

**UNIVERSIDADE DO VALE DO RIO DOS SINOS (UNISINOS)
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA
NÍVEL DOUTORADO**

FLÁVIO ELIZIARIO DE SOUZA

**DESEMPENHO DOS ESTUDANTES NO EXAME NACIONAL DO ENSINO MÉDIO
(ENEM): Uma análise a partir da Função de Produção Educacional.**

PORTO ALEGRE - RS

2023

FLÁVIO ELIZIARIO DE SOUZA

**DESEMPENHO DOS ESTUDANTES NO EXAME NACIONAL DO ENSINO MÉDIO
(ENEM): Uma análise a partir da Função de Produção Educacional.**

Tese apresentada como requisito parcial para obtenção do título de Doutor em Economia, pelo Programa de Pós-Graduação em Economia da Universidade do Vale do Rio dos Sinos (UNISINOS).

Orientador(a): Prof. Dr. Tiago Wickstrom Alves

PORTO ALEGRE - RS

2023

S729d Souza, Flávio Eliziario de.
Desempenho dos estudantes no Exame Nacional Do Ensino Médio (Enem) : uma análise a partir da função de produção educacional / por Flávio Eliziario de Souza. – 2023.
142 f. : il. ; 30 cm.

Tese (doutorado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Economia, Porto Alegre, RS, 2023.
“Orientador: Dr. Tiago Wickstrom Alves”.

1. Data envelopment analysis (DEA). 2. Eficiência. 3. Exame Nacional do Ensino Médio (Enem). 4. Função de produção educacional. 5. Estudantes. 6. Capital humano. 7. Capital físico. I. Título.

CDU: 330.356.4:373.5

FLÁVIO ELIZIARIO DE SOUZA

**DESEMPENHO DOS ESTUDANTES NO EXAME NACIONAL DO ENSINO MÉDIO
(ENEM): Uma análise a partir da Função de Produção Educacional.**

Tese apresentada como requisito parcial para obtenção do título de Doutor em Economia, pelo Programa de Pós-Graduação em Economia da Universidade do Vale do Rio dos Sinos (UNISINOS).

Aprovado em 27/09/2023.

BANCA EXAMINADORA

Prof. Dr. José Roberto Iglesias – UFRGS

Prof. Dr. Marcus Tadeu Caputi Leis - UNISINOS

Profa. Dra. Angélica Massuquetti – UNISINOS

Prof. Dr. Tiago Wickstrom Alves (Orientador) – UNISINOS

Dedico esta Tese a minha querida mãe, Ivone Maria Cabral de Souza (*In memoriam*), que sempre me incentivou a lutar e não desistir.

AGRADECIMENTOS

Meu agradecimento a Deus que sempre me deu forças para continuar as lutas da vida.

Meu agradecimento a minha mãe Ivone Maria Cabral de Souza (*in memoriam*) que desde a tenra infância me incentivou a estudar e a superar as dificuldades. Ao meu pai, João Elizario de Souza pelos ensinamentos.

Meu agradecimento especial a minha esposa, Joyce Danielle Batista Martins de Souza que de forma incondicional me apoiou durante todo esse percurso. Aos meus filhos, benção de Deus, Pedro Elizario Martins, Samuel Elizario Martins e Manuela Elizario Martins que me alegraram nos momentos mais difíceis.

Meu agradecimento especial ao meu orientador, Prof. Dr. Tiago Wickstrom Alves, que me proporcionou tantos aprendizados.

RESUMO

O objetivo desta pesquisa foi avaliar a contribuição dos insumos da função de produção educacional para o desempenho acadêmico dos estudantes no Exame Nacional do Ensino Médio (Enem). O estudo defende a tese de que tanto os recursos de capital físico como os recursos de capital humano são determinantes do desempenho acadêmico, porém, que os recursos de capital humano são mais relevantes para o desempenho dos estudantes. A Análise Envoltória de Dados (DEA) orientada a produto e com retornos variáveis de escala foi utilizada como metodologia para a determinação da eficiência dos Institutos Federais (IFs). A estatística descritiva foi aplicada e também empregado o modelo denominado DEA geral com os insumos de capital humano, capital físico e nível socioeconômico dos estudantes e como produto as notas do Enem. Além disso, foi usado o modelo denominado de DEA espacial, acrescido de indicadores regionais no nível dos estados. A análise abrangeu os dados do Enem de 2019 e contou com uma amostra de 460 campi dos Institutos Federais distribuídos em todas as unidades da federação. Os resultados permitiram identificar que há diferença na distribuição dos recursos de capital humano, tendo a região Norte o menor nível de titulação e experiência docente, seguida da região Nordeste com o segundo menor nível de titulação. Ademais, ambas as regiões concentraram os menores níveis socioeconômicos em relação aos estudantes. Apesar disso, as regiões Norte e Nordeste destacam-se pelo maior número relativo de unidades eficientes. Com a análise do modelo DEA geral, apenas um terço dos IFs alcançaram a eficiência. Com a inclusão das variáveis no nível dos estados, modelo DEA espacial, o número de IFs eficientes aumentou para 54,78%, indicando que o efeito do contexto regional impacta nos IFs de modo a torná-los mais eficientes ao mesmo tempo que altera o aproveitamento de seus insumos. Esse efeito revela que devem ser levados em consideração, na composição da função de produção educacional, não apenas os insumos no nível das instituições de ensino e no nível dos estudantes, mas, sobretudo, as diferenças regionais no nível dos estados. Além disso, foi possível identificar que os IFs ineficientes apresentam folgas em todos os insumos, contudo, de forma mais acentuada para os indicadores de infraestrutura geral e titulação, indicando que esses insumos não estão sendo utilizados de modo a ampliar o desempenho. Com isso, a pesquisa permitiu concluir que não há como precisar se o capital humano é mais relevante do que o capital físico para explicar o

desempenho, porque ambos os insumos contribuem tanto para a eficiência quanto para a ineficiência dos IFs quando não são utilizados na sua totalidade.

Palavras-chave: DEA; Eficiência; Enem; Função de Produção Educacional.

ABSTRACT

The objective of this research was to evaluate the contribution of inputs from the educational production function to the academic performance of students in the National High School Exam (Enem). The study brings the thesis that both physical capital resources and human capital resources are determinants of academic performance, however, that human capital resources are specifically more relevant to student performance. Data Envelopment Analysis focused on the product (DEA) with variable returns to scale was used as a methodology to determine the efficiency of Federal Institutes (IFs). Descriptive statistics were applied, and the model called general DEA was also used with the inputs of human capital, physical capital and socioeconomic level of students and as a product the Enem grades. Furthermore, the model called spatial DEA was used, added to regional indicators considering the States level. The analysis covered Enem data from 2019 and included a sample of 460 campuses of Federal Institutes distributed across all units of the federation. The results allowed us to identify that there is a difference in the distribution of human capital resources, with the North region having the lowest level of qualifications and teaching experience, followed by the Northeast region with the second lowest level of qualifications. Furthermore, both regions had the lowest socioeconomic levels in relation to students. Despite this, the North and Northeast regions stand out for the greater relative number of efficient units. With the overall DEA model analysis, only a third of IFs achieved efficiency. With the inclusion of variables at the state level, spatial DEA model, the number of efficient IFs increased to 54.78%, indicating that the effect of the regional context impacts on IFs in order to make them more efficient while changing the utilization of its inputs. This effect reveals that, in the composition of the educational production function, not only inputs at the level of educational institutions and at the level of students must be considered, but, above all, regional differences at the level of states. Additionally, it was possible to identify that inefficient IFs present gaps in all inputs, however, more significantly for the general infrastructure and titling indicators, showing that these inputs are not being used in order to increase performance. As a result, the research allowed us to conclude that there is no way to determine whether human capital is more relevant than physical capital in explaining performance, because both inputs contribute to both the efficiency and inefficiency of IFs when they are not used in their full capacity.

Keywords: DEA; Efficiency; Enem; Educational Production Function.

LISTA DE FIGURAS

Figura 1 -Formas possíveis para um conjunto de produção	24
Figura 2 - Eficiência e Ineficiência Técnica	28
Figura 3 - Representação comparativa das fronteiras BCC e CCR	29
Figura 4 - Modelo conceitual da tese	55
Figura 5 - Localização dos Institutos Federais de Educação, Ciência e Tecnologia no Brasil – 2021	64
Figura 6 - Tratamento dos dados da pesquisa	68

LISTA DE GRÁFICOS

Gráfico 1 - Distribuição dos indicadores de Titulação e Experiência Docente dos IFs (2019).....	72
Gráfico 2 - Distribuição dos indicadores de Titulação Docente dos IFs por regiões do país (2019).....	73
Gráfico 3 - Distribuição dos indicadores de Experiência Docente dos IFs por regiões do país (2019).....	74
Gráfico 4 - Distribuição dos indicadores de Infraestrutura Física dos IFs (2019).....	75
Gráfico 5 - Distribuição do Indicador de Infraestrutura Geral em âmbito nacional e por regiões do país.....	76
Gráfico 6 - Distribuição do Indicador de Infraestrutura de Aprendizagem dos IFs (2019) em âmbito nacional e por regiões do país.	76
Gráfico 7 - Distribuição do Indicador de Infraestrutura Audiovisual em âmbito nacional e por regiões do país.	77
Gráfico 8 - Distribuição do Indicador de Infraestrutura Profissional em âmbito nacional e por regiões do país.	78
Gráfico 9 - Indicadores Socioeconômicos por regiões de localização dos IFs (2019).	82
Gráfico 10 - Distribuição das notas no Enem por área do conhecimento dos IFs (2019).....	85
Gráfico 11 - Distribuição da média de notas dos alunos por IF, por área de conhecimento e regiões do país na prova de 2019.....	86
Gráfico 12 - Distribuição dos indicadores de titulação e experiência docente por região e localidade (2019).....	90
Gráfico 13 - Distribuição dos indicadores de capital físico classificados por região e localidade dos IFs (2019).....	91
Gráfico 14 - Distribuição dos indicadores socioeconômicos classificados por região e localidade dos IFs (2019).....	92
Gráfico 15 - Distribuição das notas no Enem classificadas por região e localidade dos IFs (2019).....	93

