ser recente a aplicação de modelos espaciais em painel, diversos estudos empíricos, em diversas áreas, adotam essa metodologia. Há dois principais métodos desenvolvidos para a estimá-los: o método da Máxima Verossimilhança (MV)[11] e o método das Variáveis Instrumentais ou dos Momentos Generalizados (GMM)[12]. O método de estimação para os dados em painel com erros espacialmente correlacionados e defasagem espacial desse estudo foi o GMM, de Piras e Millo (2012)[13]. Segundo Almeida (2012), Kelejian and Prucha (1998), Brueckner (2003) e Milimet e Rangaprasad (2007), a vantagem deste método, em comparação ao de Máxima Verossimilhança (MV), deve-se à estimação consistente de modelos espaciais que apresentam variáveis endógenas entre as variáveis explicativas, além da defasagem espacial, e aqueles que também apresentam erros espacialmente correlacionados.

De acordo com o apresentado na seção anterior, esse é exatamente o caso deste estudo. Há endogeneidade na variável explicativa de saneamento e, além disso, efeitos espaciais detectados pela Análise Exploratória dos Dados Espaciais (AEDE). Não considerar tanto os efeitos fixos quanto a dependência espacial leva a uma ampliação do viés das estimativas em relação ao modelo convencional de dados em painel. Outro fator que favorece a utilização do GMM é que, em modelos estimados por máxima verossimilhança, os parâmetros espaciais são estimados no intervalo $\left(\frac{1}{\alpha_{min}}, \frac{1}{\alpha_{max}}\right)$. O método IV/GMM não restringe esses parâmetros a este intervalo no processo de estimação (GOLGHER, 2015). No entanto, o problema do GMM é que este método não consegue determinar a significância estatística para o parâmetro ρ, que é considerado um parâmetro de distúrbio, porque é apenas usado para filtrar as variáveis (ALMEIDA, 2012). Piras e Millo (2012) se basearam em um modelo do tipo Cliff e Ord, analisado em Kapoor *et al*. (2007), e em uma extensão do estimador generalizado dos momentos (GMM) proposto por Kelejian e Prucha (1999) para estimar o modelo de erro espacialmente correlacionado. O modelo de defasagem espacial segue os estudos de Mutl e Pfaffermayr (2011) e Piras (2011). A especificação de tais modelos assume a seguinte forma:

---

[11] Elhorst (2014), Kalenkoski e Lacombe (2012) e, para o caso brasileiro, Chagas *et al.* (2014).
[12] Baylis et al. (2011).
[13] A estimação foi realizada pelo software R e o pacote utilizado foi o splm.

(4.9) $y_N(t) = \lambda Wy + X_N(t)\beta + u_N(t)$

(4.10) $u_N(t) = \rho W_N u_N(t) + \varepsilon(t)$

(4.11) $\varepsilon_N = (e_T \otimes I_N)\mu_N + \upsilon_N$

Nas equações acima, o termo $y_N(t)$ é um vetor de dimensão $(N * t \times 1)$, $X_N(t)$ é uma matriz $(N * t \times k)$ de variáveis explicativas, I_N é uma matriz identidade $(N \times N)$, W_N é a matriz de pesos espaciais cuja dimensão também é $(N \times N)$, o vetor $u_N(t)$ de erros $(N * t \times 1)$ e $\otimes$ produto Kronecker entre duas matrizes.

Quando $\lambda = 0$, os parâmetros ρ, σ_μ^2 e σ_ν^2 são estimados por GMM e os coeficientes restantes do modelo por um estimador de Mínimos Quadrados Generalizados (MQG ou GLS) factível. No entanto, ao se considerar incluídos no modelo tanto o termo da variável dependente defasada espacialmente (WY_{educ}) quanto os efeitos fixos, o procedimento passa a ser o de Mínimos Quadrados em Dois Estágios Espacial (Baltagi and Liu 2011).

De acordo com esse procedimento, além da variável dependente defasada ser instrumentalizada pelo vetor de parâmetros sugerido pelos autores, existe a possibilidade de instrumentalizar outro vetor de variáveis explicativas endógenas. Neste estudo, esse vetor é composto de uma única variável explicativa endógena, *saneamento*, que será instrumentalizada por *priv5anos*. Como apresentado em seção 3.2.1 pelos argumentos teóricos, testes implementados e exercício de robustez, esse instrumento possui as propriedades desejáveis de relevância e exogeneidade e por isso pode ser considerado também na análise espacial. Logo, reescrevendo a equação 9, o modelo de efeitos fixos pode ser expresso por:

(4.12) $y_N(t) = \lambda Wy + X_N(t)\beta + D_N\gamma + u_N(t) = Z_N\delta + u_N(t)$

Com tais modelos, tanto o problema das variáveis endógenas (a dependente defasada que é endógena por natureza e outras variáveis explicativas que também podem ser) quanto o problema de autocorrelação espacial dos erros são tratados. Segundo Anselin (1998), apesar dos problemas causados pela heterogeneidade espacial poderem ser abordados por técnicas econométricas não espaciais, o conhecimento teórico sobre a estrutura espacial dos dados pode tornar os procedimentos de estimação mais eficientes.

4.2 Resultados não-espaciais

As estimações do modelo não espacial descrito acima foram realizadas para sete diferentes variáveis dependentes: taxa de frequência escolar, taxa de abandono escolar no ensino fundamental e taxa de distorção idade-série também no ensino fundamental. Essas duas últimas foram segmentadas para o primeiro e segundo ciclo. Os resultados do primeiro estágio são apresentados na Tabela 4.3:

Tabela 4.3: Resultados do primeiro estágio

VARIÁVEL DEPENDENTE	**Saneamento**
priv5anos - Variável Instrumental	11,18***
	(0,839)
Constante	66,06***
	(5,370)
Observações	7.959
R-quadrado (within)	0,7748
R-quadrado (between)	0,5572
R-quadrado (overall)	0,5999
Número de municípios	4.347
Características socioeconômicas	SIM
Características da escola e professores	SIM
Saúde e educação	SIM
Efeito fixo de município	SIM
Erro-padrão robusto entre parênteses	
*** p<0,01; ** p<0,05; * p<0,1	

Fonte: Elaboração própria a partir de dados da pesquisa.

O primeiro estágio mostra a significância da variável instrumental em relação à variável endógena de interesse. Esse é um dos indícios de que o instrumento é relevante: ter serviços de saneamento oferecidos por uma empresa privada há mais de cinco anos é correlacionado positivamente com a variável de acesso a saneamento, como era esperado.

No entanto, como o problema de instrumento fraco pode surgir mesmo quando a correlação entre a variável endógena e o instrumento é significante aos níveis convencionais de 5% e 1%, foi realizado o teste de Stock-Yogo (2005). O resultado do teste indicou a relevância do instrumento, demonstrando que não se trata de um instrumento fraco.

Nas tabelas seguintes, a primeira estimação refere-se a um painel de efeitos fixos sem nenhum controle. Na segunda, terceira e quarta estimações, os controles foram sendo

adicionados. A quinta estimação foi realizada em dois estágios, utilizando a variável instrumental descrita anteriormente. As estatísticas dos testes de Hausman, para escolha de se estimar os dados em painel por efeito fixo ou aleatório, validade do instrumento e endogeneidade, são descritas nas últimas linhas das tabelas. A Tabela 4.4 mostra as estimações descritas acima para a variável de frequência escolar.

Tabela 4.4: Impacto do saneamento em frequência escolar 6-14 anos

VARIÁVEL DEPENDENTE	**Frequência escolar da população de 6 a 14 anos**				
	Efeito Fixo				IV Efeito Fixo
	(1)	(2)	(3)	(4)	(5)
Saneamento	0,224***	0,040***	0,022***	0,024***	0,118***
	(0,0041)	(0,0078)	(0,0085)	(0,0092)	(0,0312)
Constante	76,61***	88,83***	93,45***	93,96***	87,94***
	(0,329)	(1,926)	(2,656)	(2,746)	(3,005)
Observações	8.738	8.681	8.163	7.959	7.959
R-quadrado	0,485	0,628	0,651	0,642	-
Número de municípios	4.369	4.369	4.365	4.347	4.347
Características socioeconômicas	NÃO	SIM	SIM	SIM	SIM
Características da escola e professores	NÃO	NÃO	SIM	SIM	SIM
Saúde e educação	NÃO	NÃO	NÃO	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM	SIM
priv5anos - Variável Instrumental	-	-	-	-	SIM
Hausman - Chi2	-	-	-	832,2***	-
Durbin Wu Hausman - Chi2 (18)	-	-	-	-	36,79
Durbin Wu Hausman p-value	-	-	-	-	(0,0056)
Yogo and Stock (2005) - F statistic	-	-	-	-	154,96***

Erro-padrão robusto entre parênteses
*** p<0,01; ** p<0,05; * p<0,1

Fonte: Elaboração própria a partir de dados da pesquisa.

