

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO
DEPARTAMENTO DE ECONOMIA

NATHALIA XAVIER COSTA

**UMA ANÁLISE EMPÍRICA SOBRE A DIFERENÇA ENTRE ENSINO
PÚBLICO E PRIVADO NO RIO DE JANEIRO**

RIO DE JANEIRO

2022

NATHALIA XAVIER COSTA

**UMA ANÁLISE EMPÍRICA SOBRE A DIFERENÇA ENTRE ENSINO
PÚBLICO E PRIVADO NO RIO DE JANEIRO**

Trabalho apresentado no curso de
graduação da Pontifícia Universidade
Católica do Rio de Janeiro.

Orientador: Arthur Bragança

RIO DE JANEIRO

2022

AGRADECIMENTOS

A Deus, por me dar forças ao longo dessa caminhada e me manter firme para alcançar o objetivo final.

Aos meus pais, que me ensinaram a não desistir, sempre me incentivaram e me apoiaram.

A minha irmã, Elisa, minha maior companheira, que esteve do meu lado em todos os momentos.

Ao meu orientador, por atender tão prontamente ao meu convite para elaboração do projeto.

SUMÁRIO

| | |
|---|-----------|
| 1. INTRODUÇÃO | 5 |
| 2. REVISÃO DE LITERATURA..... | 6 |
| 3. CONTEXTUALIZAÇÃO DA EDUCAÇÃO NO RIO DE JANEIRO..... | 9 |
| 3.1 Contexto institucional..... | 9 |
| 3.2 Indicadores educacionais..... | 9 |
| 4. ANÁLISE EMPÍRICA..... | 14 |
| 4.1 Dados..... | 14 |
| 4.2 Estratégia empírica..... | 15 |
| 5. CONCLUSÃO..... | 24 |
| REFERÊNCIAS | 25 |

1 INTRODUÇÃO

Atualmente, o sistema educacional brasileiro se configura como híbrido, tendo uma coexistência de instituições públicas e privadas de ensino. Em tese, havendo uma diferença de órgão administrativo é esperado que também haja uma disparidade de resultados apresentados. Assim, o intuito do presente trabalho é validar a afirmação anterior de maneira empírica.

Antes de tudo, vale reforçar a importância da educação na sociedade como um todo. Pela Teoria do Capital Humano, a educação é um dos principais promotores de desenvolvimento econômico de um país. Considerada como um dos investimentos que podem ser feitos em capital humano, a educação é um fator que aumenta a produtividade do mesmo, de modo que o investimento em capital humano é responsável pela maior parte do aumento dos rendimentos reais por trabalhador (Schultz, 1961).

Entretanto, não é suficiente atentar apenas para quantidade. Um estudo feito por Hanushek e Kimko (2003) concluiu que, para uma amostra de cerca de 80 países, a qualidade da força de trabalho está relacionada positivamente com as taxas de crescimento econômico no período 1960-1990 e é mais importante que a média de anos de escolaridade. Sendo assim, para além de mostrar o impacto da exposição do capital humano a uma instituição de ensino, o projeto pretende discutir sobre nível de qualidade das instituições no caso brasileiro, sobretudo diferenciando escolas públicas das privadas.

Usando como amostra o Rio de Janeiro, foi feita uma contextualização do sistema educacional do estado, abordando seus principais indicadores e, em seguida, uma análise empírica da diferença qualitativa entre instituições. Com tal finalidade, serão usados dados disponibilizados pelo Inep, sobretudo o Censo Escolar.

2 REVISÃO DE LITERATURA

Em suma, o produto de uma economia é gerado por diferentes fatores. Ao longo dos anos, houve uma alteração de peso nos insumos de produção. Até o século XVIII, a terra era a base da economia. A Revolução Industrial mudou esse cenário, tornando o capital físico um elemento fundamental. Nos dias atuais, o principal recurso para desenvolvimento é o capital humano (Crawford, 1994). A sociedade pós-industrial passou a dar mais ênfase ao sujeito e seu valor intrínseco (Beyer, Soler, 2019).

A expressão capital humano apareceu pela primeira vez na literatura econômica em 1961 em um artigo intitulado Investindo em Capital Humano, publicado na *American Economic Review* escrito por Theodore Schultz - por mais que o conceito de capital humano já fosse conhecido no século XVIII por Adam Smith e outros economistas (Crawford, 1994). Schultz foi professor do departamento de economia da Universidade de Chicago e buscava explicar o retorno sob o investimento em capital humano. Assim, Schultz foi o precursor da Teoria do Capital Humano.

Capital humano para Schultz (1961) foi a aquisição de todas as habilidades e conhecimentos úteis, que é parte do investimento deliberado. Ou seja, capital humano é a capacidade de conhecimentos, competências e atributos da personalidade de uma pessoa ao desempenhar um trabalho, de modo a produzir valor econômico.

O investimento em capital humano não se resume apenas ao gasto em educação, mas também aquele em saúde, treinamento ou qualquer outro ponto que aumenta a produtividade de um indivíduo. Mas, no que caracteriza o capital humano, a educação se destaca.

Schultz (1964) trouxe a ideia de que a qualificação e o aperfeiçoamento da população via investimento em educação gera um aumento no rendimento dos indivíduos, de modo que pessoas mais educadas numa sociedade impactam positivamente sua economia. Na época, um case para Schultz foram os Estados Unidos, validando ainda mais sua ideia.

