

CAMILA GONZAGA MIZOKAMI

ESTIMANDO O VALOR DA EDUCAÇÃO PARA AS FAMÍLIAS BRASILEIRAS

Monografia apresentada ao Departamento de Economia, Universidade de São Paulo, como requisito para a obtenção do título de Bacharel em Ciências Econômicas.

Orientador: Prof. Dr. Naercio Aquino Menezes Filho
Coordenador: Prof. Dr. Luís Eduardo Negrão
Meloni

SÃO PAULO

2022

FICHA CATALOGRÁFICA

Gonzaga Mizokami, Camila

Estimando o valor da educação para as famílias brasileiras – São Paulo, 2022.

Nº de páginas: 30

Área de concentração: Economia da Educação

Orientador: Prof. Dr. Naercio Aquino Menezes Filho.

Monografia (Bacharelado) – Universidade de São Paulo, Faculdade de Economia, Administração, Contabilidade e Atuária (FEA USP), Departamento de Economia.

1. Escolha escolar; 2. Deslocamento; 3. Desempenho escolar

Aos meus pais, Nivia e André, minhas maiores referências...

Dedico

AGRADECIMENTOS

À Universidade de São Paulo e à sociedade, pelas diversas oportunidades que a universidade pública me proporcionou ao longo dos anos de graduação e me fizeram crescer acadêmica e pessoalmente.

A todos os funcionários e funcionárias da Universidade de São Paulo, que trabalham para proporcionar as condições necessárias para o estudo na instituição.

À FEA Social, entidade estudantil fundamental para a minha trajetória na FEA USP e uma das melhores escolhas que fiz.

A todos os amigos e amigas com quem tive a sorte de cruzar o caminho na FEA USP. Vivemos juntos histórias que se tornaram minhas melhores memórias.

A todos os amigos e amigas que contribuíram para o presente trabalho, fornecendo essenciais palavras de incentivo e me nutrindo com apoio.

Aos meus pais, por me apoiarem incansavelmente ao longo da vida e por serem o alicerce da minha formação enquanto indivíduo e estudante.

Ao Prof. Naercio, pela orientação atenciosa, pelo encorajamento e pela inspiração para realizar a pesquisa no campo da educação, além dos ensinamentos que extrapolam o assunto acadêmico.

“Só existirá democracia no Brasil no dia em que se montar no país a máquina que prepara as democracias. Essa máquina é a da escola pública.”

Anísio Teixeira.

SUMÁRIO

RESUMO.....	vi
ABSTRACT	vii
1 INTRODUÇÃO	8
1.1 CONTEXTO E JUSTIFICATIVA	8
1.2 OBJETIVO	9
2 REVISÃO DE LITERATURA.....	10
3 DADOS	15
4 MODELO ECONOMÉTRICO.....	19
4.1 MODELOS PROPOSTOS	19
4.2 ESTATÍSTICAS DESCRIPTIVAS.....	21
5 RESULTADOS	23
6 CONCLUSÕES.....	29
REFERÊNCIAS.....	30

RESUMO

ESTIMANDO O VALOR DA EDUCAÇÃO PARA AS FAMÍLIAS BRASILEIRAS

O objetivo deste trabalho é avaliar como a diferença da qualidade de escolas em municípios vizinhos se correlaciona com a porcentagem de alunos que se descolam para estudar no município vizinho. Foram usados dados de municípios do estado de São Paulo e de escolas públicas. Por meio de análise econômétrica, foi avaliada a magnitude da correlação entre a diferença de notas médias da prova Saeb e Enem de dois municípios e a porcentagem de pessoas que se desloca para o município com escolas que possuem maiores notas. Os resultados mostram que a diferença de notas tem correlação positiva com a porcentagem de deslocados. O resultado é ainda mais expressivo para o ensino médio, no qual nota-se coeficiente estatisticamente significante ao nível de 1%. Os resultados indicam que um aumento de 10 pontos na diferença de notas médias do Enem entre dois municípios está associado a um aumento de 0,93% no número de alunos que se deslocam do seu município de residência, *ceteris paribus*. Portanto, os resultados indicam que as famílias valorizam a qualidade acadêmica como critério para escolha da escola de seus filhos, mesmo que isso incorra em custos de deslocamento.

Descritores: Preferências na escolha escolar; Deslocamento diário intermunicipal; Qualidade acadêmica

Código JEL: I21, C01, C13

ABSTRACT

ESTIMATING THE VALUE OF EDUCATION FOR BRAZILIAN FAMILIES

The purpose of this paper is to evaluate how the difference of school effectiveness between neighboring cities is correlated with the percentage of people who commute in order to study in the nearest city. Data from public schools and cities in the state of São Paulo were used. Through econometric analysis, it was evaluated the magnitude of the correlation between the difference in average scores (from the 'Saeb' and 'Enem' tests) of two cities and the percentage of people who commuted to the city that has schools with higher scores. The results show the difference of grades is positively correlated with the percentage of commuted students. They are even more expressive for high school, in which the coefficient is statistically significant at 1%. The results indicate that an increase of 10 points in the difference of average 'Enem' scores between two cities is associated with an increase of 0.93% of people who commute from their home city in order to study, *ceteris paribus*. Therefore, the results point out to the fact that Brazilian families from São Paulo value school effectiveness as a criterion of the school choice process, even if this incurs commuting expenses.

Key words: Preferences of parental school choice; Daily commute between cities; School effectiveness

JEL classification: I21, C01, C13

1 INTRODUÇÃO

1.1 CONTEXTO E JUSTIFICATIVA

Uma ampla gama de literatura internacional tem a análise direcionada às preferências das famílias em relação à escolha de escolas para seus filhos. Os estudos analisam dados em diferentes localidades do globo sobre como essa decisão é tomada e quais os fatores que mais a influenciam. Além disso, a qualidade do sistema educacional e os fatores que podem estar associados a melhores desempenhos acadêmicos também têm atraído crescente interesse no século XXI (NGWARE ET AL., 2011). Dentre os determinantes da escolha, o fator espacial tem sido alvo de estudos por ter impacto significativo, afetando os resultados da relação positiva entre demanda e efetividade das escolas.