Os resultados da tabela acima mostram que o aumento de um por cento na proporção de pessoas que vivem em domicílios com acesso a água e esgoto gera aumento de 0,11 pontos percentuais na variável de frequência escolar da população de 6 a 14 anos. Esse resultado é compatível com o encontrado por Neri et al. (2008), de que o acesso aos serviços públicos (tanto água quanto luz elétrica) gera maiores índices de matrícula. A estimação de dados em painel por efeitos fixos, em dois estágios, utilizando variável instrumental, gerou coeficiente 4,9 vezes maior em relação ao coeficiente do modelo de efeitos fixos. O coeficiente deste modelo que era 0,024 (quarta coluna) foi para 0,118 no modelo completo expresso na quinta coluna, indicativo de que este tipo de controle pela endogeneidade é relevante para a análise em questão.

A Tabela 4.5 apresenta os mesmos tipos de estimação para a variável taxa de abandono escolar no ensino fundamental. O resultado obtido anteriormente é corroborado: um crescimento marginal na variável de saneamento provoca uma queda na taxa de abandono escolar em 0,49 pontos percentuais. Quando se controla a endogeneidade pelo uso da variável instrumental, o coeficiente torna-se oito vezes maior, em módulo, em relação àquele da estimação sem o controle. Ou seja, o coeficiente da variável de saneamento, que contém os três vetores de variáveis explicativas e o efeito fixo de município, porém, não faz o controle pela endogeneidade, é -0,0576 enquanto que o coeficiente desta mesma variável para o modelo completo da última coluna é -0,499.

Tabela 4.5: Impacto do saneamento na taxa de abandono escolar

VARIÁVEL DEPENDENTE	**Taxa de abandono escolar - ensino fundamental**				
		Efeito Fixo			IV Efeito Fixo
	(1)	(2)	(3)	(4)	(5)
Saneamento	-0,371***	-0,106***	-0,0715***	-0,0576***	-0,499***
	(0,00577)	(0,0101)	(0,0106)	(0,0109)	(0,0528)
Constante	36,35***	-15,60***	-12,09***	-14,45***	13,99***
	(0,458)	(2,615)	(3,406)	(3,432)	(5,080)
Observações	8.683	8.680	8.163	7.959	7.959
R-quadrado	0,542	0,706	0,725	0,724	-
Número de municípios	4.369	4.369	4.365	4.347	4.347
Características socioeconômicas	NÃO	SIM	SIM	SIM	SIM
Características da escola e professores	NÃO	NÃO	SIM	SIM	SIM
Saúde e educação	NÃO	NÃO	NÃO	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM	SIM
priv5anos - Variável Instrumental	-	-	-	-	SIM
Hausman - Chi2	-	-	-	408,99***	-
Durbin Wu Hausman - Chi2 (18)	-	-	-	-	895,12
Durbin Wu Hausman p-value	-	-	-	-	(0,000)
Yogo and Stock (2005) - F statistic	-	-	-	-	154,9***

Erro-padrão robusto entre parênteses
*** p<0,01; ** p<0,05; * p<0,1

Fonte: Elaboração própria a partir de dados da pesquisa.

A Tabela 4.6 também confirma os resultados anteriores, agora com respeito à taxa de distorção idade-série (TDI). O resultado indica que um aumento de um por cento na população com adequado acesso a saneamento gera uma queda de 0,96 pontos percentuais na taxa de distorção idade-série no ensino fundamental. Nesse caso, o controle da endogeneidade (modelo completo) apresentou coeficiente quatro vezes maior em relação ao modelo da quarta coluna, sem o controle pelo instrumento.

Tabela 4.6: Impacto do saneamento na taxa de Distorção Idade-Série

VARIÁVEL DEPENDENTE	**Taxa de Distorção Idade-Série Ensino Fundamental**				
	Efeito Fixo				IV Efeito Fixo
	(1)	(2)	(3)	(4)	(5)
Saneamento	-0,741***	-0,310***	-0,259***	-0,234***	-0,965***
	(0,00917)	(0,0156)	(0,0168)	(0,0180)	(0,0847)
Constante	89,42***	11,24***	19,36***	18,28***	65,39***
	(0,725)	(4,346)	(5,529)	(5,652)	(8,147)
Observações	8.589	8.586	8.163	7.959	7.959
R-quadrado	0,666	0,784	0,808	0,806	-
Número de municípios	4.369	4.369	4.365	4.347	4.347
Características socioeconômicas	NÃO	SIM	SIM	SIM	SIM
Características da escola e professores	NÃO	NÃO	SIM	SIM	SIM
Saúde e educação	NÃO	NÃO	NÃO	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM	SIM
priv5anos - Variável Instrumental	-	-	-	-	SIM
Hausman - Chi2	-	-	-	541,45***	-
Durbin Wu Hausman - Chi2 (18)	-	-	-	-	1455,68
Durbin Wu Hausman p-value	-	-	-	-	(0,000)
Yogo and Stock (2005) - F statistic	-	-	-	-	154,9***

Erro-padrão robusto entre parênteses
*** p<0,01; ** p<0,05; * p<0,1

Fonte: Elaboração própria a partir de dados da pesquisa.

A segmentação das duas últimas variáveis em primeiro e segundo ciclos é importante porque estes abrangem população de idades diferentes e com organismos e comportamentos diversos.

O primeiro ciclo é referente à população de seis a dez anos que frequenta a primeira fase do Ensino Fundamental (de primeira à quarta série ou do primeiro ao quinto ano). O segundo diz respeito à população de dez a catorze anos que frequenta a segunda fase do Ensino Fundamental (de quinta à oitava série ou do sexto ao nono ano).

Os resultados da Tabela 4.7 indicam que não houve importantes diferenças entre os ciclos em relação à magnitude dos coeficientes para a taxa de distorção idade-série. Enquanto a queda do primeiro ciclo foi de 1,0 ponto percentual, a do segundo ciclo foi de 1,04 pontos percentuais, considerando o modelo completo de efeitos fixos com vaiável instrumental.

Tabela 4.7: Impacto do saneamento na TDI (1º e 2º ciclos)

VARIÁVEL DEPENDENTE	**Distorção Idade-Série (Primeiro ciclo)**		**Distorção Idade-Série (Segundo ciclo)**	
	Efeito Fixo	IV Efeito Fixo	Efeito Fixo	IV Efeito Fixo
	(1)	(2)	(3)	(4)
Saneamento	-0,271***	-1,004***	-0,259***	-1,049***
	(0,0196)	(0,0873)	(0,0207)	(0,100)
Constante	20,58***	67,65***	26,40***	77,27***
	(5,872)	(8,422)	(7,031)	(9,644)
Observações	7.951	7.951	7.959	7.959
R-quadrado	0,77	-	0,783	-
Número de municípios	4.347	4.347	4.347	4.347
Características socioeconômicas	SIM	SIM	SIM	SIM
Características da escola e professores	SIM	SIM	SIM	SIM
Saúde e educação	SIM	SIM	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM
Efeito fixo de ano	SIM	SIM	SIM	SIM
Hausman - Chi2	582,61***	-	502,87***	-
priv5anos - Variável Instrumental	-	SIM	-	SIM

Erro-padrão robusto entre parênteses
*** p<0.01, ** p<0.05, * p<0.1

Fonte: Elaboração própria a partir de dados da pesquisa.

Para a variável abandono escolar, os resultados da Tabela 4.8 mostram que o primeiro ciclo registrou redução ligeiramente maior: neste, um aumento percentual de saneamento gera uma queda de 0,52 pontos percentuais; no segundo ciclo, essa queda é de 0,47 pontos percentuais.

Em que pese a pequena diferença quantitativa, essa evidência vai ao encontro da hipótese estabelecida de que a população de seis a dez anos pode estar mais sujeita às doenças infecto-parasitarias transmitidas pela água do que os alunos mais velhos (crianças de dez a catorze anos). Quando estão menos expostas, uma vez inseridas em melhores condições de saneamento básico, ficam menos doentes, frequentando mais as aulas, repetindo menos e, portanto, tendo menos fatores associados ao estado de saúde para abandonar os estudos.

Tabela 4.8: Impacto do saneamento na taxa de Abandono Escolar (1º e 2º ciclos)

VARIÁVEL DEPENDENTE	Abandono escolar (Primeiro ciclo)		Abandono escolar (Segundo ciclo)	
	Efeito Fixo	IV Efeito Fixo	Efeito Fixo	IV Efeito Fixo
	(1)	(2)	(3)	(4)
Saneamento	-0,112***	-0,527***	-0,00484	-0,479***
	(0,0118)	(0,0528)	(0,0139)	(0,0664)
Constante	-2,694	24,00***	-25,17***	5,433
	(3,580)	(5,083)	(4,358)	(6,390)
Observações	7.958	7.958	7.958	7.958
R-quadrado	0,678	-	0,633	-
Número de municípios	4.347	4.347	4.347	4.347
Características socioeconômicas	SIM	SIM	SIM	SIM
Características da escola e professores	SIM	SIM	SIM	SIM
Saúde e educação	SIM	SIM	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM
Efeito fixo de ano	SIM	SIM	SIM	SIM
Hausman - Chi2	373,2***	-	455,2***	-
priv5anos - Variável Instrumental	-	SIM	-	SIM

Robust standard errors in parentheses
*** p<0,01; ** p<0,05; * p<0,1

Fonte: Elaboração própria a partir de dados da pesquisa.