Tornou-se claro para mim também que nos Estados Unidos muitas pessoas que estão investindo no homem estão tendo grande influência sobre o crescimento econômico,

e que o investimento em capital humano fundamental é a educação. (T. SCHULTZ, 1963)

Becker (1993) aderiu ao mesmo pensamento e ampliou a importância do trabalho do Schultz, agregando mais no campo da formalização do modelo. Ele trouxe ainda o questionamento sobre em que condições os indivíduos determinam se investem ou não em educação. Em um de seus textos, afirma que:

Human capital analysis assumes that schooling raises earnings and productivity mainly by providing knowledge, skills and a way of analyzing problems. (G. BECKER, 1993)

Além disso, Mankiw, Romer e Weil (1992) incluem o capital humano ao modelo de Solow (1956), no intuito de evidenciar que essa variável também afeta o crescimento econômico de um país. Os autores exploram as implicações empíricas do modelo de Solow e observaram que a inclusão do capital humano ao modelo poderia melhorar a eficácia do mesmo (Jones, 2014). Agora, a função de produção passará a ser uma Cobb-Douglas, que combina capital físico, K , capital humano, H , e tecnologia aumentadora de trabalho, A , que é uma variável exógena:

$$Y = K^{\alpha}(AH)^{1-\alpha}$$

Assim, se comprovou como o investimento em capital humano também é um dos fatores que explicam as diferenças *cross-country*, sendo o capital humano é um fator fundamental para o crescimento econômico nas organizações e também nas vidas privadas (Schultz, 1961).

Até então, a medida usada para educação era a quantidade de anos de estudo. Sob outra perspectiva, Hanushek e Woessmann (2012) procuram metrificar educação de acordo com seu desempenho, uma vez que os investimentos em educação não pareciam devolver os resultados econômicos prometidos pelos modelos teóricos de crescimento. Os autores usam como variáveis notas de testes internacionais de matemática e ciências para avaliar desempenho acadêmico e relacionam ao PIB dos países em questão. Ao permitir diferenças de desempenho entre alunos com diferentes qualidades de escolaridade, mas possivelmente com a mesma

quantidade de anos de estudo, se abre a investigação sobre a relevância de diferentes políticas com a finalidade de impactar a qualidade educacional.

We find that accurately measuring differences in educational achievement, which we refer to simply as cognitive skills, dramatically improves our ability to explain variations in long-run growth across countries. (E. HANUSHEK, L. WOESSMANN, 2012)

Estudos como Pop-Eleches e Urquiola (2013) vão dar mais ênfase a importância da qualidade, ao buscar avaliar se existe benefício em frequentar escolas de alto desempenho. A partir do modelo de regressão descontínua e da amostra de alunos próximos ao ponto de corte para entrada nas chamadas escolas de alto desempenho, o acesso a uma escola melhor classificada melhora os resultados cognitivos medidos por um exame ou teste.

Lyons (2001) ainda relaciona desempenho escolar a infraestrutura da instituição onde o aluno estudou, concluindo uma correlação positiva entre as variáveis. Fatores como menor acesso a luz natural, instalações ruins e acústica desfavorável impactam atenção e performance do aluno.

Learning in the classroom requires a reasonable level of concentration, listening, writing, and reading. Individual classrooms and entire facilities need to be evaluated, not only on how they meet changing educational requirements, but also on how they meet the environmental requirements for health, safety, and security. (J. B. LYONS, 2001)

Dado o peso do capital humano e da qualidade da educação que o mesmo se expõe para uma economia como um todo, se evidencia a necessidade de analisar as políticas educacionais brasileiras e como tem afetado seus alunos. O foco do trabalho é examinar especificamente qualidade das instituições de estudo, diferenciando sobretudo ensino privado e público.

3 CONTEXTUALIZAÇÃO DA EDUCAÇÃO NO RIO DE JANEIRO

3.1 Contexto institucional

O Rio de Janeiro está entre os lugares com maior número de escolas públicas no mundo (Longaigh, 2017), se tornando uma amostra interessante para fundamentar o estudo sobre as diferenças entre escolas públicas e privadas. Dessa forma, o capítulo busca contextualizar a realidade do estado, que será fonte de dados para análise empírica.

Vale frisar que o estado segue o mesmo padrão dos outros estados brasileiros, em termos de níveis escolares, de modo que o ensino básico compreende educação infantil, ensino fundamental e ensino médio e o ideal seria o estudante concluir o último com 17 anos de idade. Além disso, desde 1988, a educação básica brasileira foi estruturada pela Lei de Diretrizes e Bases da Educação (LDB - 9.394/96) e dá autonomia às escolas no que diz respeito a sua organização administrativa, disciplinar e didática, permitindo variações de formato de ensino dependendo do tipo de direção da escola.

3.2 Indicadores educacionais

Para métricas de qualidade de ensino, o Brasil possui indicadores. Muitos deles são obtidos pelos dados publicados pelo Censo Escolar. Considerado o principal instrumento de coleta de dados da educação básica no Brasil, o Censo contém informações sobre infraestrutura, matrículas e docentes a nível escola.

Desde 1931, com a criação do Ministério da Educação e Saúde, o Brasil passou a ter o levantamento de estatísticas educacionais. No decorrer dos anos, houveram inúmeras mudanças relacionadas a maneira que eram gerados os dados, de forma que, hoje o órgão oficial encarregado pela elaboração dos dados supracitados é o Instituto Nacional de Estudos e Pesquisas Educacionais (Inep). Assim, a construção do Censo Escolar é coordenada pelo Inep com a colaboração das secretarias estaduais e municipais de educação.