Gráfico 16 - Frequência dos escores de eficiência técnica dos IFs (2019) por quantidade e percentual analisadas.....	99
Gráfico 17 - Percentuais dos potenciais de melhoria dos IFs (2019) com eficiência menor que 100 por área de conhecimento (outputs) do Modelo DEA-BCC-Output.	100
Gráfico 18 - Percentuais de folgas por indicadores de capital físico, humano e nível socioeconômico do Modelo DEA-BCC-Output para 293 IFs com eficiência menor que 100% dos IFs (2019).	101
Gráfico 19 - Distribuição percentual das folgas dos insumos de Titulação e Experiência dos 293 IFs (2019) com escores ineficientes por regiões do país.	103
Gráfico 20 - Distribuição percentual das folgas dos insumos de Infraestrutura Geral e de Aprendizado dos IFs (2019) com escores ineficientes por regiões do país.	104
Gráfico 21 - Distribuição percentual das folgas dos insumos de Infraestrutura audiovisual e Profissional dos IFs (2019) com escores ineficientes por regiões do país.	105
Gráfico 22 - Distribuição dos escores de eficiência dos IFs (2019) em âmbito nacional e por regiões do país.	107
Gráfico 23 - Escores de eficiência dos IFs (2019) por regiões do país das 460 unidades avaliadas.....	108
Gráfico 24 - Número de Institutos Federais com 50% ou mais de unidades eficientes (2019).....	109
Gráfico 25 - Frequência comparativa dos escores de eficiência dos IFs (2019) entre os modelos DEA geral e DEA espacial.	111
Gráfico 26 - Comparação da distribuição de frequência dos escores de eficiência dos IFs (2019) por regiões do país entre os modelos DEA geral e DEA espacial.	112
Gráfico 27 - Percentuais de folgas dos insumos do Modelo DEA geral e DEA espacial com dados regionais para 208 IFs com eficiência menor que 100% (2019).	113
Gráfico 28 - Distribuição dos potenciais de melhoria das áreas avaliadas no Enem de acordo com o Modelo DEA geral e DEA espacial com dados regionais para 208 IFs (2019) com eficiência menor que 100.	116
Gráfico 29 - Distribuição dos potenciais de melhoria das áreas de Linguagem e Matemática por região de acordo com o modelo DEA espacial para 208 IFs (2019) com eficiência menor que 100.....	117

LISTA DE QUADROS

Quadro 1 - Método de pesquisa utilizado nos estudos empíricos	45
Quadro 2 - Variáveis dependentes e independentes utilizadas nos estudos empíricos	47
Quadro 3 - Variáveis de recursos escolares agrupados por dimensão	60
Quadro 4 - Variáveis utilizadas no modelo DEA.....	62
Quadro 5 - Microdados servidores	66
Quadro 6 - Variáveis de controle por autor conforme Revisão Sistemática de Literatura	69

LISTA DE TABELAS

Tabela 1 - Resumo do Parâmetro de Coeficientes dos Gastos de 92 estudos de Função de Produção Educacional de Países em Desenvolvimento	37
Tabela 2 - Método de análise identificado na RSL por quantidade e percentual.....	47
Tabela 3 - Correlação de Pearson das variáveis relacionadas ao Capital Humano e ao Capital Físico dos IFs (2019).....	79
Tabela 4 - Correlação dos Indicadores Socioeconômicos por regiões de localização dos IFs (2019)	83
Tabela 5 - Correlação de Pearson das variáveis de proficiência no Enem com as variáveis socioeconômicas, capital humano e infraestrutura física dos IFs (2019) ...	88
Tabela 6 - Descrição das folgas percentuais dos indicadores de capital humano e capital físico dos IFs (2019) por regiões do país	106
Tabela 7 - Distribuição percentual das médias das folgas por região em relação aos indicadores de nível escolar, regional e individual dos Institutos Federais (IFs) ineficientes no ano de 2019 – Modelo DEA Espacial	115

LISTA DE SIGLAS

BCC	Banker, Charnes e Cooper
CCR	Charnes, Cooper e Rhodes
CH	Ciências Humanas
CN	Ciências da Natureza
CO	Centro-Oeste
CSV	<i>Comma Separated Values</i>
DE	Dedicação Exclusiva
DEA	<i>Data Envelopment Analysis</i>
DMU	<i>Decision Making Unit</i>
EAFA	Escola Agrotécnica Federal
EBTT	Ensino Básico, Técnico e Tecnológico
EJA	Educação de Jovens e Adultos
ENEM	Exame Nacional do Ensino Médio
EPF	<i>Education Production Function</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
IDEB	Índice de Desenvolvimento da Educação Básica
IDHM	Índice de Desenvolvimento Humano Municipal
IECD	Indicador de Experiência do Corpo Docente
IEE	Indicador de Efeito Escola
IFAC	Instituto Federal do Acre
IFAL	Instituto Federal de Alagoas
IFAM	Instituto Federal do Amazonas
IFES	Instituto Federal do Espírito Santo
IFF	Instituto Federal Fluminense
IFMG	Instituto Federal de Minas Gerais
IFPA	Instituto Federal do Pará
IFPI	Instituto Federal do Piauí
IFPR	Instituto Federal do Paraná
IFRO	Instituto Federal de Rondônia
IFRR	Instituto Federal de Roraima
IFs	Institutos Federais de Educação, Ciência e Tecnologia

IFSP	Instituto Federal de São Paulo
IFSul	Instituto Federal Sul-rio-grandense
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
INFRA_AP	Infraestrutura de Aprendizado
INFRA_AUD	Infraestrutura de Aprendizado
INFRA_G	Infraestrutura Geral
INFRA_PR	Infraestrutura Profissional
INSE	Índice Socioeconômico do Estudante
ITCD	Indicador de Titulação do Corpo Docente
LC	Linguagem e Códigos
MEC	Ministério da Educação
MT	Matemática
N	Norte
NE	Nordeste
NSE	Nível Socioeconômico
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
PCCTAE	Plano de Carreira dos Cargos Técnico-Administrativos em Educação
PIB	Produto Interno Bruto
PISA	<i>Programme for International Student Assessment</i>
PNAD	Pesquisa Nacional por Amostra de Domicílios
PNAES	Programa Nacional de Assistência Estudantil
PNP	Plataforma Nilo Peçanha
RED	Redação
RSC	Reconhecimento de Saberes e Competências
RSL	Revisão Sistemática de Literatura
S	Sul
SAEB	Sistema de Avaliação da Educação Básica
SE	Sudeste
SETEC	Secretaria de Educação Profissional e Tecnológica

SUMÁRIO

1 INTRODUÇÃO	18
1.1 Definição do problema de pesquisa	19
1.2 Objetivos	21
1.2.1 Objetivo geral	21
1.2.1 Objetivos específicos	21
1.3 Justificativa	21
1.4 Delimitação do tema	22
1.5 Estruturação da Tese	23
2 REVISÃO DE LITERATURA	24
2.1 Referencial teórico	24
2.1.1 Função de produção neoclássica.....	24
2.1.2 Função de produção educacional	29
2.2 Estudos empíricos dos determinantes do desempenho acadêmico	32
2.2.1 Desempenho acadêmico.....	32
2.2.2 <i>Beckground</i> familiar.....	34
2.2.3 Recursos escolares	36
2.2.4 Efeito escola.....	40
2.2.5 Efeito dos pares e clima escolar.....	42
2.2.6 Outras características que impactam no desempenho	43
2.2.7 Método de análise utilizado nos estudos empíricos	45
2.2.8 Principais variáveis dependentes e independentes dos estudos empíricos.....	47
2.3 Síntese do capítulo e formulação da tese	53
3 METODOLOGIA	57
3.1 Análise Envoltória de Dados - DEA	57
3.2 Variáveis usadas no modelo DEA	58
3.3 Dados da pesquisa	63
3.3.1 Amostra.....	63
3.3.2 Base de dados	64
3.3.3 Tratamento dos dados.....	68
3.4 Variáveis de controle	69
4 ANÁLISE DOS RESULTADOS	71
4.1 Estatística descritiva	71

4.1.1 Estatística descritiva dos indicadores de Capital Humano e Capital Físico.....	71
4.1.2 Estatística descritiva dos indicadores de nível socioeconômico.....	80
4.1.3 Estatística descritiva dos indicadores de proficiência no ENEM	84
4.1.4 Descrição das variáveis de capital humano e físico por região e localidade	89
4.1.5 Síntese da seção da análise das estatísticas descritivas	94
4.2 Análise dos determinantes da função de produção educacional.....	98
4.3 Análise DEA espacial	110
4.4 Síntese da análise de eficiência.....	118
5 CONCLUSÃO	123
REFERÊNCIAS.....	126
APÊNDICE A – PROCESSO DE SELEÇÃO DOS ESTUDOS EMPÍRICOS	138

1 INTRODUÇÃO

Testes padronizados têm sido utilizados para avaliar a eficácia de sistemas de ensino ou o desempenho acadêmico¹, como, por exemplo, as provas do Sistema de Avaliação da Educação Básica (SAEB) e o Exame Nacional do Ensino Médio (Enem). Em nível internacional, tem-se o PISA (*Programme for International Student Assessment*) promovido pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE) (JESUS; LAROS, 2004). Esses são exemplos de alguns dos testes padronizados que são utilizados no Brasil para a avaliação educacional (LAROS; MARCIANO; ANDRADE, 2012).