No Brasil, 7,4 milhões de pessoas deslocam-se diariamente para fins de estudo, de acordo com o estudo "Arranjos Populacionais e Concentrações Urbanas do Brasil", feito com base nos dados do Censo Demográfico de 2010. Um outro indicativo da relevância da questão é a mudança promovida pelo Instituto Brasileiro de Geografia e Estatística (IBGE) de 2000 para 2010 na coleta de dados do Censo, que se baseia em separar os dados sobre deslocamento pendular intermunicipal para os fins de trabalho e de estudo, os quais antes eram agregados. Essa mudança reflete a importância de analisar focalmente os dados de deslocamento para estudo e compreender o que isso nos diz sobre a educação brasileira.

Apesar disso, o tema ainda é pouco explorado pela literatura nacional. A presente pesquisa pretende, portanto, inserir o estudo dos fatores determinantes para a escolha de escolas pelas famílias no contexto brasileiro e, mais especificamente, compreender se há uma correlação positiva entre a decisão de deslocamento e a busca por escolas de melhor qualidade acadêmica. O valor que os pais atribuem à educação será analisado pela ótica dos custos (monetários e de tempo) que os mesmos estão dispostos a pagar pelo deslocamento entre escola e moradia, ao escolher escolas distantes de suas residências, mas que são mais efetivas em promover alto desempenho acadêmico dos alunos. Dessa forma, os resultados podem contribuir para o estudo das preferências na escolha de escolas e discussões referentes a políticas de democratização de um sistema de ensino público de qualidade no Brasil.

1.2 OBJETIVO

A pesquisa visa estimar o valor atribuído à educação pelas famílias brasileiras por meio da análise dos deslocamentos intermunicipais realizados pelos estudantes do Ensino Básico público entre o local de suas moradias e de seus estudos. Para tal, é necessário investigar se a diferença de desempenho acadêmico das escolas, mensurado pelas notas da Prova Saeb e Enem, impacta a quantidade de pessoas que optam por se deslocar para outro município para fins de estudo. Ou seja, comparando dois municípios, o de residência e o seu mais próximo, será investigado se o fato de este ter notas maiores do que aquele incentiva os pais a deslocarem seus filhos, apesar dos custos implicados pelo deslocamento moradia-escola.

Ademais, objetiva-se entender o impacto de outros possíveis determinantes da escolha como, por exemplo, características das escolas e da qualificação dos docentes. Dentre as características das escolas, algumas hipóteses de correlação positiva com a escolha são acesso à internet e equipamentos multimídia, existência de laboratórios (informática e ciência) e quadra de esportes, oferecimento de curso profissionalizante complementar e alimentação aos alunos. Possíveis fatores relacionados à qualificação dos docentes e correlacionados positivamente com a decisão da escola são: conclusão do Ensino Superior, de especialização, mestrado e doutorado, além de realização de modalidade de licenciatura na primeira graduação.

Outros fatores relevantes a serem considerados como controle na análise são o PIB per capita, o tipo de zona (rural ou urbana) do município de residência, e a distância percorrida no deslocamento. O recorte utilizado para a pesquisa foram os municípios de São Paulo, estado mais populoso do Brasil, com 46,6 milhões de habitantes, concentrando 21,9% da população total do país, segundo dados do Governo Federal.

2 REVISÃO DE LITERATURA

Jacob e Lefgren (2007) examinaram as preferências dos pais sobre a educação usando, pela primeira vez, informações vindas de dentro das escolas sobre as escolhas. Os dados utilizados incluem o número de solicitações de pais por professores particulares do ensino primário juntamente com informações sobre os atributos desses professores, como relatórios de diretores sobre as características que, geralmente, não são observáveis (por exemplo, se o professor é visto como um exemplo a ser seguido pelos alunos). O estudo concluiu que, em média, os pais preferem professores descritos como bons em promover a satisfação dos estudantes e dão relativamente menor valor à habilidade do professor de aumentar o padrão de desempenho em matemática ou leitura.

Entretanto, é necessário fazer aqui uma diferenciação, famílias com menor poder aquisitivo e pertencentes a minorias valorizam mais o desempenho do estudante e são praticamente indiferentes à avaliação do diretor em relação à habilidade do professor de promover a satisfação dos estudantes; ao contrário do que ocorre em famílias com maior poder aquisitivo e não pertencentes a minorias. Esses fatos são consistentes com a utilidade marginal decrescente de aprendizado básico de matemática e de leitura. Isso sem contar que os pais com menor renda e pertencentes a minorias são muito menos prováveis de solicitar qualquer professor particular.

Assim, o que os pais esperam da escola depende não apenas das preferências deles, mas também das circunstâncias familiares. Esse é um indicativo da possibilidade de que pais com distintos níveis de renda apresentem diferenças em relação a programas de educação particular, mesmo que ambos tenham essencialmente a mesma função utilidade.

Jacob e Lefgren (2007) concluem que as famílias de baixa renda não só valorizam fortemente o desempenho acadêmico, como também são capazes de identificar professores de alta qualidade. Os resultados sugerem, assim, que a homogeneidade racial e social das escolas talvez não seja reflexo de uma preferência pela segregação por si só, mas uma consequência do interesse de cada grupo por um tipo específico de currículo ou pedagogia.

É importante ressaltar que o estudo considera as decisões dos pais condicionadas à escolha da escola já feita, ou seja, os pais podem priorizar outros

fatores na escolha da escola como a proximidade escola-casa, mas ainda atribuir maior valor ao desempenho do estudante quando solicitam por um professor particular. De acordo com Hanushek et al. (2005), a variação da qualidade dos professores é maior dentro de uma mesma escola do que entre escolas, o que explicaria o fato dos pais priorizarem o critério da distância na escolha da escola.