O Censo aborda diferentes temas como acessibilidade, alimentação fornecida, recursos tecnológicos e de infraestrutura, etc. Observando os dados oriundos do Censo de 2019

compreendendo o estado do Rio de Janeiro, as instituições se equiparam em termos de infraestrutura. As figuras abaixo apresentam de maneira mais detalhada essas informações.

Figura 1 – Recursos tecnológicos nas escolas de ensino médio no Rio de Janeiro - 2019

| | Escolas Públicas | Escolas Privadas |
|----------------|------------------|------------------|
| Internet | 91% | 97% |
| Aparelho DVD | 80% | 80% |
| Impressora | 72% | 84% |
| Copiadora | 66% | 79% |
| Retroprojektor | 16% | 9% |
| Televisão | 94% | 91% |

Fonte: Elaboração própria com base nos dados do Censo Escolar

Figura 2 – Recursos relacionados à infraestrutura disponíveis nas escolas de ensino médio no Rio de Janeiro - 2019

| | Escolas Públicas | Escolas Privadas |
|----------------------------|------------------|------------------|
| Biblioteca | 43% | 59% |
| Laboratório de informática | 50% | 43% |
| Laboratório de ciências | 16% | 19% |
| Sala de leitura | 50% | 54% |
| Quadra de esportes | 47% | 52% |
| Cozinha | 98% | 72% |

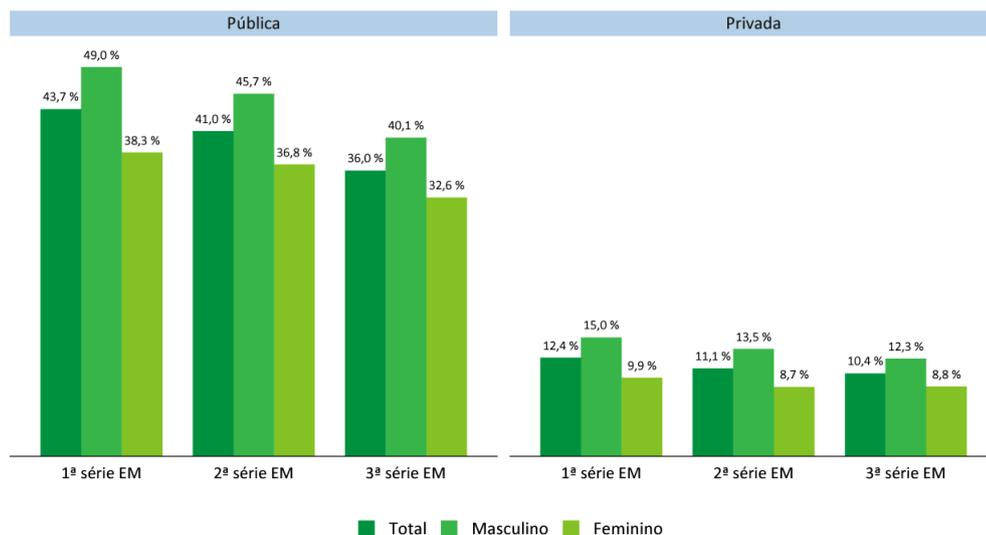
Fonte: Elaboração própria com base nos dados do Censo Escolar

A partir do Censo é possível ainda gerar indicadores como a distorção idade-série (DIS) que, segundo o Ministério da Educação (MEC), é a proporção de alunos com mais de 2 anos de atraso escolar. No Brasil, a criança deve ingressar no 1º ano do ensino fundamental aos 6 anos de idade, permanecendo no Ensino Fundamental até o 9º ano, com a expectativa de que conclua os estudos nesta modalidade até os 14 anos de idade e o conclua o ensino médio, que

compreende 3 anos, até os 17 anos. O cálculo da DIS é realizado a partir de dados coletados no Censo Escolar, uma vez que uma das informações capturadas também é sua idade.

A DIS, por sua vez, torna mais perceptível a distinção entre escolas públicas e privadas, já que esse é um problema que alcança milhares de crianças e adolescentes na educação básica, mas ocorre predominantemente em escolas públicas (Santos, 2020). Avaliando o estado do Rio de Janeiro em 2019, 40,8% dos alunos em escola pública estavam com atraso escolar, enquanto que 11,7% dos alunos em escola privada estavam na mesma situação. Conforme a figura abaixo, essa diferença é vista em cada série do ensino médio, independente do sexo.

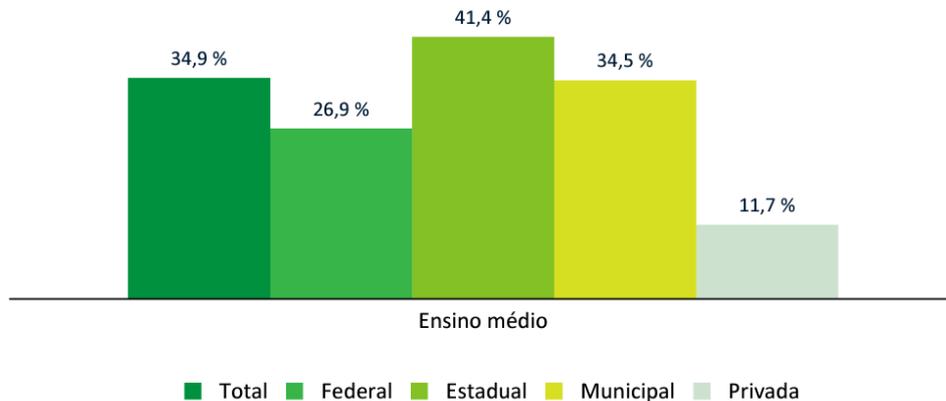
Figura 3 – Distorção idade-série do ensino médio por rede de ensino e sexo - Rio de Janeiro - 2019



Fonte: Elaboração por DEED/Inep com base nos dados do Censo Escolar

Como pode haver diferença de acordo com a dependência administrativa, é válido também avaliar a DIS entre escolas federais, estaduais e municipais, mas é aparente que essa é uma dificuldade mais recorrente em escolas públicas de maneira geral, conforme a figura a seguir, extraída do resumo técnico do Censo Escolar do estado do Rio de Janeiro.