O teste de robustez apresentado abaixo procura verificar se há correlação entre o instrumento e a porcentagem do ensino cuja provisão é municipal, buscando determinar se a privatização dos serviços de saneamento básico poderia estar associada à descentralização também do ensino no nível municipal. Ou seja, caso a correlação seja positiva e alta, governos que privatizam os serviços de saneamento poderiam ter assumido maior protagonismo também na área de ensino, encampando atividades que antes eram estaduais ou municipais. Isso poderia ocorrer por serem mais dinâmicos e eficientes, ou mais alinhados ideologicamente com a ideia do aumento do poder decisório no nível municipal. Nessa situação, o instrumento poderia ter algum impacto também nos fatores não observados das variáveis educacionais.

O teste foi feito tanto com o instrumento utilizado nas regressões (município tem os serviços de saneamento básico oferecidos por uma empresa privada há mais de cinco anos) quanto com uma variável mais geral, que se refere aos municípios que têm os serviços de saneamento básico oferecidos por uma empresa privada, independentemente do tempo. Os resultados encontram-se na Tabela 4.9 abaixo:

Tabela 4.9: Análise de robustez para o instrumento

Períodos	2000			2010		
Variáveis	priv5anos	priv	p_escol_mun	priv5anos	priv	p_escol_mun
priv5anos	1,0000	-	-	1,0000	-	-
priv	0,2594	1,0000	-	0,8987	1,0000	-
p_escol_mun	**-0,0546**	**0,0001**	1,0000	**-0,0588**	**-0,0597**	1,0000

Fonte: Elaboração própria a partir de dados da pesquisa.

A conclusão do teste é que, à medida que cresce o número de municípios que têm seus serviços oferecidos por uma empresa privada há mais de cinco anos, decresce a porcentagem do ensino que é responsabilidade do governo municipal. Mesmo que essa correlação seja de baixa magnitude nos dois anos (5,4 e 5,9%), não existe evidência de que a privatização dos serviços de saneamento básico esteja associada à descentralização do ensino no nível municipal.

A Tabela 4.10 apresenta os sinais e significâncias das variáveis explicativas utilizadas nas regressões, separadas nos três principais vetores descritos anteriormente: o primeiro deles refere-se às características socioeconômicas, o segundo às características das escolas e dos professores e o terceiro refere-se às variáveis de imunização e despesas com saúde e educação. Células em branco indicam que o coeficiente estimado não se mostrou estatisticamente diferente de zero; um sinal positivo indica que foi significante e positivo; um sinal negativo indica que foi significante e negativo. O número de sinais indica o nível de significância (um sinal, 10%; dois sinais, 5%; três sinais, 1%).

A comparação entre o modelo de efeitos fixos sem variável instrumental e o modelo de efeitos fixos estimado em dois estágios com esta variável permite concluir que controlar a endogeneidade altera os sinais das estimativas dos coeficientes em alguns casos (variáveis bolsa família e número de docentes com ensino superior). A despeito da robustez do sinal, significância e magnitude da variável de interesse para os sete tipos de especificações realizadas, algumas variáveis explicativas apresentaram resultados contra-intuitivos, problema que será contornado ao se controlar os efeitos espaciais na análise.

		Variáveis Dependentes					
	Variáveis Explicativas	**Frequencia escolar**		**Distorção idade-série**		**Abandono escolar**	
		EF	IV EF	EF	IV EF	EF	IV EF
	t_analfabetismo_25m	- - -	- - -	+ + +	+ + +	+ + +	+ + +
	pop_negra_parda	+ + +	+ + +	-	- - -	- - -	- - -
	renda_per_capita	- - -	- - -		-		- - -
(1)	gini	+	+	+	+ + +	+ + +	+ + +
	pop_55mais			+ + +	+ + +	+ + +	+ + +
	pop_15a29	- -	- - -	+ + +	+ + +	+ + +	+ + +
	bolsa_familia	+ + +	+ + +	- - -	+ + +	- - -	+ + +
	pop_total			-	- -		- - -
	urbanização	+ + +	+ + +		+ + +	+ +	+ + +
	alunos_turma_fundamental	- - -	- - -	+ + +	+ + +	+ +	+
	docente_ensino_superior_fundamental	+ + +		- - -	+ + +	- - -	+
	horas_aula_fundamental			- - -	- - -	-	-
(2)	indice de infraestrutura escolar tecnológica				- - -	- - -	- - -
	indice de infraestrutura escolar básica				+ + +		+ + +
	porcentagem de escolas públicas						
	porcentagem de escolas com diretoria	+		+ + +	+ + +	+ +	+ + +
	despesa_educação_cultura_por_habitante	+ + +	+ + +			- -	
(3)	despesa_saude_por_habitante			- -		- -	
	imunização	+					
	Observações	7,959	7,959	7,959	7,959	7,959	7,959
	Número de municípios	4,347	4,347	4,347	4,347	4,347	4,347
	Erro-padrão robusto entre parênteses						
	( ) não significante, (+ + +) ou (- - -) p<0.01, (+ +) ou (- -) p<0.05, (+) ou (-) p<0.1						

Quadro 4.1: Sinal e significância e sinais das variáveis explicativas

Para o primeiro grupo de variáveis, a maior a taxa de analfabetismo da população adulta (de 25 anos ou mais) está associada a piores indicadores educacionais para as três variáveis: menor frequência escolar, maior abandono e maior taxa de distorção idade-série. Os municípios mais desiguais (maior índice de Gini) apresentaram piores indicadores para taxa de abandono e TDI. Para a variável de frequência escolar, esses mesmos municípios (com alto grau de concentração de renda) apresentaram maiores valores para a mesma. Por outro lado, esse indicador de frequência escolar tem associação positiva com o número de beneficiários do programa Bolsa Família no município, como era esperado[14]. Em relação ao vetor de características das escolas e dos professores, a maior razão alunos por turma impacta negativamente todos os indicadores educacionais.

Por outro lado, a maior quantidade de horas-aula diminui a taxa de distorção idade-série e o abandono, mas não afeta a frequência; assim como o índice de infraestrutura tecnológica. As maiores despesas com educação e cultura aumentam a frequência escolar e diminuem a taxa de abandono escolar no ensino fundamental.

[14]A condicionalidade do Programa Bolsa Família (PBF) estabelece que crianças e adolescentes de 6 a 15 anos devem possuir 85%, no mínimo, de frequência escolar para receber o Benefício Variável Jovem (BVJ). Além disso, a família assume compromissos nas áreas de saúde, tais como: acompanhamento pré-natal e acompanhamento nutricional (Campello e Neri, 2013).

4.3 Resultados espaciais

Na Tabela 4.10 são apresentados os resultados de quatro estimações para dados em painel. A primeira delas não leva em consideração o espaço. Ou seja, trata-se do resultado obtido em seção anterior, que se utilizou de um modelo de efeitos fixos para dados em painel, com variável instrumental, em dois estágios. As outras três partem do método de estimação GMM/IV para dados espaciais relativo ao período 2000-2010. Apesar desses resultados não poderem ser analisados como o impacto marginal da variável de saneamento sobre os indicadores educacionais[15], nota-se a diferença do controle da natureza espacial dos dados em painel.

Na quase totalidade dos casos (com exceção da variável de frequência escolar), os modelos espaciais resultaram em menor magnitude para as estimativas dos coeficientes. No contexto do impacto de mudanças climáticas sobre a agricultura, Baylis et al. (2011) também encontrou redução da magnitude das estatísticas dos coeficientes ao comparar o modelo de dados em painel com efeitos fixos não-espacial com o modelo também de efeitos fixos, mas que considera as características espaciais dos dados.

Apesar da variável da frequência escolar não apresentar evidências de autocorrelação espacial nos resíduos, esse indicador varia de acordo com a realidade socioeconômica dos estados e regiões (PONTILI & KASSOUF)[16]. São tais diferenças que justificam a necessidade de se analisar essa variável, pois ainda que o país apresente médias elevadas para ela[17], há diversas regiões que apresentam baixas taxas para a mesma, carecendo de políticas educacionais direcionadas. A ESDA apontou que o efeito espacial, neste caso, ocorre sobre a forma de variável dependente defasada. Assim, a estimação do modelo de erro foi realizada apenas a título de comparação com os outros modelos.

---

[15] Não foi possível calcular os efeitos marginais, pois o comando não está implementado quando há variável endógena entre as explicativas.