Uma das limitações da pesquisa é ser baseada apenas nos pais que solicitam um professor, ou seja, uma parcela de aproximadamente 30% que é mais engajada e presente na educação dos filhos ou que tem condições financeiras para tal.

Outros estudos, por sua vez, foram direcionados às consequências geradas pela escolha dos pais de forma agregada, representada pela demanda de matrículas em cada escola. Abdulkadiroğlu et al. (2020) analisa se as escolhas dos pais por escolas geram uma pressão, pelo lado da demanda, por melhoria na produtividade e efetividade do ensino. O estudo é baseado em escolas públicas de ensino médio de Nova York. A conclusão é de que as escolhas são baseadas na comparação de desempenhos e qualidades, mas não na relação causa-efeito de cada escola para aumentar o desempenho acadêmico (efetividade causal). Os resultados sugerem, assim, que a escolha dos pais é improvável de aumentar a eficiência alocativa.

Além do impacto sobre a qualidade do ensino, uma outra consequência da escolha dos pais por escolas é a estratificação social e racial (GLAZERMAN, 1998). Primeiramente, Glazerman (1998) traz à tona a discussão a respeito do impacto da expansão das escolhas dos consumidores, ou seja, de um mercado mais competitivo na área de educação, para o aprimoramento da qualidade das escolas. Usando dados de um programa de escolha de escolas públicas em Minneapolis (MN), o estudo traz evidências de que a escolha do consumidor por si só não aumenta a performance acadêmica das escolas, corroborando com os resultados de Abdulkadiroğlu et al. (2020). Os pais em Minneapolis são mais propensos a escolher escolas relativamente mais perto de suas casas e escolas onde serão melhor representados étnico e racialmente. Assim, os pais usam esses critérios para a escolha, em detrimento de critérios que levam em conta boas notas ou relevante valor adicionado.

Por outro lado, os próprios estudantes procuram escolas que tenham desempenho igual ou ligeiramente maior do que o deles mesmos, o que pode gerar uma segregação por performance. Do mesmo modo que a procura se baseia naquilo que é

semelhante e que gera sentimento de identificação, também é possível notar a segregação por raça, etnia e posição social, o que seria agravado caso tivesse maior competitividade de escolas no mercado.

Glazerman (1998) agrupa as variáveis essenciais para a explicação das escolhas: custos, informação, atributos da escola e atributos das demais escolas. Como os custos não são explícitos e não há o custo da escola em si, por se tratar do ensino público, é considerado o custo do valor do tempo despendido andando ou indo de ônibus para a escola.

Em relação a essa variável de deslocamento pendular, a análise é aprofundada. Por volta de 66% dos pais em Minneapolis escolheram uma escola que não era vizinha, tampouco perto de casa. Mais de um quarto dos pais escolheu escolas a mais de 2 milhas (equivalente a 3,2 km) de distância, alguns escolheram escolas a 6 milhas (equivalente a 9,7 km). Na organização americana de distritos escolares, considera-se que qualquer distância até uma milha é possível de ser percorrida a pé e há oferta de transporte de ônibus gratuito para distâncias maiores que essa.

O autor modela os custos indiretos, usando a distância em milhas entre a escola e a moradia como representativa do tempo de deslocamento. Para ele, é esperado que famílias com maior nível de renda e com apenas um dos pais (mãe/pai solteiro) tenham maiores custos de tempo e paguem uma penalidade maior na função utilidade por escolherem escolas mais longes (mas ainda dentro de uma milha de distância). Já para escolas fora do perímetro da “distância a pé”, a conjuntura é diferente, pois todas as famílias têm aproximadamente a mesma penalidade para suas funções utilidade.

Além disso, os resultados apontam para a possível ocorrência de incentivo para escolher escolas mais distantes pelo fato de se tornarem elegíveis para usufruir gratuitamente do transporte de ônibus. De acordo com as estimativas, as famílias com maior poder aquisitivo se tornam elegíveis para receber o transporte de ônibus justamente no ponto em que prefeririam mudar de endereço, de qualquer forma. Uma limitação da análise reconhecida pelo próprio autor é de que não é possível observar se os estudantes utilizam o ônibus de fato ou transportes alternativos, como carros privados.

Glazerman (1998) conclui que os consumidores preferem escolas mais próximas e as diferenças importam cada vez menos na medida em que as escolas

sujeitas à comparação são muito distantes da residência. Por fim, as estimativas feitas podem ser usadas para analisar o quanto longe um estudante está disposto a se deslocar por uma unidade incremental na qualidade da escola, o que é justamente o objeto de pesquisa deste trabalho.

Seguindo a linha de pesquisa da demanda por escolas estar condicionada à dependência espacial, Torres e Prior (2014) demonstram que as famílias na Catalunha (Espanha) aparentam escolher a escola considerando o critério de localização. Os resultados confirmam que o espaço tem um efeito moderador na relação entre eficiência e demanda pela escola, apesar do efeito estar restrito a municipalidades urbanas. Isto é, a proximidade com a escola ameniza o fato dessa não ser de alta qualidade, fazendo papel de moderadora entre ineficiência e demanda pela escola.

As conclusões de Torres e Prior (2014) vão de encontro ao que é encontrado na literatura apresentada acima, na medida em que argumentam que a expansão das possibilidades de escolha das escolas pode melhorar a eficiência das escolas públicas através da competição. Essa tese também defendida por Hoxby (2000), que encontrou efeitos positivos entre a performance acadêmica e a competição. Entretanto, Zanzig (1997) chega à conclusão de que uma vez que determinado nível de competitividade é alcançado, acréscimos marginais na competitividade são irrelevantes.

O fator espacial é novamente explorado na análise, com diferentes impactos para zonas rurais e urbanas. No primeiro tipo de zona, é comum ter poucas opções de escolas públicas, restringindo a opção dos pais, logo, o fator espacial não pode ser considerado como um fator de competitividade com outras escolas. Já no espaço urbano, as escolas operam em uma situação de crescente competitividade (ZANZIG, 1997).