Existem diversos fatores intraescolares que atuam para estabelecer o nível de desempenho dos estudantes, como formação dos professores, existência de laboratórios, núcleos de formação docente e de atendimento aos alunos, entre outros (SOARES, 2004). Entretanto, há também outros elementos além das condições da escola que contribuem para o desempenho dos estudantes, entre eles, a condição socioeconômica do estudante, o nível de formação dos pais, etc. (COLEMAN *et al.* 1966; SOARES, 2004).

O tema desta tese é justamente avaliar a eficiência das instituições de ensino com base nos diferentes fatores que contribuem para explicar o desempenho dos estudantes, agregando aos tradicionais, mencionados anteriormente, também fatores da região onde se inserem as escolas, como cultura e renda, por exemplo. Esta análise estrutura-se considerando os alunos do ensino técnico integrado de nível médio dos Institutos Federais, que fazem parte da Rede de Educação Profissional e Tecnológica, e como medida de desempenho as notas desses estudantes no Exame Nacional do Ensino Médio (Enem). Esse tema está circunscrito no problema de pesquisa que se explicita na seção que segue.

¹ Desempenho acadêmico e proficiência são termos que estão sendo utilizados nesta tese com o mesmo significado, ou seja, ambos indicando o resultado dos estudantes em um teste padronizado em larga escala, como o ENEM.

1.1 Definição do problema de pesquisa

Os pesquisadores que atuam na área da educação vêm se debruçando sobre os determinantes do desempenho dos estudantes, em especial, após o Relatório Coleman, realizado em 1966. Nessa publicação, a principal evidência foi a constatação da preponderância das condições socioeconômicas dos alunos ou do *background* familiar em contraposição à relação dos fatores ligados aos recursos escolares (COLEMAN *et al.* 1966). A partir desses resultados, houve uma intensa discussão e contraposição feita por pesquisadores gerando uma rica literatura concernente ao tema.

Contudo, as pesquisas subsequentes não chegaram a um consenso sobre quais são os elementos mais influentes na determinação do desempenho dos estudantes no nível escolar (SOARES; SÁTYRO, 2008; LEE, 2010). Hanushek ao fazer uma revisão sistemática de literatura sobre o impacto dos recursos escolares verificou que havia pouca relevância estatística nas análises (HANUSHEK, 1997). O autor deu como exemplo os EUA, um país desenvolvido que, entre 1960 e 2000, triplicou os gastos em educação sem obter um incremento no desempenho de forma significativa (HANUSHEK, 1986, 2003). Essa mesma lógica tem sido evidenciada por outros autores para países em desenvolvimento (GLEWWE *et al.*, 2011).

Os esforços das pesquisas mais recentes permitem reafirmar os determinantes de nível socioeconômico, como: educação do pai, educação da mãe, renda familiar, ocupação dos pais, número de pessoas na família, cor e infraestrutura doméstica (HANUSHEK, 1989; ALBERNAZ; FERREIRA; FRANCO, 2002; BARROS *et al.*, 2001; FELÍCIO; FERNANDES, 2005; MENEZES-FILHO, 2007; GLICK; SAHN, 2009; CURI; MENEZES-FILHO, 2013; SOUZA; OLIVEIRA; ANNEGUES, 2018; ASSUNÇÃO; ARAÚJO; ALMEIDA, 2019). E ainda apontar alguns determinantes relacionados a recursos escolares, como: esforço do gestor, políticas educacionais, infraestrutura e pessoal (FRANCO *et al.* 2007; RIANI; RIOS-NETO, 2008; SOARES; SÁTYRO, 2008; COSTA; ARRAES; GUIMARÃES, 2015; TAVARES; CAMEL; PACIÊNCIA, 2018; MORAES; DIAS; MENEZES, 2020).

Os determinantes de nível socioeconômico, nesse sentido, já estão em boa medida consolidados, reafirmando o Relatório Coleman, em relação positiva e direta com o desempenho. Essa ligação evidencia a importância de no nível do estudante haver condições socioeconômicas favoráveis. Já do ponto de vista da escola, ainda

se debate os determinantes e seus efeitos, contudo, espera-se um resultado relevante dos recursos de infraestrutura e, principalmente, da atuação do professor. Esses são os efeitos positivos esperados, no entanto, condições desfavoráveis de nível socioeconômico são mais desafiadoras para a escola de acordo com suas características.

Com isso, não há clareza sobre a contribuição de cada fator no nível escolar, da eficiência no uso dos recursos e, em especial, do que falta ou sobra de recursos que contribuem ou não para o desempenho. Glewwe *et al.* (2011), em um estudo de meta-análise, identificou resultados controversos, para alguns deles, foram significativos os recursos de infraestrutura geral da escola e as variáveis relacionadas ao professor. Todavia, quando os autores diminuíram a amostra para artigos compreendidos de alta qualidade, os resultados foram inconclusivos. No contexto brasileiro, Franco e Menezes Filho (2017) alcançaram resultados similares.

Nesse mesmo sentido, Dewey, Husted e Kenny (2000), buscando um padrão nas pesquisas realizadas por outros pesquisadores, classificaram os estudos analisados em bons e ruins e excluíram a variável renda da função de produção, obtendo dois terços de estudos com impactos positivos de recursos escolares. Com isso, mesmo que não haja convergência no que se refere aos determinantes de nível escolar, é possível identificá-los a partir da literatura disponível, principalmente, variáveis relacionadas à infraestrutura escolar (capital físico) e ao professor (capital humano). Não é consenso, contudo, o seu efeito sobre o desempenho (LEE, 2010).

Portanto, considerando o contexto acerca dos efeitos dos determinantes do desempenho acadêmico no que tange ao uso adequado dos recursos, no nível escolar, e a necessidade de se compreender o impacto dos recursos de capital físico e capital humano e de verificar as carências do sistema, gerando informações que podem balizar as políticas públicas para ampliar o desempenho dos estudantes, o problema de pesquisa desta tese é expresso na seguinte questão: Quais determinantes da função de produção educacional são mais escassos e, por consequência, limitam o desempenho dos estudantes?

1.2 Objetivos

1.2.1 Objetivo geral

O objetivo geral desta tese é avaliar a contribuição dos insumos da função de produção educacional para o desempenho acadêmico dos estudantes no Exame Nacional do Ensino Médio (Enem).

1.2.1 Objetivos específicos

- a) determinar os componentes da função de produção escolar;
- b) determinar as folgas e restrições dos determinantes da eficiência escolar na função de produção.

1.3 Justificativa

Este estudo é essencial porque a educação é fundamental para o desenvolvimento dos indivíduos, o que pode ser observado a partir do ponto de vista da economia. Países com melhor nível de produtividade e, por conseguinte, renda, são aqueles em que a educação é de melhor qualidade. Melhores níveis de qualificação influenciam nos ganhos e no crescimento econômico em que o diferencial de renda pode ser explicado pelo diferencial de escolaridade (VIANA; LIMA, 2010; SALVATO; FERREIRA; DUARTE, 2010). Contudo, a significância da educação vai além dos anos de escolaridade, sendo mais importante a qualidade do processo formativo.

Hanushek e Wobmann (2007) confirmam essa tese ao defenderem que as diferenças entre as escolas quanto à qualidade acabam por impactar no crescimento econômico, distinguindo países desenvolvidos de países em desenvolvimento, isto é, a mesma escolaridade em diferentes países leva a desempenho divergente. Isso se torna mais premente porque há maior escassez de recursos em países em desenvolvimento que são aqueles com maior necessidade de elevar seus ganhos médios. Dessa forma, a avaliação da eficiência no uso dos recursos de modo a aumentar o desempenho é um tema indispensável.

O estudo contribui, assim, para a compreensão dos determinantes e seus efeitos, e da relevância e efetividade desses recursos para o desempenho dos estudantes no processo educacional. Em especial, favorece o entendimento sobre aqueles relacionados à infraestrutura física (capital físico) e aqueles relacionados a pessoal (capital humano) no nível escolar, controlados pelo nível socioeconômico dos estudantes. Além disso, destacam-se as contribuições no nível metodológico por usar variáveis da condição socioeconômica e indicadores regionais na análise envoltória de dados (DEA) como insumos no processo de eficiência das unidades de ensino analisadas.

Outros aspectos são primordiais, principalmente, porque há ainda muitas disparidades entre as regiões do país quanto ao nível de escolaridade da população, com taxas de analfabetismo em maiores proporções para as regiões Norte e Nordeste (IBGE, 2020). Evidencia-se, assim, o desafio da oferta de uma educação de qualidade num país com dimensões continentais e com realidade socioeconômica bem distinta. Diante desse contexto, o país investiu, nos últimos anos, cerca de 6% do PIB (Produto Interno Bruto), contudo, sem conseguir posições de destaque nas avaliações do PISA (*Programme for International Student Assessment*). Por essas razões, há a necessidade de melhor se compreender o processo educacional visando o desempenho com base na análise de eficiência.

Por fim, o estudo foi viável porque houve a disponibilidade de dados para análise das variáveis inerentes ao capital físico e ao capital humano das instituições de ensino, das variáveis de nível socioeconômico e das variáveis regionais no nível dos estados, conforme consta no capítulo de metodologia. Além disso, houve viabilidade porque foi possível o processamento dos dados com o suporte computacional.

1.4 Delimitação do tema

Esta pesquisa foi realizada no âmbito dos Institutos Federais de Educação, Ciência e Tecnologia (IFs), parte da Rede Federal de Educação Profissional e Tecnológica, no contexto nacional com todos os IFs nas 27 unidades federativas. O estudo foi direcionado aos estudantes da educação profissional técnica de nível médio na forma integrada. O estudo foi realizado com dados do ano de 2019, antes, portanto, da pandemia de Covid-19 ocorrida a partir de 2020.