[16] A heterogeneidade espacial não observada admite situação de instabilidade estrutural, de forma que, a resposta de certos indicadores varia consideravelmente de acordo com a região de estudo.

[17] Segundo Menezes-Filho (2007), desde meados da década de 90 o Brasil conseguiu aumentar significativamente a frequência escolar em todos os níveis, de forma que, segundo o autor, o desafio agora está em melhorar a qualidade da educação que é oferecida para alunos na rede pública.

Tabela 4.10: Determinantes espaciais para a frequência escolar (2000-2010).

VARIÁVEL DEPENDENTE	**Frequencia escolar da população de 6 a 14 anos**			
	EF Não Espacial	EF Lag	EF Erro	EF Lag e Erro
	(1)	(2)	(3)	(4)
Saneamento	0,118***	0,0351***	0,2422***	0,0262**
	(0,0312)	(0,0130)	(0,0339)	(0,0123)
Observações	7,959	7,224	7,224	7,224
Número de municípios	4,347	3,612	3,612	3,612
Características socioeconômicas	SIM	SIM	SIM	SIM
Características da escola e professores	SIM	SIM	SIM	SIM
Saúde e educação	SIM	SIM	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM
Efeito fixo de ano	NÃO	NÃO	NÃO	NÃO
priv5anos - Variável Instrumental	SIM	SIM	SIM	SIM
Lambda	-	0,1428***	-	0,1686***
	-	(0,0321)	-	(0,0308)
Rho	-	-	0,1099	-0.0729
Sigma^2	-	-	-	14,74

Erro-padrão entre parênteses
*** p<0,01; ** p<0,05; * p<0,1

Fonte: Elaboração própria a partir de dados da pesquisa.

Os três últimos modelos da Tabela 4.10 apresentam resultados condizentes àqueles encontrados com o uso de econometria convencional: o acesso adequado aos serviços de saneamento básico municipal impacta positivamente o indicador de frequência escolar da população de 6 a 14 anos. Logo, há evidências de que dentro do grupo de fatores relacionados às características familiares e comunitárias na determinação dos níveis de escolaridade alcançado pelos indivíduos, o acesso aos serviços de saneamento básico no domicílio mostrou-se importante para alcançar esse objetivo.

No entanto, com exceção do modelo de erro espacial, a magnitude dos efeitos se alteraram. No Quadro 4.2, são apresentados o sinal e a significância das variáveis explicativas quando se faz o controle pelo efeito espacial. Os resultados corrigem alguns resultados contra intuitivos que foram anteriormente obtidos no modelo em que se utilizou econometria convencional. Pontili e Kassouf (2007) estimaram um modelo *probit* para duas variáveis qualitativas de frequência e atraso escolar dos estados de São Paulo e Pernambuco para o ano 2000, indicou os fatores que aumentam ou diminuem a probabilidade dos alunos de escolas públicas se matricularem nas mesmas e se manterem na série adequada a sua idade. De forma geral, os resultados encontrados por tais autoras são corroborados pelos resultados apresentados nesse estudo para essas duas variáveis.

Variáveis Explicativas	IV EF NE	EF Lag	EF Erro	EF Lag e Erro
t_analfabetismo_25m	- - -	- - -	- - -	- - -
pop_negra_parda	+ + +	+ + +	+ + +	+ + +
renda_per_capita	- - -	+ + +	+	+ + +
gini	+	- - -	- -	- - -
pop_55mais		+ + +	+ + +	+ + +
pop_15a29	- - -	+ + +	+ + +	+ + +
bolsafam	+ + +	+ + +	+ + +	+ + +
poptot			-	
urbanização	+ + +	+ + +	+ + +	+ + +
alunos_turma_fundamental	- - -	- - -	- - -	- - -
docentes_com_ens_superior		- - -	- - -	- - -
h_a_total_fundamental		+ + +		+ + +
índice infraestrutura tecnológica		- - -	- - -	- - -
índice infraestrutura básica		+ + +	+ + +	+ + +
porcent. escolas com diretoria		+		+
porcent. escolas públicas			-	
desp_educ_cult_hab	+ + +	+ + +	+ +	+ +
desp_saude_hab		+ + +	+ + +	+ +
imunização			+	

Quadro 4.2: Sinal e significância dos determinantes espaciais - frequência escolar.

Em relação ao grupo das variáveis socioeconômicas, algumas delas merecem destaque na análise, por terem seus sinais alterados ou por terem adquirido significância estatística.

A primeira delas é a renda per capita. Ao se controlar o efeito espacial, o aumento da renda per capita aumenta a frequência escolar em todos os modelos espaciais. Este resultado corrobora os resultados encontrados por Pontili e Kassouf (2007) e Vasconcellos (2003), que encontraram efeito positivo e significativo para a variável renda em relação à probabilidade de as famílias matricularem seus filhos na escola. A outra refere-se ao índice de Gini, cujo sinal passa a ser o esperado: quanto maior a concentração de renda, menor será a taxa de frequência escolar.

E, por fim, as outras duas variáveis dizem respeito à estrutura etária da população. A maior porcentagem da população adulta de 15 a 29 anos e da população de 55 anos ou mais estão associadas, individualmente, a maiores taxas de frequência escolar da população de 6 a 14 anos. Nesse caso, pode ser que os fatores que explicam a transição demográfica, tais como a maior expectativa de vida, o planejamento familiar, as quedas das taxas de fertilidade e o próprio aumento dos anos de escolaridade da população, impactem também o indicador de frequência escolar.

As variáveis relativas ao bolsa família, à taxa de urbanização e à taxa de analfabetismo da população de 25 anos ou mais apresentaram resultados esperados para todos os modelos estimados. A primeira deve-se às condicionalidades do programa Bolsa-Família; a segunda, às evidências de que em áreas urbanas a frequência escolar é maior do que em áreas rurais e a terceira, ao fato de que quanto mais analfabeta a população nessa faixa etária, principalmente em relação à escolaridade do chefe de família, menor é a frequência escolar.

O fato da maior porcentagem da população negra e parda estar associada a maiores indicadores de frequência escolar pode ser reflexo das políticas de inclusão ocorridas nos últimos anos. Tal resultado pode evidenciar um fenômeno recente, pois Pontili e Kassouf (2007) encontraram que os negros têm menor probabilidade de frequentar a escola que os brancos, ocorrendo o mesmo com os pardos da área rural de Pernambuco e urbana de São Paulo. Vasconcellos (2003) também encontrou resultados similares, indicando que as diferenças inter-raciais no Brasil se estendem à área de educação.

Quanto maior o número médio de horas-aula, o índice de infraestrutura escolar básica e a porcentagem de escolas com diretoria, maior é a taxa de frequência escolar da população na faixa etária de 6 a 14 anos. Além do melhor desempenho escolar obtido pelos alunos que passam entre quatro e cinco horas ou mais de cinco horas na sala de aula do que aqueles que ficam menos de quatro horas (MENEZES-FILHO, 2007), esse estudo mostra que o maior tempo de aula, por aumentar o tempo que o aluno permanece na escola, também tem um efeito na taxa de frequência escolar.

Em consonância com Pontili e Kassouf (2007), os resultados indicam que a maior proporção de escolas com biblioteca, que faz parte do índice de infraestrutura escolar básica formulado nesse estudo, afeta de maneira positiva a frequência de alunos na escola. E, por fim, a variável *proxy* de gestão escolar, a porcentagem de escolas com diretoria, também apresentou efeito positivo em frequência escolar.

A variável de despesa com educação e cultura se mostrou relevante para incrementar positivamente o nível de frequência escolar. No mesmo sentido, a variável de gastos com saúde também se mostrou importante. Ou seja, o aumento da despesa com saúde gera condições mais propícias para que os alunos frequentem as escolas.

No grupo de variáveis de controle educacionais, o maior número de docentes com ensino superior e o maior índice de infraestrutura tecnológica diminui a frequência escolar de forma robusta e significativa. Em relação à última variável, de acordo com Menezes-Filho (2007), há muita discussão acerca da inclusão digital, pois os impactos da presença

de computadores para alunos, diretores e professores são dúbios sobre o desempenho dos alunos.

O outro resultado contra intuitivo é aquele relativo ao impacto negativo da variável de porcentagem de docentes com curso superior sobre a frequência escolar. Uma das explicações pode ser que a decisão dos alunos frequentarem e permanecerem na escola esteja mais relacionada às características familiares do que a questões relacionadas à didática e motivação do professor em mantê-lo na mesma. Ou ainda, como ressalta Hanuschek e Rivkin (2006), os anos de escolaridade, experiência e capacitação muitas vezes não traduzem a qualidade do professor, no sentido de que podem existir características que são não observáveis nos dados. A Tabela 4.11 apresenta os resultados das estimações para os mesmos modelos espaciais em relação à taxa de distorção idade-série da população de 6 a 14 anos do ensino fundamental.

Tabela 4.11: Determinantes espaciais da taxa de distorção idade-série (2000-2010).