Dessa forma, nota-se o crescente movimento na área acadêmica em conceber o fator espacial como essencial na análise econômica. Tradicionalmente, era dada maior importância ao papel do tempo, mas durante os anos 90, autores como Krugman renovaram o interesse por essa questão. A re-emergência da ciência regional favoreceu, por sua vez, a emergência de um novo campo teórico conhecido como econometria espacial. É com base no interesse do estudo da variável espacial que a pesquisa pretendida analisará se os pais têm preferência por matricular seus filhos em boas

escolas, em termos de resultados acadêmicos, a ponto de lidarem com os custos de um deslocamento diário maior a caminho dessas escolas.

3 DADOS

A pesquisa baseia-se majoritariamente nos microdados disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), tanto em relação ao Censo Escolar, quanto à Prova Saeb e o Enem. O Censo Escolar é o principal instrumento de coleta de dados da educação básica e a mais importante pesquisa estatística educacional brasileira, realizada anualmente. Com ele, o Inep verifica desde o número de matrículas e rendimento dos alunos até a infraestrutura das escolas. Conta com tabelas de dados referentes a escolas, alunos, docentes e turmas, os quais são fornecidos pelas próprias escolas (públicas e privadas) e redes de ensino (estaduais e municipais).

Foram utilizados os dados do Censo Escolar de 2019, anteriores ao começo da pandemia de COVID-19 no Brasil, já que essa impôs o isolamento social, alterando a dinâmica do ensino presencial e do deslocamento entre moradia e escola, além de representar um período em que a já existente desigualdade de acesso à educação de qualidade foi ainda mais aprofundada. Como mencionado anteriormente, foi feito um recorte dos alunos que residem e estão matriculados em municípios do estado de São Paulo, totalizando amostra possível de 645 municípios. Foram retiradas observações nas quais o aluno reside em São Paulo, porém está matriculado em escola de outro estado, pois nesses casos, há uma possível desatualização do cadastro do aluno, já que não se trata de um deslocamento diário. Dentro desse recorte e com os dados disponíveis, foram considerados 236.238 alunos matriculados nas escolas públicas do estado de SP.

É importante ressaltar que o escopo da pesquisa contempla apenas a rede de ensino pública, que engloba escolas federais, estaduais e municipais, pois a escolha da escola é essencialmente diferente entre famílias de alunos de escola pública e privada por estar condicionada a fatores relacionados ao nível socioeconômico dos pais, assim como foi concluído pela literatura existente. Dados como o código do município de residência do estudante, a zona em que se localiza (rural e urbana), o código do município da escola em que está matriculado e a cor dos alunos foram agregados em nível municipal.

Ademais, foram coletados dados das tabelas de escolas e de docentes, que refletem características que possivelmente têm correlação positiva com a escolha da

escola. A hipótese é de que municípios com escolas com mais recursos tecnológicos, instalações melhores e oferecimento de educação profissional atraiam mais os pais. Por isso, foram coletados dados da existência de laboratório de informática, laboratório de ciências, quadra de esportes; acesso a equipamentos multimídia; acesso à internet; oferecimento de alimentação e de educação profissional. A última variável compreende curso técnico concomitante e/ou subsequente e/ou curso FIC (Formação Inicial e Continuada) concomitante em classes comuns, de acordo com o dicionário de dados.

Já a hipótese relativa às características dos docentes é de que municípios com escolas com professores mais qualificados sejam mais atrativos, já que representa um fator relevante para a qualidade do ensino.¹ Para testar tal hipótese, foram coletados dados dos docentes sobre a conclusão do ensino superior, realização de curso modalidade licenciatura como primeira graduação, realização de especialização, mestrado e doutorado.

Todas as variáveis acima mencionadas são disponibilizadas, respectivamente, em nível de escola e docente, e como dummies que assumem valor 0 caso a resposta seja "não" e 1, caso "sim". Essas foram agrupadas em nível municipal e foi calculada a porcentagem de escolas ou docentes que têm a característica em questão (valor igual a 1) em cada município.

A fim de coletar informações sobre a qualidade acadêmica das escolas públicas paulistas, foram utilizados dados da Prova Saeb (Sistema Nacional de Avaliação da Educação Básica) de 2017 e do Enem (Exame Nacional do Ensino Médio) de 2018, recorte temporal escolhido com base no fato de que os pais tomam a decisão de qual escola matricular seus filhos a partir de informações previamente disponíveis sobre as instituições. Para ambas as bases, o recorte foi o mesmo do Censo, escolas públicas localizadas em municípios do estado de São Paulo.

¹ O Plano Nacional de Educação (PNE), que estabelece diretrizes e metas tidas como base para a política educacional dos governos federal, estadual e municipal, evidenciou em seu texto que a qualificação dos docentes é determinante para a qualidade do ensino. Nesse sentido, o PNE para o decênio 2014/2024 definiu como uma das vinte metas a formação de 50% dos professores da educação básica das redes pública e privada de ensino do país em nível de pós-graduação (especialização, mestrado e doutorado) até o ano final da sua vigência.

A Prova Saeb é uma avaliação bienal aplicada desde 1995 com o objetivo de coletar informações que contribuam para que gestores públicos tenham uma visão mais ampla da qualidade da educação básica brasileira. Fornece dados sobre o desempenho dos alunos do 5º e 9º anos do Ensino Fundamental em Matemática e Língua Portuguesa. Com realização amostral, participaram ao todo do Saeb 2017 mais de 73 mil escolas e aproximadamente 5,4 milhões de estudantes de escolas públicas e privadas. O total de alunos participantes do Saeb 2017 corresponde a cerca de 68% daqueles matriculados no ensino fundamental e médio, conforme dados do Censo da Educação Básica de 2017. As notas foram agregadas em nível municipal e foi criada uma variável adicional com a média entre as notas de matemática e língua portuguesa.