1.5 Estruturação da Tese

Além desta introdução, a tese foi estruturada em quatro capítulos. O capítulo 2 contém uma discussão acerca da teoria da função de produção educacional e aborda os estudos empíricos que tratam dos determinantes do desempenho acadêmico, concluindo o capítulo com a formalização da tese. Na sequência, tem-se o capítulo 3 que trata da metodologia da pesquisa, o método de análise envoltória de dados (DEA) utilizado, a construção dos indicadores e as bases de dados. No capítulo 4, abordam-se os resultados a partir da análise descritiva demonstrando as principais características das unidades analisadas, a análise de eficiência geral denominada de DEA geral e a análise com a inclusão de indicadores regionais no nível dos estados denominado de DEA espacial. Por fim, no capítulo 5, apresenta-se a conclusão da tese.

2 REVISÃO DE LITERATURA

Este capítulo tem por objetivo a revisão da literatura com uma abordagem teórica sobre a função de produção neoclássica e a função de produção educacional; e uma abordagem empírica, trazendo estudos que analisam o desempenho acadêmico dos estudantes.

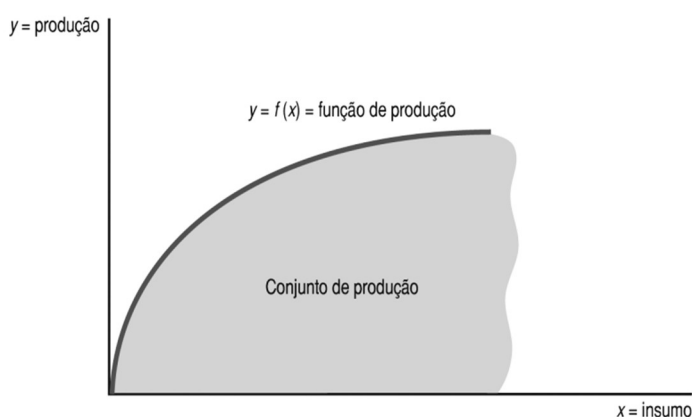
2.1 Referencial teórico

Esta seção teórica está dividida em duas partes, a primeira analisa a função de produção neoclássica e a segunda analisa a função de produção educacional.

2.1.1 Função de produção neoclássica

A função de produção educacional deriva da função de produção microeconômica da firma. Essa função relaciona basicamente o produto máximo que uma empresa produz, dada uma combinação específica de insumos (BESANKO; BRAEUTIGAM, 2010; VARIAN, 1992, 2012; PINDYCK; RUBINFELD, 2013). A função de produção descreve o limite possível para todas as combinações de insumos, conforme apresentado na figura 1. Consiste, portanto, no limite máximo de produção (VARIAN, 2012). Trata-se de uma relação em que há uma escolha na quantidade do insumo x , que resulta em um produto y , sendo assim uma relação $y = f(x)$. Essa função permite verificar o limite de produção, dados a tecnologia e os insumos pela firma, em que diferentes fatores de produção determinam um resultado.

Figura 1 -Formas possíveis para um conjunto de produção



Fonte: Varian (2012).

Mas-Colell (1995) expõe algumas propriedades desses conjuntos: (i) Y é não vazio, significa que a firma apenas tem um plano de produção. (ii) Y é fechado, o conjunto Y inclui o seu limite, que é a fronteira tecnologicamente possível, que pode ser expresso como $y^R \rightarrow y$ e $y^n \in Y$ implica que $y \in Y$. (iii) *No free lunch*, a ideia é que não é possível produzir algo do nada, assim $y \in Y$ e $y \geq 0$, então $y = 0$. (iv) Possibilidade de inação, $0 \in Y$, ou seja, a firma pode escolher não operar desde que não haja tomado nenhuma decisão de produção, caso contrário, operará em custos irre recuperáveis. (v) Livre disposição (*free disposal*), a possibilidade de usar mais insumos sem alterar a quantidade produzida é sempre possível, se $y \in Y$ e $y' \leq y$, então $y' \in Y$.

Ainda, de acordo com Mas-Colell, tem-se, (vi) irreversibilidade, sendo $y \in Y$ e $y \neq 0$, então $-y \notin Y$, isto é, um vetor de um conjunto de produção que transforma insumos em produtos não pode operar de forma inversa, produtos em insumos. (vii) Retornos não crescentes de escala, isto é, $y \in Y$ e $\alpha y \in Y$ para todos os escalares $\alpha \in [0,1]$ (tecnologia divisível). (viii) Retornos não decrescentes de escala, para $y \in Y$ em que $\alpha y \in Y$ para toda escala $\alpha \geq 1$ (tecnologia replicável). (ix) Retornos constantes de escala, em que $y \in Y$ implica $\alpha y \in Y$ para algum escalar² $\alpha \geq 0$ (tecnologia replicável e divisível). (x) Livre entrada, a ideia é que dois planos de produção factíveis podem ser executados de formas independentes, assim se $y \in Y$ e $y' \in Y$ então essa propriedade requer que $y + y' \in Y$. (xi) Convexidade, em que Y é convexo; assim se Y tem essa propriedade, então um vetor de produção $y, y' \in Y$ e $\alpha \geq 0$ e $\beta \geq 0$, tem-se $\alpha y + \beta y' \in Y$.

Dadas as propriedades de um conjunto de produção factível, a função de produção, simplificada, pode ser representada a partir da combinação de dois insumos (PINDYCK; RUBNFIELD, 2013). Na análise com dois fatores de produção, por exemplo, trabalho (L) e capital (K), tem-se:

$$Y = F(L, K) \tag{1}$$

A equação (1) permite verificar que a função de produção representa o limite de produto que se pode produzir para qualquer combinação possível de insumos, dada a tecnologia de produção. Em outras palavras, "as funções de produção

² MAS-COLELL, A.; WHINSTON, M. D.; GREEN, J. R. (1995).

descrevem o que é tecnicamente viável quando a empresa opera eficientemente, ou seja, quando utiliza cada combinação de insumos da forma mais eficaz possível” (PINDYCK; RUBNFIELD, 2013, p.195).

Shephard (1970) apresenta a função de produção como uma forma matemática definida no conjunto de insumos de produção de uma tecnologia com propriedades que seguem a família do conjunto $L(u), u \in [0, +\infty)$, em que $L(u)$ consiste no conjunto de insumos de produção. Segundo o autor, para qualquer vetor de insumo $x \in D$ deve-se considerar uma função $\Phi(x)$ definida no conjunto $L(u)$, sendo que a função de produção $\Phi(x)$ significa a maior taxa de produção possível dados os insumos x , como segue:

$$\Phi(x) = \text{Max}\{u | x \in L(u), u \in [0, +\infty)\}, x \in D \quad (2)$$

Quanto à relação temporal, na microeconomia, o curto prazo se refere a quando pelo menos um fator de produção é fixo (ALBUQUERQUE, 1986; VARIAN, 2012; PINDYCK; RUBNFIELD, 2013;). Assim, essa função (1) com dois insumos operada no curto prazo, dá a ideia de que pelo menos um fator é fixo. Nesse sentido, a relação entre capital e trabalho dependerá do tempo e da atividade desenvolvida, pois em um espaço curto de tempo a firma tende a manter Capital (K) fixo em detrimento da variação do fator trabalho (L). Como exemplo, para uma dada firma intensiva em bens de capital, a decisão de aplicação de mais de um fator não será instantânea. Por outro lado, no longo prazo, todos os fatores são variáveis. A produtividade da firma nesse processo pode ser mensurada por meio do produto médio e marginal do trabalho (BESANKO; BRAEUTIGAN, 2010). O produto médio do fator trabalho é dado por:

$$\frac{Y}{L} > 0 \quad (3)$$

Em que Y é o produto total e L a quantidade total de trabalho utilizado, sendo que a relação dá a medida de produtividade em que se tem o resultado médio de produto por unidade de trabalho (ALBUQUERQUE, 1986). Quanto maior a quantidade produzida por menor unidade de trabalho mais produtiva é a firma. Essa relação (4) é apresentada por Besanko e Braeutigam (2010) da seguinte forma:

$$AP_L = \frac{\text{produto total}}{\text{quantidade de trabalho}} = \frac{Y}{L} \quad (4)$$

Outra medida de produtividade consiste no produto marginal que se constitui na taxa resultante da variação total do produto dividido pela variação total na quantidade do fator de produção da firma. Sendo o trabalho empregado, tem-se, matematicamente, a relação dada por $\frac{\partial F}{\partial L}$, que dá a inclinação da reta tangente à função de produção (ALBUQUERQUE, 1986). Didaticamente, de acordo com Besanko e Braeutigam (2010), pode ser dada por:

$$MP_L = \frac{\text{mudança no total de produto}}{\text{mudança na quantidade de trabalho}} = \frac{\Delta Y}{\Delta L} \quad (5)$$

Em que MP_L é a produtividade marginal do trabalho; ΔY a variação na quantidade; e ΔL a variação no fator trabalho. Essa medida dá o comportamento do produto em relação à adição de uma unidade de trabalho. Quando o produto marginal é igual ao produto médio, tem-se que o produto marginal é mais alto do que a produtividade do fator de produção (ALBUQUERQUE, 1986). Outra ideia que está vinculada à função de produção diz respeito aos rendimentos marginais decrescentes, pois, mantendo-se um insumo constante, ao se adicionar uma unidade do insumo variável, os aumentos na produção serão gradativamente menores (PINDYCK; RUBNFIELD, 2013). Ou seja, a lei dos rendimentos marginais decrescentes consiste no “princípio segundo o qual quando o uso de um insumo produtivo aumenta, mantendo-se os demais insumos fixos, a partir de dado momento, as resultantes adições ao produto serão cada vez menores” (PINDYCK; RUBNFIELD, 2013. p.200).