VARIÁVEL DEPENDENTE	**Taxa de Distorção Idade-Série Ensino Fundamental**			
	EF Não Espacial	EF Lag	EF Erro	EF Lag e Erro
	(1)	(2)	(3)	(4)
Saneamento	-0,965***	-0,2065***	-0,3114***	-0,2000**
	(0,0847)	(0,0312)	(0,0589)	(0,0296)
Observações	7,959	7,224	7,224	7,224
Número de municípios	4,347	3,612	3,612	3,612
Características socioeconômicas	SIM	SIM	SIM	SIM
Características da escola e professores	SIM	SIM	SIM	SIM
Saúde e educação	SIM	SIM	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM
Efeito fixo de ano	NÃO	NÃO	NÃO	NÃO
priv5anos - Variável Instrumental	SIM	SIM	SIM	SIM
Lambda	-	0,0806***	-	0,0842***
	-	(0,0166)	-	(0,0162)
Rho	-	-	0,1269	0.03
Sigma^2	-	-	-	56,5181

Erro-padrão entre parênteses
*** p<0,01; ** p<0,05; * p<0,1

Fonte: Elaboração própria a partir de dados da pesquisa.

No que diz respeito à taxa de distorção idade-série, os resultados também se mantiveram e indicaram que o maior acesso às condições adequadas de saneamento básico reduz a distorção idade-série. Além do acesso a esse serviço essencial manter as crianças na escola, ainda fornece melhores condições cognitivas, por estarem livres das

doenças relacionadas ao saneamento inadequado (DRSAI), de forma que sejam aprovadas e não sejam retidas em séries inadequadas à sua idade.

Neste caso, todos os modelos espaciais apresentaram coeficientes de menor magnitude em relação ao modelo não espacial da primeira coluna. No Quadro 4.3 estão dispostos os sinais e significâncias dos efeitos das variáveis de controle.

Variáveis Explicativas	IV EF NE	EF Lag	EF Erro	EF Lag e Erro
t_analfabetismo_25m	+++	+++	+++	+++
pop_negra_parda	---	+++	+++	+++
renda_per_capita	-	---	-	-
gini	+++	+++	+++	+++
pop_55mais	+++	+++	+++	+++
pop_15a29	+++	+++	+++	+++
bolsafam	+++	---	---	---
poptot	--			
urbanização	+++	--	-	-
alunos_turma_fundamental	+++	+++	+++	+++
docentes_com_ens_superior	+++	---	---	---
h_a_total_fundamental	---	---	---	---
índice infraestrutura tecnológica	---	---	---	---
índice infraestrutura básica	+++	---	--	--
porcent. escolas com diretoria		+++	+++	+++
porcent. escolas públicas	+++	++	-	-
desp_educ_cult_hab				
desp_saude_hab		+	-	+
imunização				

Quadro 4.3: Sinal e significância determinantes espaciais da distorção idade-série.
Fonte: Elaboração própria a partir de dados da pesquisa.

Para a taxa de distorção idade-série (TDI), o coeficiente da renda per capita, o índice de Gini e a taxa de analfabetismo não apresentaram reversão de sinais. No entanto, em relação à variável étnica referente à porcentagem da população negra e parda, quanto maior a porcentagem dessa população maior a taxa de distorção idade-série, revelando que as diferenças étnicas-raciais são relevantes para explicar o atraso escolar. Além disso, o maior número de famílias beneficiadas pelo programa Bolsa Família indicou menor taxa de distorção idade-série, ou seja, além das crianças frequentarem mais as escolas, o atraso escolar também diminui. Ainda que essa evidência não implique em qualidade do ensino, tais resultados revelam o alcance do programa em termos educacionais.

A taxa de urbanização também teve seu sinal alterado: de fato, regiões mais urbanizadas tendem a ter melhores indicadores educacionais que, nesse caso, refletem-se

em menores taxas de distorção idade-série. No tocante à estrutura etária da população, conforme a porcentagem da população adulta e da população de 55 anos ou mais aumenta, a taxa de distorção idade-série aumenta também, pois a probabilidade de atraso escolar é maior conforme aumenta a idade do estudante (PONTILI e KASSOUF, 2007).

A maior razão de alunos por turma leva a menor taxa de frequência escolar, para todos os modelos, indicando que turmas muito numerosas podem captar a deficiência do aprendizado individual, que levaria a reprovações e aumentos no atraso escolar. De forma contrária ao resultado encontrado para a taxa de frequência escolar, quando se considera a taxa de distorção idade-série, a porcentagem de docentes com ensino superior é importante para diminuir o atraso escolar entre as séries no ensino fundamental. Além disso, o índice de infraestrutura escolar tecnológica, o índice de infraestrutura escolar básica e o número médio de horas-aula também se mostraram fatores importantes para diminuir essa variável.

Neste caso, o resultado que não era esperado diz respeito à variável que representa a gestão da escola (porcentagem de escolas com diretoria): esta contribui de forma a aumentar o atraso dos alunos de ensino fundamental. O efeito da porcentagem de escolas públicas e das despesas com saúde por habitante, por terem apresentado sinais inconclusivos entre os três modelos, não foram consideradas na análise. Em linha com o formato das estimações espaciais das duas variáveis dependentes anteriores, a Tabela 4.12 mostra os resultados das estimações para a variável referente à taxa de abandono escolar da mesma população do ensino fundamental.

Tabela 4.12: Determinantes espaciais da taxa de abandono escolar (2000-2010).

VARIÁVEL DEPENDENTE	**Taxa de abandono escolar - ensino fundamental**			
	EF Não Espacial	EF Lag	EF Erro	EF Lag e Erro
	(1)	(2)	(3)	(4)
Saneamento	-0,499***	-0,0563***	-0,1223***	-0,0497**
	(0,0528)	(0,0148)	(0,0304)	(0,0142)
Observações	7,959	7,224	7,224	7,224
Número de municípios	4,347	3,612	3,612	3,612
Características socioeconômicas	SIM	SIM	SIM	SIM
Características da escola e professores	SIM	SIM	SIM	SIM
Saúde e educação	SIM	SIM	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM
Efeito fixo de ano	NÃO	NÃO	NÃO	NÃO
priv5anos - Variável Instrumental	SIM	SIM	SIM	SIM
Lambda	-	-0,0154***	-	-0,0108***
	-	(0,0247)	-	(0,0243)
Rho	-	-	0,1269	0,0489
Sigma^2	-	-	-	14,9124

Erro-padrão entre parênteses
*** p<0,01; ** p<0,05; * p<0,1

Fonte: Elaboração própria a partir de dados da pesquisa.

Novamente, a variável relativa ao maior acesso aos serviços de saneamento básico adequados impacta de forma negativa a taxa de abandono escolar. Ou seja, trata-se de resultado positivo em termos da melhora dos indicadores educacionais, haja vista os resultados também favoráveis obtidos anteriormente. Nota-se que, para essa variável, a magnitude dos coeficientes diminuiu de forma mais intensa. Ou seja, no modelo de Efeito Fixo não espacial o valor era de -0,49 e, com a introdução dos efeitos espaciais, esse valor passou para -0,056 e -0,049.

O abandono escolar está associado com altas taxas de reprovação e, consequentemente, com o atraso escolar e com menores ganhos no mercado de trabalho (EIDE e SHOWALTER, 2001). Por outro lado, Jacob e Lefgren (2009) sugerem que a reprovação no ensino fundamental não afeta a probabilidade do aluno concluir o ensino médio, porém a reprovação associada ao baixo rendimento no último ano do ensino fundamental aumenta substancialmente a probabilidade de que esses alunos abandonem os estudos no ensino médio. O fato dos municípios que apresentam adequadas condições de saneamento básico também apresentarem menores taxas de abandono escolar pode significar que, uma vez garantidas melhores condições de saúde às crianças, essas terão melhor rendimento escolar, o que pode implicar em menores taxas de abandono escolar.

O Quadro 4.4 apresenta os resultados referentes aos sinais e níveis de significância das variáveis explicativas após o controle pelos efeitos espaciais.

Variáveis Explicativas	IV EF NE	EF Lag	EF Erro	EF Lag e Erro
t_analfabetismo_25m	+++	+++	+++	+++
pop_negra_parda	---	+++	++	+++
renda_per_capita	---	---	---	---
gini	+++	+++	+++	+++
pop_55mais	+++			
pop_15a29	+++	+++	+	+++
bolsafam	+++	---	---	---
poptot	---			
urbanização	+++	+++	+++	+++
alunos_turma_fundamental	+	+++	+++	+++
docentes_com_ens_superior	+	---	---	---
h_a_total_fundamental	-	---	---	---
índice infraestrutura tecnológica	---	---	---	---
índice infraestrutura básica	+++	--		-
porcent. escolas com diretoria		+++	+++	++
porcent. escolas públicas	+++	+	+	
desp_educ_cult_hab				
desp_saude_hab				
imunização				

Quadro 4.4: Sinal e significância dos determinantes espaciais do abandono escolar.