O Enem, por sua vez, é o Exame Nacional do Ensino Médio realizado anualmente e utilizado desde 2004 para ingresso no ensino superior. O modelo de prova do Enem engloba questões em cinco grandes áreas de conhecimento, sendo elas: i) Linguagens, Códigos e suas Tecnologias; ii) Ciências Humanas e suas Tecnologias; iii) Ciências da Natureza e suas Tecnologias; iv) Matemática e suas Tecnologias e v) Redação.

As notas do Enem são disponibilizadas em nível de escola, mas foram reagrupadas de modo a possibilitar a análise: foi calculada uma nota média por escola (média aritmética entre as notas obtidas nas cinco áreas do conhecimento) e, posteriormente, uma nota média por município, agregando todas as escolas participantes de cada município.

Para que fosse possível a inclusão do fator espacial na análise por meio da investigação do papel da distância na escolha da escola, foram coletadas as coordenadas geográficas (latitude e longitude) dos centroides de cada município, disponibilizadas pelo IBGE, e foi utilizada uma função em software de programação para cálculo de distância. A função calcula as distâncias geodésicas - isto é, a distância considerando a curvatura da Terra - entre todos os pares possíveis de municípios em quilômetros e determina qual é a menor, ou seja, indica o município mais próximo.

Com base no cruzamento desses dados com a base do Censo Escolar referente à matrícula, foi criada variável dummy que indica se o aluno se desloca para o município mais próximo ao de sua residência para estudar. Posteriormente, foi

calculada, para cada município de residência, a porcentagem de alunos que realizava esse deslocamento.

Além disso, foram selecionadas as variáveis de população e de PIB dos municípios de residência para uso como controle. Foram utilizados dados da população estimada para 2019 e do PIB per capita de 2019, os quais são disponibilizados pelo IBGE em nível municipal e enviados para o TCU (Tribunal de Contas da União) após publicação no DOU (Diário Oficial da União). A hipótese é de que municípios mais populosos e mais ricos têm maior oferta de escolas de boa qualidade, diminuindo a probabilidade dos alunos deslocarem-se dos mesmos para estudo.

Por fim, foi feito o cruzamento de todas as bases de dados acima mencionadas, usando como chave o código do município de residência e/ou de escola do estudante. O código identifica os municípios brasileiros de forma inequívoca, por seguir padrão estabelecido pelo IBGE, segundo o qual os dois primeiros dígitos são referentes à unidade da federação e os quatro seguintes, à identificação única do município. Foram excluídas observações com dados faltantes.

4 MODELO ECONOMÉTRICO

4.1 MODELOS PROPOSTOS

A metodologia será pautada na análise econométrica da base de dados construída, fazendo uso também da econometria espacial, pois pretende-se relacionar aspectos geográfico-espaciais à dinâmica social e populacional. Para a estimativa dos parâmetros do modelo, será utilizado o método dos mínimos quadrados ordinários, de forma a minimizar a soma dos quadrados dos resíduos e a maximizar o grau de ajuste do modelo aos dados observados.

Assim, por meio de modelos de regressão linear múltipla, pretendeu-se analisar qual a influência de fatores como a qualidade acadêmica do ensino, infraestrutura das escolas e qualificação dos professores sobre a escolha de escola em outro município, incorrendo em deslocamento diário intermunicipal.

Para efeito de simplificação, as duplas de municípios residência-escola consideradas nessa análise são aquelas em que a pessoa estuda no município mais próximo ao de sua residência.

O primeiro modelo a ser estimado é:

$$Y_{pr} = \alpha + \beta Dif_{pr}(N) + \gamma Dist_{pr} + \mu \ln(PIB) + \theta zona + \varepsilon \quad (1)$$

onde

Y representa a porcentagem de alunos que moram em r e estudam em p;

p representa o município mais próximo ao de residência;

r representa o município de residência;

p e r estão localizados no estado de São Paulo;

$Dif_{pr}(N)$ representa diferença de notas:

$$Dif_{pr}(N) = \text{Nota p} - \text{Nota r}$$

Dist representa distância (em km) entre p e r;

PIB representa PIB per capita do município de residência (em reais);

Zona representa uma variável dummy que assume o valor 1 quando o município de residência se localiza em zona urbana e 0, caso contrário;

ε representa erro;

$\alpha, \beta, \gamma, \mu, \theta$ são parâmetros.

Esse modelo foi estimado separadamente para as notas médias do 5º ano, 9º ano (Saeb) e do ensino médio (Enem), pois há a hipótese de que exista considerável colinearidade entre as variáveis. Por fim, foi estimado um modelo considerando todas as avaliações para analisar qual efeito se sobressai.

Foi proposto também um segundo modelo que considera as características das escolas e dos docentes:

$$Y_{pr} = \alpha + \beta Dif_{pr}(N) + \lambda Dif_{pr}(E) + \omega Dif_{pr}(D) + \gamma Dist_{pr} + \mu \ln(PIB) + \theta \text{ zona} + \varepsilon \quad (2)$$

onde

$Dif_{pr}(X)$ representa diferença entre porcentagens:

$$Dif_{pr}(x) = X_p - X_r$$

r representa o município de residência;

p representa o município mais próximo ao de residência;

E é um vetor de porcentagem de escolas com as seguintes características: i) possui laboratório de informática, ii) possui laboratório de ciências, iii) possui quadra de esportes, iv) possui acesso à internet, v) fornece alimentação, vi) possui equipamento multimídia, vii) oferece ensino profissional;

D é um vetor de porcentagem de docentes com as seguintes características: realizou i) ensino superior, ii) especialização, iii) licenciatura como primeira graduação, iv) mestrado, v) doutorado;

as variáveis já mencionadas no modelo 1 se mantêm.

De forma análoga, o modelo foi estimado separadamente para as notas médias do 5º ano, 9º ano (Saeb) e do ensino médio (Enem) e, posteriormente, considerando todas as anteriores.