Figura 4 – Distorção idade-série do ensino médio segundo a dependência administrativa - Rio de Janeiro - 2019



Fonte: Elaboração por DEED/Inep com base nos dados do Censo Escolar

Outro importante indicador nacional é o Índice de Desenvolvimento da Educação Básica (IDEB), obtido pela multiplicação entre dois fatores: desempenho dos alunos e a taxa de rendimento escolar. O IDEB foi criado em 2007 pelo Inep e tem o objetivo de monitorar a qualidade de educação.

As médias de desempenho utilizadas para cálculo do IDEB são as da Prova Brasil, para escolas e municípios, e do Sistema de Avaliação da Educação Básica (Saeb), para os estados e o País, realizados a cada dois anos. Basicamente, é feita uma média aritmética entre as notas de português e matemática das avaliações citadas, buscando compreender o nível de aprendizado dos alunos.

A taxa de rendimento é a porcentagem de alunos que foram aprovados no período em questão e, por si só, é um importante indicador educacional, podendo também ser denominada fluxo ou taxa de aprovação. Em 2019 temos que, em escolas privadas, houve uma taxa de aprovação de 94,3%, enquanto que, em escolas públicas a taxa foi de 78,5%.

Dessa forma, o cálculo do IDEB para escolas privadas será:

$$5,94 \times 0,95 = 5,7$$

Da mesma maneira, o cálculo do IDEB para escolas públicas será:

$$4,44 \times 0,79 = 3,5$$

Assim, embora não exista grandes diferenças nos quesitos de infraestrutura, os indicadores educacionais sugerem que escolas privadas tem maior qualidade, o que já foi comentado por Naercio Menezes (2012):

“O sistema público de ensino tende a ter uma qualidade média baixa, insuficiente para formar os indivíduos para o mercado de trabalho, para o sucesso profissional e também pessoal. Por isso, surge o mercado privado de educação, que muitas vezes exige um alto dispêndio por parte das famílias, para prover um ensino de melhor qualidade.” (MENEZES, 2012)

O próximo capítulo visa fazer uma análise empírica para validar tal afirmação.

4 ANÁLISE EMPÍRICA

4.1 Dados

Com o intuito de analisar se o fato de estudar em instituição pública ou privada impacta em indicadores educacionais, foi gerado um histórico dos alunos com os dados do Censo Escolar, disponibilizados publicamente pelo Inep. O histórico reflete o tipo de escola em que determinado aluno estudou nos anos do seu ensino médio e, na sua construção, foram escolhidos apenas os anos de 2016 e 2017 por serem os únicos incluindo variável que possibilite uma associação entre anos. Como já citado, foram usadas as informações referentes ao estado do Rio de Janeiro e as mesmas foram extraídas diretamente do site da ONG Base dos Dados, que trata diferentes categorias de dados, facilitando o acesso a dados públicos.

A variável explicativa do modelo é a quantidade de anos de estudo em escola privada do estudante, derivada do histórico desenvolvido, podendo assumir os valores 0, 1 e 2, já que temos informações de apenas dois anos. A variável explicada, por sua vez, é a distorção idade-série, decorrente da subtração entre a idade em que o aluno está concluindo o ensino médio em 2017, o que é comunicado pelo Censo Escolar, e a idade em que o aluno deve concluir sua fase escolar – 17 anos, segundo o MEC. Cabe salientar que foi feita uma limpeza permitindo na base apenas os indivíduos que estavam no último ano do ensino médio em 2017, por conta da variável endógena do modelo.

O Censo Escolar foi de suma importância por entregar ainda características pessoais dos alunos, como sexo, raça e região onde vive. Assim, é viável incluir variáveis de controle no modelo.

Ao final, temos uma base totalmente criada a partir dos dados fornecidos pelo Censo Escolar formando uma cross-section com observações de 128.227 indivíduos, todos residentes do estado do Rio de Janeiro e terminando seu ensino básico no ano de 2017. Segue a estatística descritiva das variáveis principais:

Tabela 1 - Estatística Descritiva

| Statistic | N | Mean | St. Dev. | Min | Pctl(25) | Pctl(75) | Max |
|----------------------------------|---------|-------|----------|-------|----------|----------|-------|
| Anos de estudo em escola privada | 128,245 | 1.098 | 1.584 | -3 | 0 | 1 | 58 |
| Distorção idade-série | 128,227 | 0.523 | 0.871 | 0.000 | 0.000 | 2.000 | 2.000 |

Fonte: Elaboração própria com base nos dados do Censo Escolar

4.2 Estratégia empírica

A ideia fundamental é captar o efeito da variável explicativa - x_i - na variável explicada - y_i . Para isso, inicialmente, foi usado o modelo de regressão linear bivariada.

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Em que, i indexa alunos, o coeficiente β_0 é o termo constante da regressão e o coeficiente β_1 é o parâmetro de maior interesse, uma vez que mostra a relevância de x_i para y_i . Já a variável u_i é o termo de erro.