Os resultados dos sinais e significância das variáveis explicativas para a variável dependente taxa de abandono escolar foram semelhantes à análise para a variável anterior (taxa de distorção idade-série), mas com duas exceções. Em relação à primeira delas, a maior taxa de urbanização gerou maior abandono. A segunda faz parte do vetor de características educacionais. Segundo resultado expresso no Quadro 4.4 acima, a maior porcentagem de escolas públicas mostrou-se favorável para o registro de maiores taxas de abandono escolar. Além disso, cabe observar que nenhum tipo de despesa com saúde, educação, cultura e nem a cobertura de imunização afetam estatisticamente as taxas de abandono escolar do ensino fundamental.

4.4 Comparação com o método de Máxima Verossimilhança.

Millo e Piras (2012), a partir de um exemplo de modelo sem endogeneidade, compararam os coeficientes estimados por Máxima Verossimilhança (MV) e pelo método

generalizado dos momentos (GMM). Segundo os autores, nesse caso, os parâmetros ficaram bem próximos, ressaltando pouca diferença entre eles. Espera-se que, pelo controle da endogeneidade, as estimativas dos coeficientes sejam diferentes nesse estudo.

Na Tabela 4.13 são comparados o modelo SAR de efeitos fixos estimado por GMM, com a identificação da variável endógena entre as explicativas além da variável dependente defasada, com o modelo SAR de efeitos fixos estimado por MV. As outras duas comparações, terceira e quarta colunas, referem-se aos modelos SAC de efeitos fixos estimados por GMM e MV para a taxa de distorção idade-série (TDI) por GMM. A mesma comparação foi feita para as duas últimas colunas considerando a taxa de abandono escolar.

Tabela 4.13: Comparação entre os modelos estimados por GMM e MV

VARIÁVEL DEPENDENTE	Frequência escolar (GMM)	Frequência escolar (MV)	TDI (GMM)	TDI (MV)	Abandono Escolar (GMM)	Abandono Escolar (MV)
	(1)	(2)	(3)	(4)	(5)	(3)
Saneamento	0,0351***	0,0241***	-0,20***	-0,23***	-0,049***	-0,058***
	(0,0130)	(0,0046)	(0,0296)	(0,0097)	(0,0142)	(0,0062)
Observações	7.224	7.224	7.224	7.224	7.224	7.224
Número de municípios	3.612	3.612	3.612	3.612	3.612	3.612
Características socioeconômicas	SIM	SIM	SIM	SIM	SIM	SIM
Características escola e professores	SIM	SIM	SIM	SIM	SIM	SIM
Saúde e educação	SIM	SIM	SIM	SIM	SIM	SIM
Efeito fixo de município	SIM	SIM	SIM	SIM	SIM	SIM
Efeito fixo de ano	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO
priv5anos - Variável Instrumental	SIM	NÃO	SIM	NÃO	SIM	SIM
Lambda	0,1428***	0,0180	0,0842***	-0,007	-0,0108***	0,019
	(0,0321)	(0,0125)	(0,0162)	(0,0118)	(0,0243)	(0,0147)
Rho	-	-	0,03	0,011	0,0489	0,0053
	-	-	-	(0,021)	-	(0,023)

Erro-padrão entre parênteses
*** p<0,01; ** p<0,05; * p<0,1

Fonte: Elaboração própria a partir de dados da pesquisa.

As estimativas dos coeficientes espaciais se mostraram diferentes, como era esperado, pois na estimação por MV tais coeficientes não foram significativos. Com exceção da variável de frequência escolar, a magnitude dos coeficientes estimados por MV foram ligeiramente maiores.

Conclusões

No Brasil, as elevadas taxas de incidência de internações por Doenças Relacionadas ao Saneamento Ambiental Inadequado (DRSAI), os insatisfatórios indicadores educacionais e de acesso aos serviços de saneamento básico apontam para uma realidade socioeconômica desigual e concentrada através das regiões. Apesar da evolução positiva de tais indicadores ao longo do tempo, os desafios a serem superados em tais áreas ainda são de grande magnitude. A identificação do efeito do acesso aos serviços de saneamento básico sobre os indicadores educacionais, por meio do impacto que tal acesso gera no perfil de morbidade da população em idade escolar, é uma contribuição desse trabalho para superar tais desafios.

A partir de dados municipais, o efeito do saneamento sobre educação foi obtido. Os dados municipais foram utilizados para identificar o efeito do maior percentual de pessoas que residem em domicílios que possuem banheiro e água canalizada por rede geral de distribuição sobre indicadores educacionais da população de 6 a 14 anos. A identificação desse efeito se valeu de análise não espacial e espacial para dados em painel (2000 e 2010) em dois estágios com variável instrumental.

Os resultados finais da identificação do efeito do saneamento sobre educação, após todos os procedimentos de estimação serem realizados, indicam que o efeito do aumento de 1% no acesso a saneamento está associado a um aumento de 0,11 pontos percentuais na taxa de frequência escolar, a uma queda de 0,31 pontos percentuais na taxa de distorção idade-série e a também a uma queda de 0,12 pontos percentuais na taxa de abandono escolar.

Os dados municipais utilizados foram obtidos de diversas fontes para os anos 2000 e 2010. A análise não espacial de dados em painel com modelo de efeitos fixos, ao utilizar variável instrumental em dois estágios, indicou que saneamento importa para educação. Um aumento de 1% no acesso aos serviços de saneamento básico está associado a um aumento de 0,11 pontos percentuais na taxa de frequência escolar da população de 6 a 14 anos e a diminuições de 0,49 pontos percentuais na taxa de abandono escolar e de 0,96 pontos percentuais na taxa de distorção idade-série no ensino fundamental para essa mesma população.

Os testes dos resíduos do modelo não espacial, que geraram os resultados acima, foram implementados para identificar possíveis efeitos espaciais. Os resultados de tais testes indicaram erros autocorrelacionados espacialmente para as variáveis de distorção idade-série e abandono escolar e evidências da presença de autocorrelação espacial na forma de

defasagem espacial da variável dependente para as três variáveis de educação (incluindo frequência escolar). Dado que ignorar os efeitos espaciais pode gerar viés e ineficiência das estimativas dos coeficientes, o método mais adequado para o controle dos efeitos espaciais, neste caso, é o método das Variáveis Instrumentais ou dos Momentos Generalizados (GMM). A vantagem deste método é estimar consistentemente os modelos espaciais que apresentam variáveis endógenas entre as variáveis explicativas, além da defasagem espacial, e aqueles que também apresentam erros espacialmente correlacionados.

Na quase totalidade dos casos, os Modelos de Erro Autoregressivo Espacial (SEM), de Defasagem Espacial (SAR) e o de Defasagem Espacial com Erro Autoregressivo Espacial (SAC) estimados por GMM apresentaram menor magnitude para as estimativas dos coeficientes em relação aos modelos não-espaciais, igualmente à evidência obtida por Baylis et al. (2011). Porém, por ser área recente de estudo e poucos trabalhos terem usado esse método, não foi possível comparar os efeitos marginais dos modelos não espaciais com os modelos SAR e SAC, por uma limitação computacional do método[18]. Dessa forma, para esses dois últimos modelos não foi factível precisar o impacto do saneamento sobre educação somente com os resultados dos coeficientes, pois o parâmetro positivo e significativo (ρ), que mede o *spillover* das variáveis de educação, é capaz de potencializar o impacto de uma variável independente sobre a variável dependente no efeito marginal. Com as informações obtidas, foi possível concluir somente sobre a direção do sinal.

Frente a essa situação, a escolha pelo modelo espacial recaiu sobre o modelo de erro (SEM) para as variáveis de abandono escolar e distorção idade-série. Pelo fato desse modelo não possuir o termo que determina o *feedback* na variável dependente, os parâmetros estimados sob tal modelo possuem interpretação direta. Além disso, teoricamente é razoável supor, pela literatura apresentada, que os indicadores educacionais apresentam respostas distintas de acordo com as diversas realidades socioeconômicas dos estados e regiões. Apesar dos problemas causados pela heterogeneidade espacial poderem ser abordados por técnicas econométricas não espaciais, considerar a estrutura espacial dos dados torna os procedimentos de estimação mais eficientes. Além disso, vários sinais e significância dos coeficientes das variáveis

[18]Para o GMM, o cálculo do efeito marginal quando existe endogeneidade entre as explicativas não está implementado no *software* R ou em outro programa estatístico.

explicativas deixaram de ser inconclusivos, como eram na estimação do modelo não espacial.

Para a variável de frequência escolar, o modelo escolhido é o não-espacial, pelo fato de não existir evidências de efeitos espaciais sobre a forma de autocorrelação dos resíduos. Logo, o efeito do aumento de 1% no acesso a saneamento está associado ao aumento de 0,11 pontos percentuais na taxa de frequência escolar e a quedas de 0,31 e 0,12 pontos percentuais na taxa de distorção idade-série e abandono escolar, respectivamente.