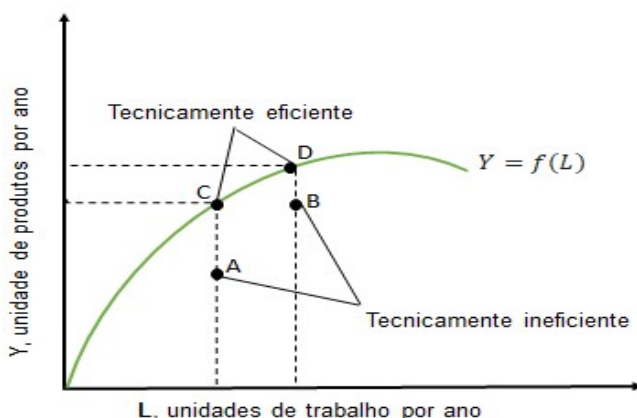
Quando se analisa as decisões das firmas a longo prazo, como visto, todos os fatores de produção são variáveis. Então, um aumento no produto dar-se-ia a partir da liberdade para combinar os insumos. Assim, ao passo que esses dobrem, esperar-se-ia também um produto duas vezes maior. Isso nem sempre ocorre e a explicação econômica está nos retornos de escala: “Os retornos de escala referem-se à proporção de aumento do produto quando os insumos aumentam proporcionalmente entre si” (PINDYCK; RUBNFIELD, 2013. p.213). Eles podem ser de três formas: retornos de escala crescentes, constantes e decrescentes.

De forma simples, os retornos crescentes de escala constituem a situação em que havendo um aumento proporcional em todos os insumos, há um aumento mais que proporcional na produção; já os retornos constantes de escala, numa situação em que haja um aumento proporcional em todos os insumos, há o mesmo aumento

proporcional na produção; e os retornos decrescentes de escala indicam a situação em que havendo um aumento proporcional em todos os insumos, há um aumento menos do que proporcional na produção (PINDYCK; RUBNFIELD, 2013). Por sua vez, Varian (1992) apresenta que os retornos dependem da tecnologia, ou seja, serão retornos constantes ou não.

A função de produção é o limite da capacidade produtiva, ela representa a própria fronteira de eficiência produtiva. A figura 2 reproduz essa ideia, ao evidenciar as firmas C e D como eficientes e as firmas A e B como ineficientes. A partir da fronteira de eficiência, surge a Análise Envoltória de Dados (*Data Envelopment Analysis - DEA*), principalmente, decorrente das contribuições de Farrell (1957), Charnes, Cooper e Rhodes (1978) e Banker, Charnes e Cooper (1984), estes com o avanço nos modelos de fronteira de produção determinística por meio de técnicas de programação linear, para vários recursos e resultados.

Figura 2 - Eficiência e Ineficiência Técnica

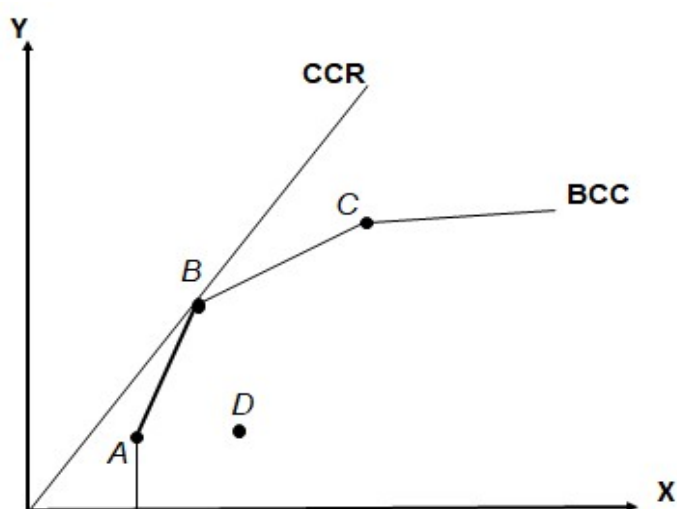


Fonte: Adaptado de Besanko e Braeutigan (2010).

A DEA é uma técnica não paramétrica por meio de programação matemática com a finalidade de construir fronteiras de produção através de processos tecnológicos aproximadamente homogêneos (CASADO, 2007). Tem-se o modelo CCR em referência às iniciais dos nomes dos autores (CHARNES, COOPER e RHODES, 1978), orientado a insumos (*input*) e a produtos (*output*) com retorno constante de escala. Há ainda o modelo BCC, em referência a Banker, Charnes e Cooper (1984), podendo ser orientado tanto a insumos (*inputs*) quanto a produtos (*outputs*) para retornos variáveis de escala (CHARNES; COOPER; RHODES, 1978; BANKER; CHARNES; COOPER, 1984; MELLO et al., 2005; CASADO, 2007).

A diferença entre o modelo CCR e o BCC está nos retornos de escala. Enquanto o modelo CCR trabalha com retornos constantes de escala, o modelo BCC atua com retornos não constantes de escala (retornos crescentes de escala e retornos decrescentes de escala). A figura 3 permite a comparação entre os dois modelos, o ponto B apresenta a intersecção entre os dois, os demais pontos seguem um retorno não constante, enquanto o ponto D é tido como ineficiente nesse processo.

Figura 3 - Representação comparativa das fronteiras BCC e CCR



Fonte: Adaptado de Mello *et al.* (2005).

A partir da função de produção da firma, tem-se, portanto, a fronteira de eficiência teórica, segundo a teoria microeconômica. Conforme observado na figura 2, nem todas as unidades produtivas operam de forma eficiente. Há algumas firmas que podem operar dentro do conjunto de possibilidades de produção abaixo da curva de eficiência. Com a análise envoltória de dados, é possível verificar a eficiência relativa e, assim, construir uma fronteira de eficiência de acordo com as unidades de decisão que possuem características semelhantes (*Decision Making Unit – DMU*). Dadas as principais implicações da função de produção microeconômica, passa-se à função de produção educacional na próxima seção.

2.1.2 Função de produção educacional

Quando se analisa uma instituição educacional, busca-se logo saber qual o seu objetivo e, por conseguinte, qual o seu resultado esperado. Esse objetivo segue uma

dada tecnologia que possibilita transformar um conjunto de insumos em um dado resultado, como o resultado acadêmico em testes de proficiência. Nessa perspectiva, a função de produção educacional permite verificar a relação entre insumos (*input*) e resultados (*output*) (ALBERNAZ; FERREIRA; FRANCO, 2002; AMMERMÜLLER; HEIJKE; WÖßMANN, 2005; CURI; SOUZA, 2015). Contudo, o escopo de uma instituição de ensino é diferente do de uma firma, basicamente porque a oferta é essencialmente estatal (ALTINOK, 2007). O produto ou o resultado, nesse sentido, pode ser definido por estudantes com determinado grau de proficiência, conforme planejamento curricular, no tempo esperado.

Todo esse processo conduz naturalmente ao debate sobre a qualidade da educação, algo bastante discutido e por vezes sem uma definição clara. Para a teoria da função de produção educacional (*Education Production Function -EPF*), essa medida pode ser alcançada pelo indicador de proficiência dos estudantes ou pelo indicador de desempenho no mercado de trabalho. Dada a dificuldade de uso deste último, o primeiro é comumente utilizado, pois há a identificação de correlação com o desempenho econômico dos países (ALTINOK, 2007). A EPF, portanto, fornece uma estrutura conceitual orientadora da escolha lógica das variáveis, permitindo uma interpretação coerente de seus efeitos (TODD; WOLPIN, 2003).

O Relatório Coleman (COLEMAN *et al.* 1966), denominado de Igualdade de Oportunidades Educacionais (título original *Equality of Educational Opportunity*), valeu-se da função de produção educacional para explicar as diferenças entre características escolares com o desempenho para diferentes segmentos estudantis. O estudo é tido como precursor na pesquisa dos determinantes do desempenho acadêmico pela sua repercussão quanto aos resultados, principalmente, pela preponderância dos fatores socioeconômicos em relação ao peso dos recursos escolares (WALTENBERG, 2006; CURI; SOUZA, 2015).

A EPF é fundamental na identificação dos determinantes do desempenho educacional (qualidade da educação), daí o porquê de se investir em pesquisas nessa área pelo fato de que as informações de avaliação esclarecem por onde se deve caminhar, confirmando os fatores impactantes no desempenho acadêmico dos estudantes por meio de evidências científicas. Em outros termos, o conhecimento oriundo a partir da utilização da função de produção educacional é essencial para a alocação de recursos (BOWLES, 1970). Assim, a partir das ideias elencadas sobre EPF, Bowles (1970) formalmente apresenta-a como:

$$A = f(X_1, \dots, X_m, X_{m+1}, \dots, X_v; X_w, \dots, X_z) \quad (6)$$

Em que A define alguma medida de saída da escola (*output*); X_1, \dots, X_m são variáveis de medida do ambiente escolar, como as características de infraestrutura da escola, qualificação dos professores e tempo de exposição dos estudantes aos recursos escolares (*input*); X_{m+1}, \dots, X_v são variáveis representativas do aprendizado fora da escola, como realizado pelos pais; e X_w, \dots, X_z são variáveis representativas das habilidades inatas dos estudantes antes de sua entrada na escola.