4.2 ESTATÍSTICAS DESCRIPTIVAS

Tabela 1 – Estatísticas descritivas

Variável	Observações	Média	Mínimo	Máximo	Desvio
Porcentagem deslocados	547	0,0262	0	1	0,1229
Zona	547	0,841	0	1	0,3661
Dif Saeb 5 ano	546	-0,5901	-43,4	48,535	14,6461
Dif Saeb 9 ano	546	0,1962	-56,265	56,265	16,1851
Dif Enem	542	-3,778	114,001	114,0009	36,0961
Dif Internet	547	-0,0001	-0,8333	0,8333	0,2624
Dif Lab. Informática	547	0,0093	-0,7	0,75	0,2328
Dif Profissionalizante	547	-0,0019	-0,1429	0,25	0,0334
Dif Mestrado	547	0,0002	-0,1353	0,1412	0,033
Dif Doutorado	547	-0,0003	-0,0409	0,0495	0,0105
Distância	547	12,1629	1,1074	47,3231	5,8233
Ln PIB	547	10,2571	9,1088	12,9669	0,5531

Fonte: Elaboração própria

A fim de analisar as características gerais da base de dados construída, são apresentadas na Tabela 1 as estatísticas descritivas considerando as principais variáveis. Como mencionado, a variável dependente representa a porcentagem de alunos de determinado município de São Paulo que se desloca, em forma decimal; variando, portanto, entre 0 e 1. A parcela média de pessoas que se desloca para o município mais próximo ao de sua residência para estudar é de 2,62%. Nesse sentido, não se espera que os coeficientes estimados sejam de grande magnitude, apesar da expectativa de confirmação das hipóteses propostas. Pode ser considerada em futuras análises a inclusão de outros municípios de destino, o que aumentaria a amostra, apesar de incorrer por outro lado em dificuldades para definição de quais devem ser considerados.

A diferença de porcentagem de professores com mestrado, doutorado ou outra especialização entre os municípios são pequenas, o que também é reflexo da baixa quantidade de docentes com essa formação. De acordo com dados da Secretaria Municipal de São Paulo, por exemplo, apenas 1,2% dos professores da rede municipal de ensino da cidade de São Paulo realizou mestrado e 0,08% concluiu doutorado. As diferenças mais significativas estão relacionadas às notas médias, em especial, àquelas do Enem.

Como mencionado na seção anterior, as diferenças são calculadas na mesma ordem para todas as variáveis: município de escola menos o de residência. Assim, as

variáveis de diferença podem assumir valores negativos nos casos em que a variável de interesse do município de moradia é maior do que a de escola; ou valores positivos, caso contrário (variável de interesse do município de matrícula é maior).

5 RESULTADOS

Como já exposto, ambos os modelos foram estimados para cada uma das séries separadamente (5º ano, 9º ano e ensino médio) e, posteriormente, de forma conjunta para análise de qual efeito tem predominância. Para análise mais detalhada dos resultados, o segundo modelo foi segmentado em modelo 2, que contempla as características das escolas em adição às variáveis de diferença de notas, distância, PIB e zona do modelo 1; e o modelo 3, que adiciona as características dos docentes ao modelo anterior. Todos os coeficientes estimados foram multiplicados por 1000 para melhor visualização e interpretação dos resultados.

Os resultados para o 5º ano do ensino fundamental são apresentados na Tabela 2, na qual nota-se a significância estatística do coeficiente relacionado à diferença de notas na prova Saeb ao nível de 5% em todos os modelos propostos. Como pode ser visto no terceiro modelo, dado um aumento de 10 pontos na diferença das notas médias do Saeb de dois municípios, haverá um aumento de 0,83% de pessoas se deslocando para estudar no município de maior nota, *ceteris paribus*.

No primeiro modelo, o PIB per capita possui coeficiente estatisticamente significante ao nível de 10% e o seu sinal negativo corrobora com a hipótese de que quanto mais rico um município, maior a oferta de escolas de qualidade e, consequentemente, menor a probabilidade de que os residentes desse precisem se deslocar em busca de uma escola.

Nos dois modelos finais, destaca-se também a descoberta de três outros coeficientes significantes: o acesso à internet nas escolas, o oferecimento de ensino profissionalizante e a formação de docentes em mestrado.²

²É comum encontrar matérias na internet com recomendação de critérios para a escolha de uma boa escola para os filhos. Dentre eles, pode-se notar sugestões de busca por escolas com bons docentes e com adequadas instalações físicas, sendo citada a utilização de computadores para auxiliar no aprendizado. Outros fatores comumente citados são a distância em relação à residência e a identificação da família com os valores da escola. Além disso, os artigos citam que dois indicadores podem ser consultados: o Ideb (Índice de Desenvolvimento da Educação Básica), que utiliza as notas do Saeb, e o Enem. Assim, os resultados estão coerentes com alguns dos fatores expostos nesses canais, os quais podem ter influência na decisão dos pais, uma vez que são websites com alto número de visualizações.

Dado que um município tem 10% a mais de escolas que oferecem ensino profissional, há 6,5% a mais de pessoas que se deslocam para estudar nesse, ceteris paribus (Tabela 2). A diferença percentual de municípios com oferecimento de ensino profissionalizante segue significante mesmo ao nível de 1%. Além disso, dado que um município tem 10% a mais de professores com mestrado, há 3,48% de pessoas a mais que se deslocam para estudar nele, ceteris paribus.