Em relação à sensibilidade dos efeitos para as populações das duas diferentes faixas etárias, pela magnitude das estimativas dos parâmetros dos efeitos marginais, a população de zero a 14 anos é a mais atingida pelo acesso a condições inadequadas dos serviços de saneamento básico, o que confirma uma das hipóteses estabelecidas nesse estudo, de que essa população é a mais frágil e a mais suscetível a contrair as DRSAI. Nesse sentido, outra evidência que pode ser destacada refere-se ao modelo não espacial, quando se desagrega o ensino fundamental em dois ciclos, sendo o primeiro formado pela população de 6 a 10 anos e o segundo constituído pela população de 10 a 14 anos. Comparativamente, o primeiro ciclo apresenta maior redução na taxa de abandono escolar (0,52 pontos percentuais, contra 0,47 pontos percentuais do segundo ciclo). Para a taxa de distorção idade-série, os resultados foram similares. Em que pese a pequena diferença quantitativa, essa evidência vai ao encontro do resultado anterior. A população mais jovem, de seis a dez anos, está mais sujeita às DRSAI do que os alunos mais velhos.

Nesse contexto, políticas públicas voltadas para a mitigação das condições precárias de saneamento básico devem levar em consideração os impactos negativos que tais condições exercem sobre a saúde dos indivíduos, em particular das crianças, e os impactos educacionais decorrentes das dificuldades que as mesmas têm de se manter na escola, avançar nos estudos e não desistir dos mesmos, por terem a saúde comprometida em razão de condições impróprias de abastecimento de água, coleta e tratamento de esgoto.

Referências Bibliográficas

ALBERNAZ, A; FERREIRA, F.H.G.; FRANCO, C. *Qualidade e equidade no ensino fundamental brasileiro*. **Pesquisa e Planejamento Econômico**, v.32, n.3, p.453-476, 2002.

ALMEIDA, E. *Econometria Espacial Aplicada*, Campinas, SP: Editora Alínea, 2012.

ANSELIN, L.; BERA, A. *Spatial dependence in linear regression models with an introduction to spatial econometrics*. In: ULLAH A.; GILES D. E. (eds). **Handbook of applied economic statistics**, Nova York: Marcel Dekker, p. 237-289, 1998.

ARAÚJO, B. S.; SANTOS, J. F.; NEIVA, T. S.; MAGALHÃES-FILHO, R. R. e RIOS, D. S. *Associação das parasitoses intestinais com anemia e eosinofilia em escolares do povoado de Matilha dos Pretos*, Feira de Santana, Bahia, Brasil. **Sitientibus Série Ciências Biológicas**, 9 (1): 3-7, 2009.

BALTAGI, B. H.; Liu L. *Instrumental Variable Estimation of a Spatial Autoregressive Panel Model with Random Effects*, **Economics Letters**, 111, 135-137, 2011.

BANERJEE, A. V.; DUFLO, E. *A. Economia dos Pobres*, 1 ed, 2011.

BARROS, R. P. de; MENDONÇA, R. Os determinantes da desigualdade no Brasil. A economia brasileira em perspectiva: 1996. Rio de Janeiro, **IPEA**, v. 2, p. 421- 474, 1996.

BARROS, R.P. de; MENDONÇA, R.; SANTOS, D.D.; QUINTAES, G. Determinantes do desempenho educacional no Brasil. Rio de Janeiro: **IPEA**, 33p., Texto para discussão, n. 834, 2001.

BAYLIS, K, PAULSON, N. E PIRAS, G. *Spatial Approaches to Panel Data in Agricultural Economics: A Climate Change Application*, **Journal of agricultural and applied economics**, v.3, n.43, p.325-338, 2011.

BISHOP, J. *Incentives for learning: Why American high school students compare so poorly to their counterparts overseas*. Ithaca, NY: Cornell University, **Center for Advanced Human Resource Studies**, 1989.

BLEAKLEY, H. *Disease and Development: Evidence from Hookworm Eradication in the American South*, **The Quarterly Journal of Economics**, 122 (1): 73-117, 2007.

BORJAS, G. *Ethnicity, Neighborhoods, and Human-Capital Externalities*. **American Economic Review**, v.3, n.85, p.365-390, 1995.

BRUECKNER, J. K. *Strategic Interaction Among Governments: An Overview of Empirical Studies*. **International Regional Science Review**, Champaign, IL, 2003.

BRITO, L. L.; Mauricio, L. B.; SILVA, R. C. R.; ASSIS, A. M. O.; REIS, M. G.; PARRAGA, I. e BLANTON, R. E. *Fatores de risco para anemia, por deficiência de ferro em crianças e adolescentes parasitados por helmintos intestinais*. **Revista Panam Salud Publica,** Pan AM J Public Health, 14(6), 2003.

CAIRNCROSS, S.; FEACHEM, R. G. *Environmental health engineering in the tropics: an introductory text*. John Wiley & Sons, Chichester, 1990.

CAIRNCROSS, S.; R. FEACHEM. *Environmental Health Engineering in the Tropics*. 2. ed. Chichester, U.K.: John Wiley & Sons, 1993.

CARD, D.; KRUEGER, A. B. *Does school quality matter? returns to education and the characteristics of public schools in the United States*, **The Journal of Political Economy**, v. 100, n. 1, 1-40, 1992.

CARNEIRO, P., CUNHA, F. e HECKMAN, J. *Interpreting The Evidence of Family Influence on Child Development*, **Economics of early childhood conference**, Minneapolis Fed, 2003

CARVALHO, S.; FIRPO, S. *O regime de Ciclos de aprendizagem e a Heterogeneidade de seus efeitos sobre a proficiência dos alunos*. **Revista de Economia Aplicada**, v. 18, n. 2, pp. 199-214, 2014.

CHAGAS, ANDRÉ L. S.; ALMEIDA, ALEXANDRE N.; AZZONI, CARLOS R. *Sugar Cane Burning and Human Health: a Spatial Difference-in-Difference Analysis*, Department of Economics – FEA/USP. **Working Paper Series** Nº 2014-20. 2014.

COLEMAN, J. S. et al. *Equality of educational opportunity*. Washington, 1966.

CRAGG, J. G., e DONALD, S. G., *Testing Identifiability and Specification in Instrumental Variable Models*, **Econometric Theory**, 9, 222–40., 1993.

CURI, A.; MENEZES-FILHO, N. *A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil*, **Estudos Econômicos**, São Paulo, vol.39 no.4, São Paulo, 2009.

CUTLER, D., & MILLER, G. *The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States*. **Demography**, 42 (1), 1-22, 2005.

CARD, D. *Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems*. **Econometrica**, Vol. 69, No. 5, p. 1127-1160, 2001.

DAVIDSON, R. and MACKINNON, J. G., *Econometric Theory and Methods*, Oxford University Press, 2004.

DOMENE, S. M. A. *O Papel do ferro sobre a nutrição e a saúde*. Serviço de Informação da Carne, **Comitê Técnico do SIC**, PUC, Campinas, 2004.

DUFLO, E. *Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment*. **American Economic Review**, Vol. 91 (4), pp 795-813, 2001.

DUFLO, E.; MICHAEL GREENSTONE, RAYMOND GUITERAS, AND THOMAS CLASEN. *Toilets Can Work: Short and Medium Run Health Impacts of Addressing Complementarities and Externalities in Water and Sanitation*, **NBER Working Paper,** No. 21521, 2015.

EIDE, E. R.; SHOWALTER, M. H. *The effect of grade retention on educational and labor market outcomes*. **Economics of Education Review**, v. 20, pp. 563–576, 2001.

ELHORST, J. *Applied spatial econometrics: Raising the bar*. **Spatial Economic Analysis**, v.5, n.1, p.9-28, 2010.

ELHORST, J. P. *Spatial econometrics: from cross-sectional data to spatial panels. Heidelberg*, New York, Dordrecht, London: Springer-Verlag. 2014.

FERRÃO, M. E. et. al. *Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz*. **Revista Brasileira de Estudos de População**, v. 18, n.1/2, jan/dez, 2001.

FEWTRELL, L., KAUFMANN, R. *Water Sanitation and Hygiene Interventions to Reduce Diarrhoea in Developed Countries: A Systematic Review and Meta-Analysis*. **Lancet Infectious Diseases**, 5, 2005.

FRANCO, A.M de P; MENEZES FILHO, N.A. Os determinantes do aprendizado com dados de um painel de escolas do SAEB. **Encontro ANPEC**, 2009**.**

GALVÃO JR., A. C. et al. *Marcos regulatórios estaduais em saneamento básico no Brasil*. **Revista de Administração Pública**, Rio de Janeiro, v. 43, n. 1, p. 207-227, 2009.

GERUSO, M.; SPEARS, D. *Neighborhood Sanitation and Infant Mortality*, **National Bureau of Economic Research Working Paper Series**, n. 21184, 2015.