Hanushek (1979), por sua vez, ao fazer referência à função de produção educacional, critica a abordagem de Coleman por não apresentar um modelo conceitual subjacente. Para o autor, dadas as diferenças do contexto da firma, especialmente, por alguns insumos em educação serem fixos (a exemplo de trabalho e capital), vê a necessidade de uma lógica subjacente no ambiente educacional, desse modo, conceitualmente apresenta o seguinte modelo:

$$A_{it} = f(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i) \quad (7)$$

Assim: (i) A_{it} é o desempenho do estudante i no tempo t ; (ii) $B_i^{(t)}$ é um vetor da influência do *background* familiar no tempo t ; (iii) $P_i^{(t)}$ é o vetor de influência dos pais no tempo t ; (iv) $S_i^{(t)}$ é o vetor de insumos escolares acumulado no tempo t ; (v) I_i é um vetor de habilidades inatas. Como esse modelo é analisado em dado tempo, ele é uma função de produção cumulativa, ou seja, os insumos pretéritos têm efeitos duradouros os quais diminuem com a distância temporal (HANUSHEK, 1979). No entanto, em vez de se ter uma medida de resultado num dado ponto no tempo, o autor propõe uma mudança na análise do desempenho que considere as diferenças entre o tempo t e t^* , ou seja, o foco está no valor adicionado, correspondente à mudança no produto de um momento para o outro, como segue:

$$A_{it} = f^*(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)}, I_i, A_{it^*}) \quad (8)$$

De forma mais simples, no trabalho de Albernaz, Ferreira e Franco (2002), a EPF é apresentada da seguinte forma:

$$y = F(c, m, g, p, s) \quad (9)$$

Em que y denota o desempenho dos alunos sujeito aos seguintes fatores: características pessoais dos alunos como raça e gênero (c); características de suas

famílias como renda ou nível socioeconômico (m); características dos colegas de escola (g); características dos professores como escolaridade e salário (p); e outras características da escola (s). Ou seja, o produto é basicamente o desempenho e/ou proficiência do estudante a partir de características dos estudantes, da família, dos colegas, dos professores e da escola.

A função de produção educacional tem a ideia básica de possibilitar a análise da transformação de insumos em um resultado mensurável, como o desempenho acadêmico medido por meio de testes padronizados e, conforme visto, a maioria de suas versões são bem próximas. Complementarmente, outros estudos, portanto, seguem essa mesma lógica (GYIMAH-BREMPONG; GYAPONG, 1991; FIGLIO, 1999; COOPER; COHN, 1997; PRITCHETT; FILMER, 1999; TODD; WOLPIN, 2003; GLEWWE; KREMER, 2006; GLEWWE *et al.* 2011).

Para a compreensão das principais variáveis inerentes à atividade educacional, a próxima seção apresenta as variáveis tratadas nos estudos empíricos.

2.2 Estudos empíricos dos determinantes do desempenho acadêmico

Esta seção apresenta os estudos empíricos relacionados aos determinantes do desempenho dos estudantes. Aborda-se o desempenho acadêmico, o *background* familiar, os recursos escolares e o efeito escola. Para esse fim, utilizou-se tanto os principais artigos sobre o tema como a Revisão Sistemática de Literatura (RSL), conforme Kitchenham (2004), apêndice A.

2.2.1 Desempenho acadêmico

Ao se tratar de desempenho acadêmico vem à tona o tema sobre a qualidade da educação, o qual traz consigo a dificuldade do conceito, longe de consenso (TRAVITZKI, 2013). Algumas indagações, a partir da leitura de Travitzki (2013), são importantes. A primeira diz respeito à dificuldade conceitual de se definir o que é uma pessoa bem-educada. A segunda refere-se ao que é uma pessoa inteligente, dados os vários tipos de inteligência. Os dois questionamentos levam a escolhas, o que representa bem o papel da escola. Portanto, as escolhas devem ser realizadas no âmbito educacional, levando em consideração a limitação de recursos.

Nessa linha, Soares (2009) afirma que a educação e a qualidade são delimitadas a partir da educação escolar com a priorização de competências necessárias à vida, de natureza cognitiva, como compreensão leitora, matemática, conhecimentos científicos e culturais, além das competências de ordem social, afetiva e ética. O autor refere-se ao professor Simon Schwartzman, “o professor Simon é pioneiro no Brasil na defesa de que os resultados concretos na educação consistem no aprendizado dos alunos”, além disso, aponta que os fatos sociais devem se fundamentar em sólidas evidências empíricas (SOARES, 2009, p. 217).

Verifica-se, assim, que as habilidades cognitivas são significativas no processo de aprendizado escolar e, portanto, podem ser mensuradas com fundamento em evidências empíricas. Mas qual seria um indicador factível nesse processo? Altinok (2007) apresenta o desempenho no mercado de trabalho como uma primeira possibilidade, isto é, os salários são o retorno sobre o investimento em capital humano acumulado ao longo do tempo (MINCER, 1996). De forma geral, o autor apresenta a seguinte relação:

$$\ln w_{ij} = \alpha_{i0} + r_i K_{ij} \quad (10)$$

Sendo que: (i) w_{ij} é o salário do indivíduo i no tempo j ; (ii) K_{ij} é o produto acumulado do investimento do indivíduo i no período j (capital humano acumulado período); (iii) r_i representa a taxa de retorno; (iv) α_{i0} representa o log individual dos indivíduos sem capital humano acumulado, que, segundo Mincer (1996), são baixos em economias industriais.

Nessa abordagem, a educação da população a partir dos níveis de qualificação acaba por influenciar nos ganhos, na produtividade e no crescimento econômico (VIANA; LIMA, 2010). Salvato, Ferreira e Duarte (2010), ao utilizarem dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 1999, com o intuito de analisar a relação entre a escolaridade e a distribuição de renda, indicam que entre 12% e 36% do diferencial de renda é explicado pelo diferencial de escolaridade. Cangussu, Salvato e Nakabashi (2010), ao analisarem o capital humano sobre o nível de renda dos estados brasileiros, apontam para a relevância dos resultados da teoria do capital humano na explicação do PIB *per capita*.

Contudo, a análise direta do mercado de trabalho não se mostra prática, pois os resultados são vistos após anos da escolaridade realizada, dificultando assim avaliar essa relação (HANUSHEK, 2006). Por isso, uma segunda possibilidade ao

produto do processo educacional, a partir de Altinok (2007), pode ser a proficiência dos estudantes em testes internacionais, pois, segundo esse autor, há estudos que demonstram que as habilidades cognitivas são bons indicadores da renda futura dos alunos. Por conseguinte, no contexto microeconômico, sugere-se como factível o desempenho educacional a partir de testes padronizados como medida de qualidade.

Sendo, portanto, a proficiência um indicador possível, cabe olhar para o interior da escola, para o resultado a partir da qualidade dos serviços ofertados. Hanushek e Wobmann (2007) apresentam a qualidade da educação como significativa para o crescimento econômico, havendo, por exemplo, diferença no nível da qualidade da escolaridade entre países desenvolvidos e países em desenvolvimento. A mesma quantidade de anos de estudo pode ser diferente em termos de qualidade entre as nações. Isso desloca o foco da escolaridade para o que acontece no interior da escola, para a forma como os recursos são aplicados.

Uma série de estudos apontam para o uso da proficiência verificada por meio de testes padronizados, sendo a grande maioria para testes em matemática e linguagens (de acordo com a língua materna) (AGÜERO; BELECHE, 2013; BARTHOLO; COSTA, 2016; BECKER, 2011; BENEVIDES; SOARES, 2020; BERNAL; MITTAG; QURESHI, 2016; BEUCHERT *et al.*, 2018; BLANCO, 2011; CARRASCO; MARTÍN, 2012; CHUDGAR; QUIN, 2012; DELPRATO; CHUDGAR, 2018; DRIESSEN; AGIRDAG; MERRY, 2016; GALVÃO, 2018; HÆGELAND; RAAUM; SALVANES, 2012).

Após os apontamentos sobre o desempenho acadêmico com a constatação de sua utilidade como medida de qualidade da educação, o próximo tópico trata dos fatores que podem afetá-lo, em outras palavras, sobre quais as variáveis comumente discutidas pela literatura que têm relação com o resultado dos estudantes.

2.2.2 *Background* familiar

De acordo com a literatura de economia da educação e áreas afins, o *background* familiar impacta significativamente no desempenho acadêmico dos estudantes. Segundo Souza, Oliveira e Annegues (2018, p.134), “o *background* familiar consiste em um conjunto de características do ambiente familiar no qual os indivíduos se desenvolvem (educação e ocupação dos pais, renda familiar etc.)”. Em outros termos, *background* familiar e nível socioeconômico (NSE) são praticamente

equivalentes. O relatório Coleman (1966) inaugurou o debate de que a proficiência dos estudantes é afetada quase que exclusivamente pelo *background* familiar ante a influência da escola, trata-se de um trabalho complexo, com números expressivos e dados coletados de 4.000 escolas, 66.000 professores e aproximadamente 600.000 alunos (HILL, 2017).

O documento foi tão importante por seus resultados inesperados que uma geração se debruçou a partir de um seminário em Harvard, organizado por Daniel Patrick Moynihan, com 80 pesquisadores, para revisar as análises realizadas, um deles foi Eric Hanushek (HANUSHEK, 2016; HILL, 2017). Os resultados apontaram para a força de *background* familiar, fatores como raça, renda e educação dos pais, confirmando as conclusões apontadas por Coleman (HILL, 2017). No entanto, Hanushek (2016), a partir de questões metodológicas do estudo, considerou os achados não como determinantes, mas sim como hipóteses a serem avaliadas.

Todo esse cenário indica que, desde a década de 1960, e após várias pesquisas, o *background* familiar se consolidou como importante fator no desempenho acadêmico. A partir da revisão sistemática de literatura, foi possível evidenciar essa realidade. Masci *et al.* (2017) demonstram que, na Itália, fatores internos das escolas levam a diferentes resultados para o desempenho em leitura e matemática. Para eles, o efeito escola é mais forte em matemática do que em leitura, concluem ainda que é possível identificar uma boa escola, levando-se em consideração a variabilidade dentro das classes de acordo com as três regiões estudadas. A situação socioeconômica dos alunos e o contexto familiar pesam positivamente no desempenho dos estudantes.