Tabela 2 - Regressão linear 1

| <i>Dependent variable:</i> | |
|----------------------------------|--------------------------------|
| Distorção idade-série | |
| Anos de estudo em escola privada | -0.362*** (0.005) |
| Constant | 1.287*** (0.005) |
| Observations | 128,227 |
| R ² | 0.040 |
| Adjusted R ² | 0.040 |
| Residual Std. Error | 1.552 (df = 128225) |
| F Statistic | 5,286.892*** (df = 1; 128225) |
| <i>Note:</i> | * p<0.1; ** p<0.05; *** p<0.01 |

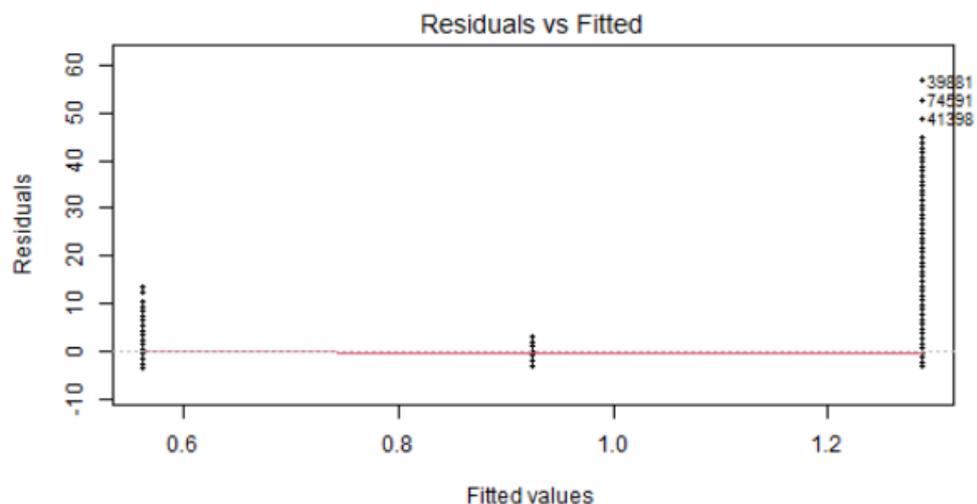
Fonte: Elaboração própria com base nos dados do Censo Escolar

A tabela 2 apresenta os resultados da regressão, sugerindo que, conforme o aluno estuda por mais anos em uma instituição privada, ele tem menor distorção idade-série, pois o coeficiente β_1 é negativo, o que era um resultado esperado. Além disso, existe uma boa significância tanto em β_0 , como em β_1 .

É possível ainda avaliar o coeficiente de determinação R^2 , que representa a qualidade do ajuste. Estando sempre entre 0 e 1, em termos gerais, R^2 mostra quanto da variável dependente é explicado pelo regressor. Entretanto, o R^2 observado no modelo é 0,04, o que é muito baixo.

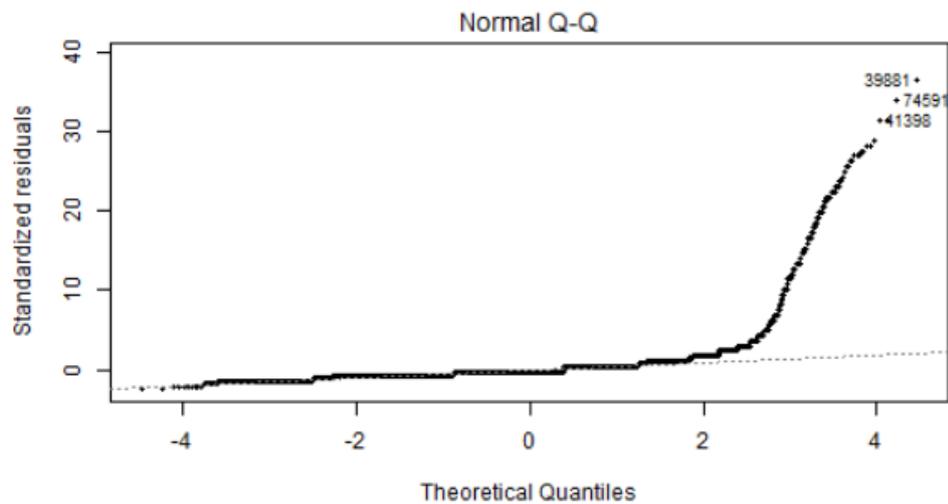
Podemos notar ainda que o modelo não segue as condições requeridas de um modelo de regressão linear. Basicamente, os resíduos devem ter uma distribuição normal, ter variância constante para todos os valores de x_i (homocedasticidade) e ser independentes do regressor. Entretanto, os gráficos abaixo refletem o oposto. No primeiro gráfico, temos os resíduos em função dos valores estimados e, no segundo gráfico, temos a distribuição dos resíduos e a distribuição normal – em linha pontilhada – para avaliar se ambas são similares. O modelo não segue as hipóteses, podendo ser considerado um modelo fraco.