GOLGHER, A. B. *Introdução à Econometria Espacial*, Jundiaí, Paco Editorial: 2015.

GREENE, W. H. *Econometric Analysis*, 7 Edition, Prentice Hall, 2003.

GURGEL, R. Q. et al. *Creche: ambiente expositor ou protetor nas infestações por parasitas intestinais em Aracaju*, *SE*. **Revista da Sociedade Brasileira de Medicina Tropical**, Uberaba, v. 38, n. 3, p. 267-269, maio/jun. 2005.

HANUSHEK, E. *The Economics of Schooling: Production and Efficiency in Public Schools*, **Journal of Economic Literature**, v. 24, No. 3, p 1141-1177, 1986.

HANUSHEK, E. e KIMKO, D. *Schooling, Labor-Force Quality, and the Growth of Nations*, **The American Economic Review**, 90, 1184-1208, 2000.

HANUSHEK, E. A.; RIVKIN, S. G. *Teacher Quality* in: **Handbook of the Economics of Education**, v.2, Amsterdam: North Holland, pp. 1052-1078, 2006.

KELEJIAN, H. H.; PRUCHA, I. R. *Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model*. **International Economic Review**, Vol. 40, No. 2, pp. 509-533, 1999.

HAUSMAN, J. A. *Specification Tests in Econometrics*, **Econometrica**, Vol. 46, No. 6, p. 1251-1271, 1978.

HECKMAN, J. *The Lessons from technology of skill formation*, **NBER Working Paper**, 11142, 2005.

HECKMAN, J., LAYNE-FARRAR, A.; TODD, P. *The schooling quality-earnings relationship: using economic theory to interpret functional forms consistent with the evidence*. **National Bureau of Economic Research Working paper**, n. 5288, 1996.

JACOB, B. A.; LEFGREN, L. *The Effect of Grade Retention on High School Completion*. **CLOSUP Working Paper Series**, n. 12, 2009.

KAPOOR M, KELEJIAN H.H, PRUCHA I.R. *Panel Data Model with Spatially Correlated Error Components*. **Journal of Econometrics**, 140(1), 97-130, 2007.

KASSOUF, A. L. *Saneamento e educação: bens substitutos ou complementares*, **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.25, n.2, p. 359-372, ago 1995.

KASSOUF, A. L. *Trabalho infantil: escolaridade x emprego*, **Economia**, v. 2, n. 2, p. 549-586, jul./dez. 2001.

KELEJIAN, H. H.; PRUCHA, I. R. *A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances*. **J. Real State Finance Econ**., Dordrecht, v.17, n.1, p.99-121, 1998.

KELEJIAN, H. H.; PRUCHA, I. R. *A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model*, **International Economic Review**, Vol. 40, No. 2, pp. 509-533, May, 1999.

KUNZ, J. M. O.; VIEIRA, A. S.; VARVAKIS, T.; GOMES, G. A.; ROSSETO, A. L.; BERNARDINI, O. J.; ALMEIDA, M. S. S.; ISHIDA, M. M. I. *Parasitas intestinais em crianças de escola municipal de Florianópolis, SC*, Educação ambiental e em saúde, **Biotemas**, v. 21, n. 4, p. 157-162, 2008.

LUDWIG, K. M.; FREI, F.; ALVARES FILHO, F.; RIBEIRO-PAES, J. T. *Correlação entre condições de saneamento básico e parasitoses intestinais na população de Assis*, Estado de São Paulo, **Rev. Soc. Brasileira de Medicina Tropical**, 32(5):547-55, 1999.

MACEDO, G. A. *Fatores associados ao rendimento escolar de alunos da 5ª série (2000): uma abordagem do valor adicionado e da heterogeneidade*. 124 f. Dissertação (Mestrado em Demografia), Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2004

MARA, D. D.; FEACHEM, R. G. A. *Water and excreta related diseases: unitary environmental classification*. **Journal of Environmental Engineering**, n.125, 1999.

MENEZES-FILHO, N.A. *Os determinantes do desempenho escolar no Brasil. Instituto Futuro Brasil*, IBMEC São Paulo e Faculdade de Economia e Administração da Universidade de São Paulo. Sumário Executivo. 2007.

MENEZES-FILHO, N. A. *Educação e desigualdade*. In: MENEZES-FILHO, N. A.; LISBOA, M. (Eds.). **Microeconomia e sociedade**. Rio de Janeiro: Contracapa, 2001.

MENEZES-FILHO, N.; VASCONCELLOS, L.; WERLANG, S. *Avaliando o impacto da progressão continuada no Brasil*. São Paulo: **Fundação Itaú Social**, 2008.

MIGUEL,E; KREMER, M. *Worms: identifying impacts on education and health in the presence of treatment externalities*, **Econometrica**, 72 (1), 159-217, jan 2004.

MILLO E PIRAS, *splm: Spatial Panel Data Models in R*, **Journal of Statistical Software**, v. 47, Issue 1, 2012.

MILLIMET, D. L.; RANGAPRASAD, V. *Strategic Competition Amongst Public Schools*, **Regional Science and Urban Economics**, 37, 199-219, 2007.

MUTL, J.; PFAFFERMAYR, M. *The Hausman test in a Cliff and Ord panel model*, **Econometrics Journal**, volume 14, pp. 48–76, 2011.

NAIDOO A.G.V, VAN EEDEN, MUNCH Z 2014. *Spatial variation in school performance, a local analysis of socio-economic factors in Cape Town*. **South African Journal of Geomatics**, 3(1): 78-94, 2014.

NERI *et. al.*, TRATA BRASIL: *Saneamento, Educação, Trabalho e Turismo*, 2008.

NEVES, R.; PAZELLO, E. T. *O efeito de políticas de não repetência sobre o desempenho dos estudantes do Ensino Fundamental*. In: Anais 40 **Encontro Nacional de Economia**, 2012, Porto de Galinhas - RE.

PONTILI, R. M.; KASSOUF, A. L. *Fatores que afetam a frequência e o atraso escolar, nos meios urbano e rural, de São Paulo e Pernambuco*, **Revista de Economia e Sociologia Rural**, v. 45, n. 1, Brasília, jan/mar 2007.

PRADO, M. S.; BARRETO, L. M.; STRINA, A.; FARIA, J. A. S.; NOBRE, A. A. e JESUS, S. R. *Prevalência e intensidade da infecção por parasitas intestinais em crianças na idade escolar na Cidade de salvador (Bahia, Brasil)*. **Revista da sociedade Brasileira de Medicina Tropical**, 34 (1): 99-101, Jan/Fev 2001.

RIVKIN, S. G.; HANUSHEK, E. A.; KAIN, J. F. *Teachers, schools, and academic achievement*. **Econometrica**, v. 73, n. 2, p. 417-458, 2005.

RODRIGUEZ-POSE, A.; TSELIOS, V. *Education and Income Inequality in the Regions of the European Union,* **Spatial Economics Research Centre** (SERC), SERC Discussion paper 11, November 2008.

UMAR, H. M. Regional Inequality of Educational Attainment in Nigeria, ,**International, British Journal of Economics**, Management & Trade, 4(3): 420-430, 2014.

SAIANI, C. C. S.; AZEVEDO, P. F. Privatização como estratégia política: evidências para o saneamento básico brasileiro. In: ENCONTRO NACIONAL DE ECONOMIA, 40, 2012.Anais… Porto de Galinhas, Pernambuco: ANPEC, 2012.

SCHADY, N. *Early childhood development in Latin America and the Caribbean*, **World Bank Policy Research**, Working Paper 3869, March 2006.

SEROA DA MOTTA, R. As opções de marco regulatório de Saneamento no Brasil. **Plenarium**, Brasília, n.3, p. 100-116, set. 2006.

STOCK, J. H.; YOGO, M. *Testing for weak instruments in linear IV Regression. Identification and Inference for Econometric Models*, Cambridge, **UK. Cambridge University Press**, 2005.

PIRAS, G. *Estimation of Random Effects Spatial Panel Data Models: Some Additional Evidence*. **Unpublished Manuscript**, 2011.

PLENZLER, N. L. *Student performance and educational resources: a spatial econometric examination*, Theses and Dissertations, **The University of Toledo Digital Repository**, 2004.

TOPEL, R. *Labor Markets and Economic Growth* in: **Handbook of Labor Economics**, Amsterdam: Elsevier Science B.V., 2943-29, 1999.

VASCONCELLOS, C. dos S. *Avaliação da aprendizagem: práticas de mudanças por uma práxis transformadora*, 5. ed., São Paulo: **Libertad**, 2003.

WATSON, T. *Public Health Investments and the Infant Mortality Gap: Evidence from Federal Sanitation Interventions on U.S. Indian Reservations*, **Journal of Public Economics**, 90: 1537-1560, 2006.

WOOLDRIDGE, J. M. *Introductory Econometrics*. Pioneira Thompson Learning, 20ª ed., 2002.