Tabela 2 - Resultados 5º Ano

	<i>Variável dependente:porcentagem de deslocados para município mais próximo</i>		
	Modelo (1)	Modelo (2)	Modelo (3)
Dif notas	0.907** (0.359)	0.920** (0.362)	0.829** (0.364)
Distância	-1.092 (0.910)	-0.902 (0.895)	-1.011 (0.895)
Ln PIB	-18.548* (9.651)	-14.359 (9.540)	-10.929 (9.612)
Zona	0.908 (14.646)	0.575 (14.412)	-2.204 (14.412)
Dif Lab Informática		-26.603 (22.601)	-25.538 (22.574)
Dif Lab Ciências		-9.251 (40.360)	-12.812 (40.386)
Dif Internet		42.735** (20.629)	47.056** (20.838)
Dif Equip Multimídia		-7.718 (17.348)	-9.864 (17.316)
Dif Profissionalizante		702.022*** (159.498)	650.426*** (160.405)
Dif Ensino Superior			-76.551 (114.859)
Dif Mestrado			347.703** (173.312)
Dif Doutorado			265.667 (530.754)
Constante	229.545** (99.377)	186.237* (98.182)	154.718 (98.729)
Observações	546	546	546
R ²	0.022	0.066	0.077
R ² Ajustado	0.015	0.050	0.056
Erro Padrão Residual	122.146 (df = 541)	119.918 (df = 536)	119.552 (df = 533)
Estatística F	3.023** (df = 4; 541)	4.204*** (df = 9; 536)	3.696*** (df = 12; 533)

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

Fonte: Elaboração própria

Tabela 3 - Resultados 9º Ano

	<i>Variável dependente:porcentagem de deslocados para município mais próximo</i>		
	Modelo (1)	Modelo (2)	Modelo (3)
Dif notas	0.591* (0.326)	0.601* (0.327)	0.518 (0.330)
Distância	-1.182 (0.911)	-0.993 (0.896)	-1.101 (0.895)
Ln PIB	-18.877* (9.678)	-14.565 (9.569)	-11.086 (9.637)
Zona	1.481 (14.731)	1.067 (14.493)	-1.937 (14.492)
Dif Lab Informática		-22.973 (22.619)	-22.391 (22.581)
Dif Lab Ciências		-7.546 (40.627)	-11.505 (40.662)
Dif Internet		42.911** (20.701)	47.241** (20.896)
Dif Equip Multimídia		-4.096 (17.272)	-6.553 (17.242)
Dif Profissionalizante		709.582*** (159.994)	656.694*** (160.915)
Dif Ensino Superior			-62.819 (114.949)
Dif Mestrado			344.047** (174.956)
Dif Doutorado			374.835 (533.676)
Constante	232.884** (99.679)	188.334* (98.511)	156.575 (99.012)
Observações	546	546	546
R ²	0.016	0.061	0.072
R ² Ajustado	0.009	0.045	0.051
Erro Padrão Residual	122.494 (df = 541)	120.259 (df = 536)	119.855 (df = 533)
Estatística F	2.238* (df = 4; 541)	3.843*** (df = 9; 536)	3.453*** (df = 12; 533)

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

Fonte: Elaboração própria

Os resultados referentes ao 9º ano do ensino fundamental são apresentados na Tabela 3 e vão ao encontro do que foi apresentado para o 5º ano. O sinal dos impactos é mantido e os mesmos coeficientes são estatisticamente significantes, sendo o nível de significância do coeficiente relacionado às notas a única diferença, devido ao aumento do p-valor em 5%.

Já os resultados para o ensino médio são apresentados na Tabela 4. A partir desses, pode-se notar que o impacto das notas é ainda mais evidente, uma vez que o coeficiente da diferença de notas no Enem é significante mesmo ao nível de 1%.

Tabela 4 - Resultados Ensino Médio

	<i>Variável dependente: porcentagem de deslocados para município mais próximo</i>		
	Modelo 1 (1)	Modelo 2 (2)	Modelo 3 (3)
Dif notas	0.939*** (0.144)	0.840*** (0.160)	0.777*** (0.165)
Distância	-1.043 (0.887)	-0.996 (0.884)	-1.065 (0.885)
Ln PIB	-8.068 (9.545)	-7.129 (9.520)	-5.505 (9.582)
Zona	1.860 (14.285)	0.760 (14.235)	-1.147 (14.287)
Dif Lab Informática		-22.114 (22.422)	-21.790 (22.463)
Dif Lab Ciências		12.634 (39.881)	8.210 (40.112)
Dif Internet		50.667** (20.454)	53.304** (20.724)
Dif Equip Multimídia		-1.189 (16.842)	-3.086 (16.888)
Dif Profissionalizante		309.401* (174.227)	306.132* (174.373)
Dif Ensino Superior			-22.196 (114.080)
Dif Mestrado			224.475 (175.312)
Dif Doutorado			245.677 (524.659)
Constante	123.857 (98.254)	115.081 (97.923)	100.732 (98.425)
Observações	542	542	542
R ²	0.083	0.101	0.106
R ² Ajustado	0.076	0.086	0.086
Erro Padrão Residual	118.705 (df = 537)	118.042 (df = 532)	118.078 (df = 529)
Estatística F	12.107*** (df = 4; 537)	6.669*** (df = 9; 532)	5.222*** (df = 12; 529)

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

Fonte: Elaboração própria

A magnitude do coeficiente associado à diferença percentual de escolas com acesso à internet entre municípios é a maior dentre os resultados apresentados neste

trabalho. Na tabela 4, é possível observar que dado um município com 10% a mais de escolas com acesso à internet, 0,53% pessoas a mais se deslocam para estudar no município em questão. Entretanto, o coeficiente relacionado ao mestrado deixa de ser significante. Além disso, é importante notar que, embora ainda reduzido, o R^2 ajustado dos modelos do ensino médio apresentam aumento em relação aos modelos que consideram o 5º e 9º anos do ensino fundamental.

É possível atribuir esses resultados à importância do ensino médio na vida escolar de uma pessoa, já que é o passo que precede e prepara o aluno para o ingresso no ensino superior. Ademais, os estudantes dessa fase escolar possuem maior autonomia para se deslocarem por conta da idade em comparação com alunos do 5º e 9º ano.