Gráfico 1 – Resíduos e valores ajustados da regressão 1



Fonte: Elaboração própria com base nos dados do Censo Escolar

Gráfico 2 – Resíduos da regressão 1 e distribuição normal



Fonte: Elaboração própria com base nos dados do Censo Escolar

É razoável incluir variáveis de controle associadas, formando um modelo de regressão linear multivariada. O novo modelo é:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + u_i$$

Onde x_{1i} é a variável de anos em escola privada e x_{ki} , onde $k \in [2,4]$, são dummies apontando para as características pessoais dos alunos dadas pela base do Censo Escolar, sendo:

$$x_{2i} \begin{cases} 1, \text{ se o aluno } i \text{ for do sexo masculino} \\ 0, \text{ caso contrário} \end{cases}$$

$$x_{3i} \begin{cases} 1, \text{ se o aluno } i \text{ for negro} \\ 0, \text{ caso contrário} \end{cases}$$

$$x_{4i} \begin{cases} 1, \text{ se o aluno } i \text{ residir em zona urbana} \\ 0, \text{ caso contrário} \end{cases}$$

Nesse caso, os resultados serão:

Tabela 3 - Regressão linear 2

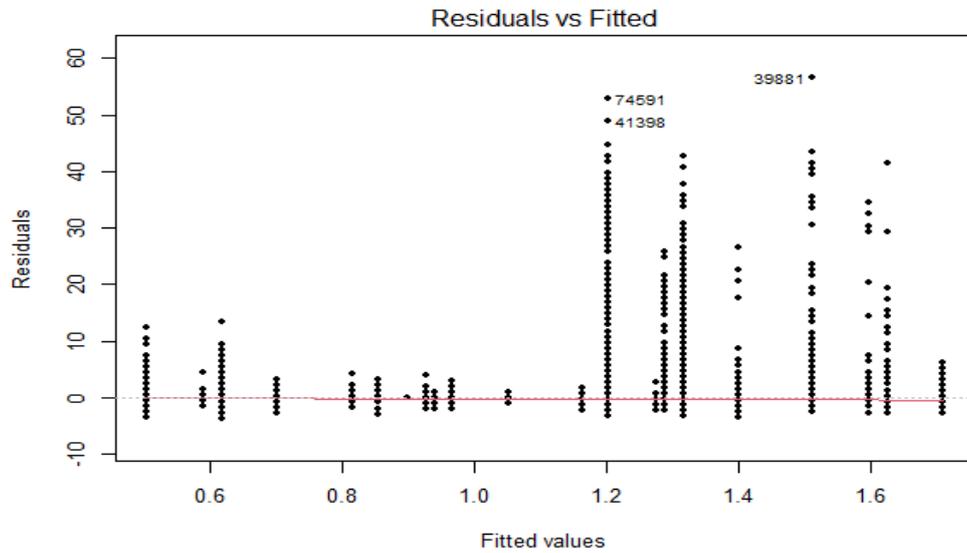
| | <i>Dependent variable:</i> |
|----------------------------------|--------------------------------|
| | Distorção idade-série |
| Anos de estudo em escola privada | -0.349*** (0.005) |
| Sexo | 0.112*** (0.009) |
| Cor | 0.310*** (0.017) |
| Zona residencial | -0.085*** (0.022) |
| Constant | 1.288*** (0.021) |
| Observations | 128,227 |
| R ² | 0.044 |
| Adjusted R ² | 0.043 |
| Residual Std. Error | 1.549 (df = 128222) |
| F Statistic | 1,457.955*** (df = 4; 128222) |
| <i>Note:</i> | * p<0.1; ** p<0.05; *** p<0.01 |

Fonte: Elaboração própria com base nos dados do Censo Escolar

Fazendo uma breve análise, é constatado que os coeficientes da regressão permanecem com uma boa significância e β_1 continua sendo negativo, mas é menor em termos absolutos, dando a entender que na regressão anterior havia um viés de variável omitida e o coeficiente foi superestimado. Aparentemente a inclusão de variáveis de controle também aumentou a qualidade do resultado, pois o R² teve um pequeno acréscimo.

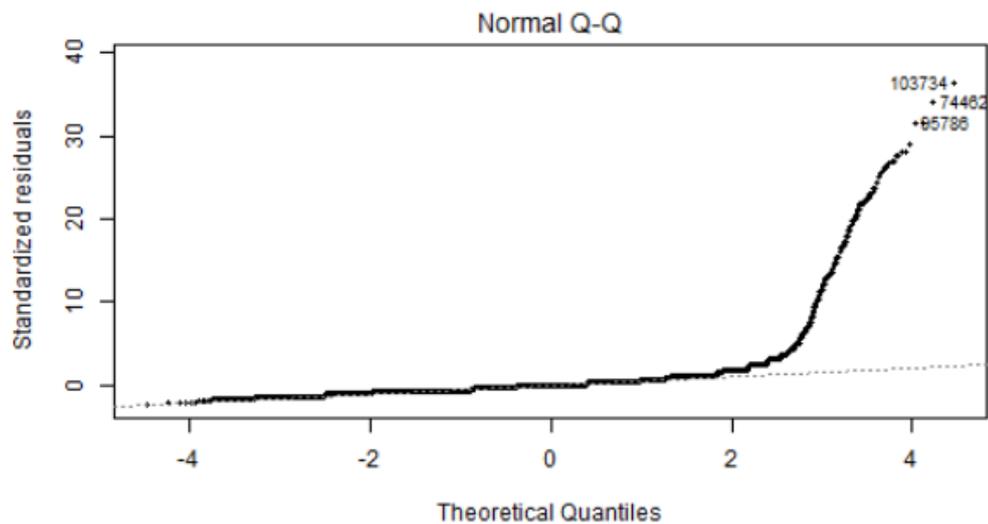
Explorando as hipóteses do modelo de regressão, conforme os gráficos a seguir, houve uma melhora nos quesitos de homocedasticidade e independência em relação aos regressores, tornando a expectativa dos resíduos condicional aos x_{ki} com $k \in [1,4]$ mais próxima de zero. Já a hipótese de normalidade de resíduos continua longe da realidade de acordo com segunda gráfico a seguir.