Vardardottir (2015), usando uma subamostra suíça dos dados do PISA, investiga a produção educacional com foco na influência que o *status* socioeconômico dos colegas de classe exerce sobre o resultado acadêmico dos alunos em um sistema de *streaming*. Como novidade, o estudo relata que os pares têm um efeito significativo nas habilidades de resolução de problemas dos alunos; o *background* socioeconômico médio dos colegas tem efeito positivo nas pontuações de matemática e resolução de problemas. Além disso, a heterogeneidade dos colegas tem efeito significativo em matemática, o que aponta para a importância das interações sociais no contexto escolar.

Becker (2011), ao estudar as habilidades de linguagem de crianças de 3 a 5 anos, no Reino Unido, ressalta que o desempenho depende da educação dos pais e

que a educação pré-escolar atua positivamente no desenvolvimento dessas competências, ou seja, tanto o nível socioeconômico como a escola se mostram importantes. Seabra, Carvalho e Ávila (2019) buscam compreender, no contexto português, se a composição da escola afeta os resultados dos alunos e se os efeitos variam de acordo com o nível de origem nacional. Os autores concluem que o efeito da composição socioeconômica de uma escola é mais importante no que tange ao desempenho acadêmico dos estudantes do que o efeito de sua composição étnica.

Uma das implicações do peso do nível socioeconômico é que ele afeta o poder de escolha. Sampaio *et al.* (2011), no caso brasileiro, ao estudarem as decisões de evasão de estudantes universitários em relação à renda, ao desempenho no vestibular, a deixar um curso atual e em tentar uma nova carreira por meio de novo vestibular, entendem que a renda é condição essencial na proporção de melhores condições de estudo por meio de acesso a escolas privadas, cursinhos e no poder de escolha da carreira. Acrescentam ainda que o fato de o estudante ter frequentado a escola pública reverbera em menor desempenho e em menores chances de escolha, enquanto ter frequentado um cursinho repercute em melhor desempenho e maiores chances de escolha de um curso superior.

Para a Grã-Bretanha, Schmitt e Wadsworth (2006) identificaram que a posse de computador doméstico tem impacto significativo no desempenho dos estudantes. Esse é um exemplo da relevância da condição socioeconômica, ao evidenciar que famílias mais abastadas têm maior poder de consumo de bens, o que está associado ao desempenho. Destacam-se, portanto, as seguintes variáveis quanto ao *background* familiar: educação do pai, educação da mãe, renda familiar, ocupação dos pais, número de pessoas na família, cor e infraestrutura doméstica (HANUSHEK, 1989; ALBERNAZ; FERREIRA; FRANCO, 2002; BARROS *et al.*, 2001; FELÍCIO; FERNANDES, 2005; MENEZES-FILHO, 2007; GLICK; SAHN, 2009; CURI; MENEZES-FILHO, 2013; SOUZA; OLIVEIRA; ANNEGUES, 2018; ASSUNÇÃO; ARAÚJO; ALMEIDA, 2019).

2.2.3 Recursos escolares

Coleman (1966), ao atestar a preponderância do nível socioeconômico sobre o desempenho dos estudantes, não encontrou evidências quanto aos recursos escolares. Esse foi sem dúvida o resultado mais controverso, gerando os inúmeros

estudos subsequentes na direção de entender a escola na performance dos estudantes. Hanushek (1995), ao analisar o impacto dos recursos escolares por meio de meta-análise, afirma a ineficiência da organização escolar e dos gastos de recursos, não contribuindo para o desempenho do estudante. A Tabela 1 sistematiza essas conclusões, exceto para instalações físicas. O autor afirma que as correções dessas ineficiências não são simples e que não existe um projeto de escola de qualidade que possa ser entregue aos elaboradores de políticas públicas.

Tabela 1 - Resumo do Parâmetro de Coeficientes dos Gastos de 92 estudos de Função de Produção Educacional de Países em Desenvolvimento

Input	Número de estudos	Estatisticamente significante		Estatisticamente insignificante
		Positivo	Negativo	
Taxa professor-aluno	30	8	8	14
Educação do professor	63	35	2	26
Experiência do professor	46	16	2	28
Salário do professor	13	4	2	7
Gasto por aluno	12	6	0	6
Infraestrutura	34	22	3	9

Fonte: Harbinson e Hanushek (1992).

Hanushek segue com os trabalhos de meta-análise. O autor atualizou os estudos até 1994 e os ampliou (estudos anteriores, Hanushek, 1981, 1986, 1989). Após o procedimento, as conclusões demonstraram que não há força nem consistência na relação entre recursos escolares e desempenho dos estudantes (HANUSHEK, 1997). Nessa mesma linha, Glewwe *et al.* (2011), a partir de meta-análise sobre o impacto dos recursos escolares no desempenho dos estudantes em países em desenvolvimento, período compreendido de 1990 a 2010, apresentaram em sua maioria resultados controversos. Em uma amostra de 79 trabalhos, verificou-se que, em alguns deles, os recursos de infraestrutura geral da escola e as variáveis do professor são significativos. Dentre as variáveis tidas como significativas, destacam-se: livros e materiais didáticos, quantidades de mesas e carteiras, despesa com eletricidade, índice de infraestrutura escolar, nível de escolaridade do professor,

conhecimento da disciplina que ministra, treinamento dos professores e tamanho da sala de aula (com ressalva).

No entanto, quando os autores utilizam um filtro para 43 estudos tidos como de elevada qualidade, os resultados são inconclusivos. Ainda assim, a pesquisa evidencia que a descoberta mais importante talvez seja a de que ter uma escola em pleno funcionamento com melhor infraestrutura pode ser propício à aprendizagem (GLEWWE *et al.*, 2011). Nos países em desenvolvimento, portanto, a educação se preocupou inicialmente com a expansão do ensino, isto é, com a quantidade de matrículas nas escolas. Contudo, a partir da aferição da qualidade por meio dos resultados em testes de proficiência, o direcionamento foi dado para a qualidade da educação (GLEWWE; KREMER, 2006).

Nessa mesma lógica de pouca relação identificada entre recursos escolares e desempenho, Franco e Menezes Filho (2017), no Brasil, encontraram resultados favoráveis para as características dos alunos, atraso escolar e condição socioeconômica da família, contudo, sem encontrar significância para as variáveis da escola. Outros estudos, no entanto, apontam para a significância dessa relação. Já para os EUA, alguns estudos identificam a importância do professor (ROCKOFF, 2004; RAVKIN; HANUSHEK; KAIN, 2005; AARONSON; BARROW; SANDER, 2007). Ademais, com evidências da importância econômica, isto é, um docente com um desvio padrão acima da eficácia média pode gerar ganhos marginais anuais de mais de \$400.000 de valor presente dos ganhos futuros de seus alunos (HANUSHEK, 2011).

Hægeland, Raaum e Salvanes (2012), usando variações nos recursos escolares ocasionadas por maior nível de receitas de localidades com usinas hidrelétricas, na Noruega, examinam os efeitos dos gastos escolares sobre o desempenho dos alunos aos 16 anos. Os autores, por meio de estimativas de mínimos quadrados, indicam pouca relação entre recursos financeiros e desempenho. Todavia, ao usarem as receitas das hidrelétricas como variável instrumental, os recursos escolares foram significativamente positivos nas pontuações dos testes, indicando que 10% a mais nas horas de professor por aluno elevaram a nota média em cerca de 0,25, além disso, um aumento estimado em US \$ 1.500,00 de despesa por aluno eleva a nota no exame em 0,2.

Oswald e Backes-Gellner (2014), por sua vez, investigam, na Suíça, o efeito dos incentivos financeiros no desempenho dos alunos e como esse efeito é moderado

pelas preferências temporais. Os autores apontam que os incentivos são mais eficazes no início de um programa educacional, quando os benefícios reais do mercado de trabalho estão mais distantes, pois os alunos impacientes têm melhor resposta a incentivos monetários do que estudantes relativamente pacientes. Em outros termos, há diferença no desempenho a partir da alocação de recursos financeiros aos estudantes, contudo, dependente das preferências temporais.

Bernal, Mittag e Qureshi (2016) estimam o efeito da qualidade da escola no desempenho dos alunos nos EUA por meio de características da escola, a partir de uma extensão de variável *proxy* usada por Black e Smith (2006). Os autores encontram uma relação significativa e positiva entre a qualidade da escola e o desempenho dos alunos se as características da escola, como tamanho da classe e escolaridade dos professores, forem tratadas como medidas ruidosas de qualidade escolar. Afirmam ainda que a qualidade da escola produz resultados, contudo, sendo latente e não observada.

Rocha e Funchal (2019), ao analisarem a relação entre custos diretos e organização da oferta escolar com o desempenho em escolas públicas, a partir de uma amostra composta por escolas brasileiras, indicam que há grande variabilidade nos custos escolares diretos dentro de uma mesma rede de ensino, o que sugere desigualdade na distribuição desses recursos; os custos escolares diretos apresentaram correlações estatisticamente significantes, porém, com coeficientes com baixo poder explicativo.

Nessa mesma linha, Galvão (2018) investiga a relação entre gastos escolares e desempenho escolar em São Bernardo do Campo, no Brasil. Os resultados evidenciam que a relação entre gastos e desempenho não é homogênea entre as categorias de gastos e que os resultados do modelo melhoram na medida da desagregação dos dados. Destacam-se os gastos com coordenação pedagógica em relação positiva e significativa no que tange ao desempenho em matemática e relação negativa entre gastos por aluno com assistentes de classe. Kim e Rhee (2019), ao estudarem os efeitos do fornecimento de banheiro escolar sobre a taxa de frequência à escola primária, no Quênia, mostram, de forma consistente, que um aumento na disponibilidade de banheiro escolar por aluno eleva significativamente a taxa de frequência à escola entre meninos e meninas, sendo o efeito maior para as meninas.