Tabela 5 - Resultados para todos os anos

<i>Variável dependente: porcentagem de deslocados para município mais próximo</i>		
	Modelo (1)	Modelo (2)
Dif notas 5ano	0.521 (0.395)	0.469 (0.393)
Dif notas 9ano	0.115 (0.355)	0.120 (0.354)
Dif notas ensino médio	0.897*** (0.147)	0.739*** (0.166)
Distância	-0.964 (0.888)	-1.007 (0.884)
Ln PIB	-7.816 (9.551)	-5.168 (9.524)
Zona	3.359 (14.352)	0.080 (14.298)
Dif Internet		49.983** (19.564)
Dif Profissionalizante		338.302** (169.104)
Dif Mestrado		232.115 (163.283)
Constante	119.246 (98.364)	95.419 (97.921)
Observações	541	541
R^2	0.087	0.108
R^2 Ajustado	0.077	0.092
Erro Padrão Residual	118.734 (df = 534)	117.743 (df = 531)
Estatística F	8.515*** (df = 6; 534)	7.109*** (df = 9; 531)

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

Por fim, os resultados da análise econométrica incluindo o ensino fundamental e o ensino médio são apresentados na Tabela 5. Essa evidencia que o efeito das notas do Enem se sobressai aos demais, corroborando com a argumento acima defendido. Conforme pode ser visto na Tabela 5, o coeficiente referente a essa variável é estatisticamente significante mesmo ao nível de 1%. Mantém-se a significância das variáveis de acesso à internet e oferecimento de ensino profissionalizante.

É importante pontuar que algumas variáveis mencionadas no capítulo de metodologia foram suprimidas dos modelos aqui apresentados para que as tabelas de resultados se mantivessem mais enxutas, já que não se mostraram estatisticamente significantes quando adicionadas à regressão econométrica (por exemplo, existência de quadra de esportes, diferença percentual de docentes formados em licenciatura no primeiro curso).

A fim de contribuir para possíveis extensões da presente pesquisa, destacam-se algumas limitações encontradas que devem ser consideradas. Caso fosse pretendida uma análise em nível individual e de escolas, ao invés de municipal como foi realizado, haveria um obstáculo metodológico: definir o contrafactual a ser utilizado na comparação de notas. Isso porque não é possível definir qual escola específica seria a escolha dos pais caso optassem por não deslocarem seus filhos, sendo uma possibilidade viável a utilização da nota média do município de residência, apesar de não ser o dado mais preciso nessa situação. Além disso, outra limitação seria a impossibilidade de utilizar notas desagregadas do Enem para essa comparação, uma vez que a divulgação do Enem por escola foi descontinuada em 2017. Uma alternativa seria coletar os dados desagregados do Ideb (Índice de Desenvolvimento da Educação Básica) por escola, porém isso poderia trazer perdas significativas para a análise, já que a diferença de notas no Enem se demonstrou estatisticamente significante em todos os modelos, como mencionado.

6 CONCLUSÕES

Primeiramente, ressalta-se uma limitação da pesquisa devido à ausência de dados sobre a localização das residências de uma parcela de alunos matriculados no estado de São Paulo no Censo Escolar de 2019, cujas causas poderiam ser alvo de análise, já que se trata de uma base de dados fundamental para a pesquisa no campo da educação. Além disso, alguns alunos matriculados em São Paulo têm municípios de residência cadastrados fora do estado. Esses casos não representam deslocamentos diários e, provavelmente, tratam-se de dados desatualizados a respeito da moradia do estudante. Outra limitação de dados é o fato de que o Censo Demográfico, base fundamental para diversas pesquisas, não apenas para o campo da educação, foi realizado pela última vez em 2010. Essas restrições impedem que sejam feitas análises mais acuradas do cenário atual.

Ainda assim, a partir dos resultados dos modelos estimados, pode-se concluir que é verdadeira a hipótese de que a qualidade acadêmica está correlacionada positivamente com uma maior quantidade de pessoas se deslocando em direção a municípios com escolas que obtiveram notas maiores em provas como o Saeb e o Enem. Para todos os anos escolares avaliados, que engloba séries do ensino fundamental e o ensino médio, o resultado é mantido estatisticamente significante, mostrando-se ainda mais relevante no cenário do ensino médio. Além disso, o fato da escola ter acesso à internet, oferecer cursos profissionalizantes e possuir docentes capacitados com realização de mestrado são outros fatores relevantes para a escolha da escola. Sendo assim, conclui-se que esses resultados são indicativos de que as famílias paulistas valorizam a educação, já que ao escolher escolas para seus filhos, optam por escolas de melhor qualidade, apesar dos custos do deslocamento diário intermunicipal.

A análise desenvolvida na presente pesquisa pode servir como contribuição para que a literatura brasileira enfoque estudos sobre a preferência dos pais e/ou alunos na decisão da escola. Ademais, pode ser utilizada como insumo para reflexão e construção de políticas públicas com o objetivo de promover o ensino público de qualidade de forma democrática e distribuída geograficamente.

REFERÊNCIAS

ABDULKADIROĞLU, Atila et al. Do parents value school effectiveness?. **American Economic Review**, v. 110, n. 5, p. 1502-39, 2020.

BURGESS, Simon; GREAVES, Ellen; VIGNOLES, Anna; WILSON, Deborah **What Parents Want: School Preferences and School Choice**. *The Economic Journal, Royal Economic Society*, 2014

GLAZERMAN, Steven M. **School Quality and Social Stratification: The Determinants and Consequences of Parental School Choice**. San Diego, CA: 1998

HANUSHEK, Eric A. et al. The market for teacher quality. National Bureau of Economic Research, 2005.

HOXBY, Caroline M. Does competition among public schools benefit students and taxpayers?. **American Economic Review**, v. 90, n. 5, p. 1209-1238, 2000.

JACOB, Brian A.; LEFGREN, Lars. What do parents value in education? An empirical investigation of parents' revealed preferences for teachers. **The Quarterly Journal of Economics**, v. 122, n. 4, p. 1603-1637, 2007.

NGWARE, Moses W.; OKETCH, Moses; EZEH, Alex C. Quality of primary education inputs in urban schools: Evidence from Nairobi. **Education and Urban Society**, v. 43, n. 1, p. 91-116, 2011.

TORRES Laura López; PRIOR Diego **Measuring school demand in the presence of spatial dependence. A conditional approach**. Barcelona: 2014

ZANZIG, Blair R. Measuring the impact of competition in local government education markets on the cognitive achievement of students. **Economics of Education Review**, v. 16, n. 4, p. 431-441, 1997.