Gráfico 3 – Resíduos e valores ajustados da regressão 2



Fonte: Elaboração própria com base nos dados do Censo Escolar

Gráfico 4 – Resíduos da regressão 2 e distribuição normal



Fonte: Elaboração própria com base nos dados do Censo Escolar

A tabela 3 compila as informações sobre as regressões que foram rodadas até o momento para comparativo mais adequado.

Tabela 3 - Regressões simples e composta

| | <i>Dependent variable:</i> | |
|----------------------------------|--------------------------------|-------------------------------|
| | Distorção idade-série | |
| | (1) | (2) |
| Anos de estudo em escola privada | -0.362*** (0.005) | -0.349*** (0.005) |
| Sexo | | 0.112*** (0.009) |
| Cor | | 0.310*** (0.017) |
| Zona residencial | | -0.085*** (0.022) |
| Constant | 1.287*** (0.005) | 1.288*** (0.021) |
| Observations | 128,227 | 128,227 |
| R ² | 0.040 | 0.044 |
| Adjusted R ² | 0.040 | 0.043 |
| Residual Std. Error | 1.552 (df = 128225) | 1.549 (df = 128222) |
| F Statistic | 5,286.892*** (df = 1; 128225) | 1,457.955*** (df = 4; 128222) |
| <i>Note:</i> | * p<0.1; ** p<0.05; *** p<0.01 | |

Fonte: Elaboração própria com base nos dados do Censo Escolar

Para obter um produto mais crível, podemos manter o modelo de regressão composta, mas atentar para pontos mais específicos. Para além das características pessoais dos alunos, pode-se incluir variáveis de controle referentes ao meio em que vivem, como um efeito fixo de município. Para isso, foram criadas dummies para cada município do estado do Rio de Janeiro, exceto um município qualquer, para não ter o problema de multicolinearidade perfeita. Ainda, reparando a estatística descritiva, apresentada na Tabela 1, é observada a presença de alguns outliers. Existem alunos com distorção idade-série até 58 anos, então é aceitável restringir a base de dados para indivíduos com a distorção até 15 anos, para que não sejam captados os casos muito particulares.

$$y_i = \beta_0 + \beta_1 x_{1i} + \sum \beta_k x_{ki} + u_i$$

Onde, $k \in [2,94]$, por englobar também os municípios do estado.

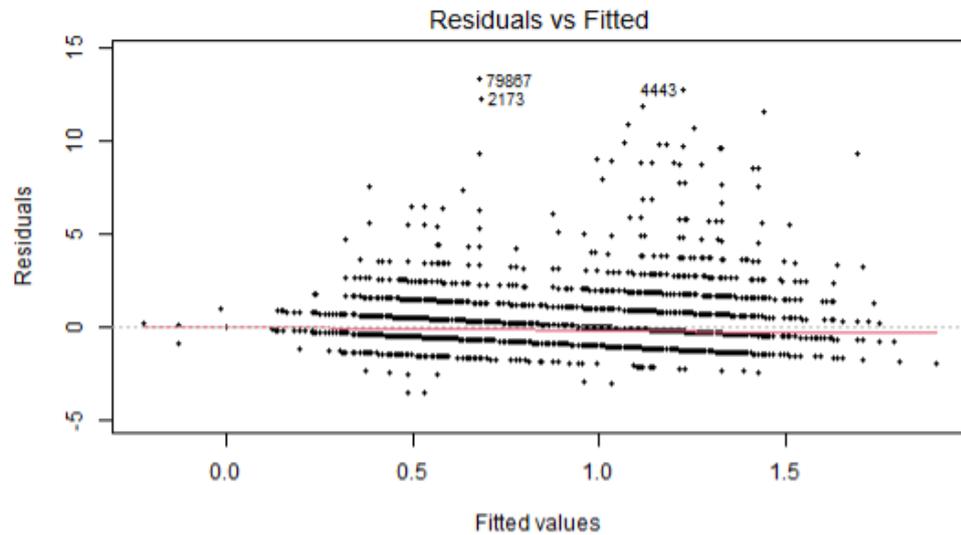
Tabela 3 - Regressão linear 3

| | <i>Dependent variable:</i> |
|----------------------------------|--------------------------------|
| | Distorção idade-série |
| Anos de estudo em escola privada | -0.324*** (0.007) |
| Sexo | 0.114*** (0.009) |
| Cor | 0.211*** (0.028) |
| Zona residencial | -0.014 (0.033) |
| Constant | 1.045*** (0.042) |
| Observations | 33,374 |
| R ² | 0.096 |
| Adjusted R ² | 0.093 |
| Residual Std. Error | 0.857 (df = 33278) |
| F Statistic | 37.156*** (df = 95; 33278) |
| <i>Note:</i> | * p<0.1; ** p<0.05; *** p<0.01 |

Fonte: Elaboração própria com base nos dados do Censo Escolar

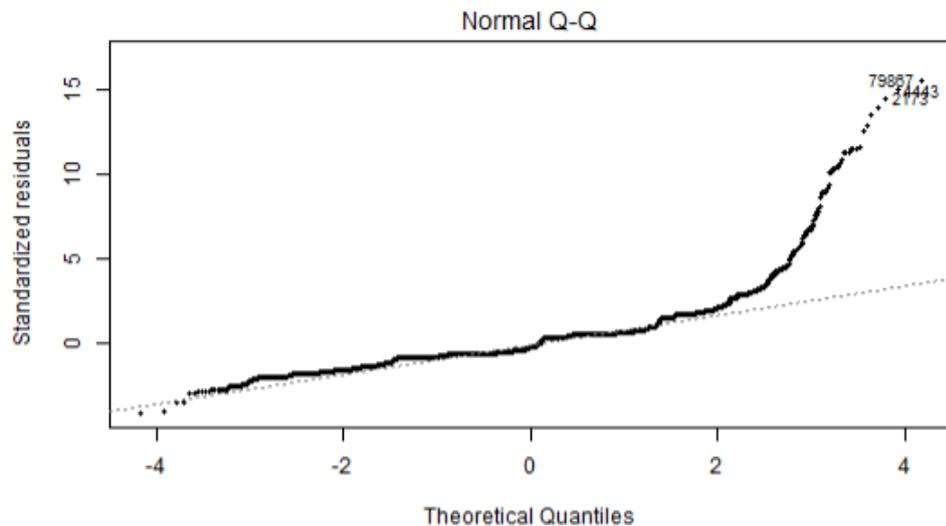
A amostra acaba por sofrer uma redução, principalmente por conta da variável de efeito fixo de município, pois boa parte dos alunos não tem essa informação compartilhada, enquanto que restrição da variável de distorção idade-série retirou apenas 222 observações. Por outro lado, há um aumento mais substancial do R², passando a ser 0,096, o β_1 continua sendo inferior a zero e as hipóteses de regressão linear são melhor atendidas, como provam os gráficos em seguida, embora ainda tenha uma fuga da distribuição normal em pontos mais extremos.

Gráfico 3 – Resíduos e valores ajustados da regressão 3



Fonte: Elaboração própria com base nos dados do Censo Escolar

Gráfico 4 – Resíduos da regressão 3 e distribuição normal



Fonte: Elaboração própria com base nos dados do Censo Escolar

Assim, a estratégia de regressão múltipla se adequa de maneira mais apropriada a situação tratada, mas existem problemas que precisam ser comentados. É possível afirmar que existe uma relação negativa entre as variáveis, confirmando que alunos que estudam por mais tempo em escolas particulares tem distorção idade-série menor, mas não é, necessariamente, uma relação de causalidade. Outro ponto importante é que nem todas as variáveis que impactam indicadores educacionais, como renda familiar e nível de educação dos pais, foram incluídas no modelo por não ser um dado disponibilizado pelo Censo, podendo gerar um viés. Ainda existe

uma dificuldade em garantir aleatoriedade, pois, em geral, o perfil de aluno que opta pela escola pública é diferente daquele que opta pela escola privada.

5 CONCLUSÃO

O objetivo do trabalho é constatar a presença de um gap no que diz respeito a qualidade de ensino público e privado. A pesquisa feita no capítulo anterior busca fundamentar a discussão usando como indicador de qualidade a distorção idade-série e examinando como estudo em escola privada se relaciona ao mesmo.

O resultado dos modelos evidenciou uma relação significativa e negativa entre as variáveis, embora existam simplificações que podem ser uma problemática, como foi citado no final do último capítulo. Mesmo assim, é um fato que um indicador adverso se correlaciona contrariamente ao estudo em escola privada. Ou seja, tomando a escola privada como um tratamento em relação a administração governamental, existe um retorno positivo nesse tratamento, confirmando as suposições feitas a princípio.

REFERÊNCIAS

BECKER, G. S. Human Capital Revisited. 1993.

BEYER, A. L., SOLER, R. O Capital Humano nas Organizações: Uma Revisão Bibliográfica. 2019.

CRAWFORD, R. Na Era do Capital Humano: O Talento, A Inteligência e O Conhecimento Como Forças Econômicas, Seu Impacto nas Empresas e nas Decisões de Investimento. 1994.

HANUSHEK, E. A., WOESSMANN, L. Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. 2012.

HANUSHEK, E. Education Quality And Economic Growth, In The Bush Institute. The Four Percent Solution. 2012

JONES, C. I. Introdução à Teoria do Crescimento Econômico. 2014.

LYONS, J. B. Do School Facilities Really Impact a Child's Education? 2001

MANKIW, N. G. , ROMER D. , WEIL D. A Contribution to the Empirics of Economic Growth. Quarterly Journal of Economics. 1992.

NERI, F. Schooling, Labour Force Quality and the Growth of Nations: Comment. Austrália: Department of Economics University of Wollongong. 2003.

POP-ELECHES, C., URQUIOLA, M., Going to a Better School: Effects and Behavioral Responses. 2013.

SCHULTZ, T. Investment in Human Capital, The American Economic Review. 1961.

SOLOW, ROBERT M. A contribution to the theory of economic growth. The quarterly journal of economics. 1956.

Longaigh, Ciara Ní. Uma Introdução à Educação Pública no Rio de Janeiro. RioOnWatch, 17 de fev. de 2017. Disponível em: <
<https://rioonwatch.org.br/?p=24544#:~:text=No%20Rio%20a%20educa%C3%A7%C3%A3o%20segue,alunos%20de%206%2D14%20anos>>.

<http://cdn.novo.qedu.org.br/uf/33-rio-de-janeiro>

<https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar>

<http://portal.mec.gov.br/conheca-o-ideb>

Santos, Valéria. Distorção Idade-Série Em Meio Às Políticas Educacionais Neoliberais. 2020.

MENEZES, N. Estimando os Gastos Privados Com Educação No Brasil. 2012.

WOOLDRIDGE, J. M. Introdução à Econometria. Uma Abordagem Moderna. 2010.