Por fim, alguns estudos alinham-se no sentido de corroborar a importância dos recursos escolares. Costa, Arraes e Guimarães (2015) identificam a importância do

professor no desempenho do aluno. Além disso, outros autores demonstram que esforço do gestor, políticas educacionais, infraestrutura, pessoal, isto é, recursos escolares em geral, melhoram o desempenho médio dos estudantes (FRANCO *et al.* 2007; RIANI; RIOS-NETO, 2008; SOARES; SÁTYRO, 2008; COSTA; ARRAES; GUIMARÃES, 2015; TAVARES; CAMEL; PACIÊNCIA, 2018; MORAES; DIAS; MENEZES, 2020).

2.2.4 Efeito escola

A seção anterior tratou dos recursos escolares de forma geral e de seu impacto no desempenho dos estudantes. Nesta, articula-se o conceito de efeito escola e suas implicações para o desempenho dos estudantes. Para Ferrão e Fernandes (2003), o efeito escola é uma derivação metodológica de valor adicionada à educação originada da economia. De acordo com Moraes, Dias e Menezes (2020), uma característica essencial é que a unidade de análise consiste na escola e não no estudante individualmente, ou seja, a unidade de ensino deve ser analisada com o fim de se observar o quanto acrescenta ao aprendizado dos alunos dado o desempenho esperado.

O efeito escola, assim, trata da questão de como as características escolares impactam no desempenho de seus alunos e, desse modo, podem corroborar a tese de que a escola faz a diferença (LEE, 2010). Pode ainda ser dada a pontuação em uma escala de desempenho atribuída às práticas internas de uma instituição de ensino (SOARES; CANDIAN, 2007). Ou ainda verificar até que ponto frequentar determinada escola altera o resultado de um estudante (RAUDENBUSH; WILLMS, 1995). Segundo os autores, essas são iniciativas que vários países utilizam para manter escolas responsáveis pelas contribuições ao aprendizado dos alunos.

O efeito escola, segundo Willms e Raudenbush (1989), pode ser do Tipo A e do Tipo B. O primeiro utiliza variáveis de controle associadas ao aluno, incluindo o efeito das políticas e práticas escolares, até os efeitos e as influências sociais mais amplas. Há interesse dos pais diante da expectativa do resultado da unidade de ensino sobre a formação de seus filhos. O segundo estima a performance de determinada escola em comparação com outras escolas dadas as suas semelhanças, com a inclusão de variáveis associadas ao contexto do estudante e de fatores de

composição do corpo estudantil. É uma medida comparativa das escolas ponderada pelo seu público. Esse efeito interessa mais a gestores e tomadores de decisão.

Uma das dificuldades de se medir o efeito escola está no produto utilizado, pois os testes de proficiência refletem tanto a ação da escola como a de outras variáveis inerentes a outros contextos da vida do estudante (SOARES; CANDIAN, 2007). Segundo os autores, esse efeito só pode ser estimado com dados longitudinais em que se compara o conhecimento do aluno desde a sua entrada até a sua saída da unidade de ensino. Não obstante essa restrição, Soares e Candian (2007) explicam que, com os dados que se tem hoje, já existe literatura que se baseia neles. Barbosa e Fernandes (2001), por exemplo, encontram resultados satisfatórios do efeito escola da educação básica brasileira referente à proficiência em matemática para os alunos da 4ª série.

Quando se analisa os estudos empíricos sobre o efeito escola nos países desenvolvidos, pode-se observar que um caso de sucesso na educação foi o da Coreia do Sul. Nesse país, os alunos eram alocados às unidades de ensino de forma aleatória, essa especificidade ocorreu porque, nos anos de 1970, o país incrementou um sistema de nivelamento das condições das escolas, dando iguais oportunidades aos seus estudantes (CHOI; MOON; RIDDER, 2014). Do ponto de vista empírico, esses autores derivam uma função de produção educacional em que agregam funções de resultados individuais que dependem de insumos escolares observados e não observados, os quais interagem com habilidades individuais não observadas, apontando para o efeito escola unissex e maior que os estudos anteriores daquele contexto.

Vinha e Laros (2016), ao analisarem como os dados de avaliação educativos para 27 estudos são utilizados por Brasil, Argentina e Chile, apresentam que existe efeito escola significativo para os países analisados, com efeito bruto variando de 19,4% a 46%, e efeito líquido oscilando entre 8,1% e 19%. Rodríguez e Correa (2019), ao avaliar a relevância do contexto municipal sobre o rendimento acadêmico individual na Colômbia, mostram que as escolas chegam a ser responsáveis por 30% das diferenças nos resultados individuais, ao passo que os municípios explicam 5% da variabilidade no rendimento acadêmico. Apontam ainda para a importância do contexto escolar e da escola sobre o desempenho acadêmico como significativo no estudo.

Outra forma de se avaliar o efeito da escola consiste na mensuração do resultado de determinados tipos de organização escolar. Nessa direção, Benevides e Soares (2020), ao examinarem se o efeito escola militar ainda pode ser comprovado em escolas públicas do estado do Ceará, apontam que os resultados são favoráveis e se mostram significativos, mesmo quando se controla os vieses de seleção. É importante destacar que a falta de controle para o viés de seleção e para desempenhos passados aumenta o efeito escola militar em uma vez e meia quando são usados modelos tradicionais.

Especificamente para o Brasil, vários outros estudos corroboram o efeito escola (GAVIRIA; MARTÍNEZ-ARIAS; CASTRO, 2004; HARBISON; HANUSHEK, 1992; LOCKHEED; BRUNS, 1990; MELLO; SOUZA, 2005). Além disso, na América Latina, na Colômbia, Rodríguez-Jiménez e Murillo-Torrecilla (2011) encontraram um efeito escola entre 6% e 23%, maior para leitura do que para matemática. Essa é uma relação que se diferencia dos resultados de países desenvolvidos, principalmente, no que tange à tradição americana de Hanushek. Há imensa desigualdade tanto no nível do aluno, como no nível das escolas. Essa realidade acaba por impactar no desempenho dos estudantes.

De uma perspectiva mais ampla, o efeito escola se alarga a outros resultados dentro do campo educacional. Griew *et al.* (2010), ao avaliar o efeito escola nas práticas de atividade física em crianças, no Reino Unido, concluem que há um efeito significativo da escola na atividade física dos seus estudantes, respondendo por um efeito positivo e significativo de 14,5%. No Canadá, Hobin *et al.* (2010) demonstram que há um efeito da escola nas matrículas da disciplina de educação física. Sansani (2011) explicita, nos EUA, a relação positiva entre qualidade escolar e saúde, isto é, aumentar o salário relativo do professor, em um desvio padrão, resulta em cerca de 1,9 mortes a menos por mil pessoas por ano extra de educação.

2.2.5 Efeito dos pares e clima escolar

A sala de aula e mais a sua composição têm impacto no desempenho. McVicar, Moschion e Ryan (2018) estudam o efeito dos pares na Austrália em relação ao desempenho em alfabetização e fornecem evidências fortes da existência de efeito dos pares, com resultado pequeno, mas positivo, e estatisticamente significativo. O clima também se mostra importante. Melo e Moraes (2019) verificam a relação entre

clima escolar e desempenho em escolas públicas de um estado brasileiro, a partir da percepção de alunos, professores e gestores. Apontam que o clima escolar atenua os efeitos dos baixos níveis socioeconômicos, denotando-se como fator protetivo que possibilita aos alunos melhores condições na conquista de um maior rendimento acadêmico.

Brault, Janosz e Archambault (2014) pesquisam sobre o efeito de clima educacional, composição socioeconômica, étnica e acadêmica das escolas nas expectativas dos docentes quanto ao sucesso dos alunos, no Canadá. Os autores indicam que as percepções dos professores sobre a capacidade dos seus alunos não estavam determinadas apenas por indicadores estruturais dessa capacidade, mas também por uma percepção coletiva de qual envolvimento escolar e desempenho poderiam ser esperados dos estudantes. Os autores relatam que as percepções dos professores estão associadas aos resultados dos alunos.

Diferentemente dos demais estudos, Blanco (2011), para o caso mexicano, investiga se as escolas são relativamente independentes dos condicionamentos do seu ambiente. O autor acaba por demonstrar que grande parte da variância da aprendizagem é explicada por variáveis que levam em consideração a interface escola-ambiente bem como pela interação entre esses fatores. Verifica também outras influências, como a confirmação da alegação de que em parte o efeito escola pode ser explicado pela composição social das escolas, como analisado por Bartholo e Costa (2016) (composição social das escolas), ao pesquisarem a implementação do sistema de avaliação federal e municipal no Brasil, na cidade do Rio de Janeiro.

2.2.6 Outras características que impactam no desempenho

Além de *background* familiar, recursos escolares e efeito escola, outras características foram encontradas nos estudos empíricos que afetam o desempenho acadêmico dos estudantes. Jurado (2013) realiza análise para determinar as variáveis que incidem no desempenho em matemática para a Colômbia, sendo: o tipo de escola, o gosto pela matemática, a zona de localização da escola, o gosto pela escola, além das variáveis de características pessoais e da instituição. Cita como efeito positivo e significativo a presença de computadores nos lares, algumas características da escola e da aula, o fato de a escola ser urbana e privada. Apresenta ainda que o modelo nulo indica uma variância de 42,6% entre as escolas, ao se incluir todos os grupos de

variáveis no nível do aluno e da escola, o resultado se reduz para 32,9%, com considerável diminuição da variância explicada.

Jurado (2016), ao analisar o desempenho em matemática, para 15 países da América Latina, a fim de determinar quais as variáveis são importantes, aponta: sexo, tipo de escola, setor, recursos de consulta, repetência, pertencimento e PIB investido. Do ponto de vista do aluno, repetição de série, mudança de instituição de ensino, comportamento desrespeitoso em sala de aula e origem social são aspectos que trazem desvantagem em relação aos demais. Para as escolas, os efeitos são maiores para aquelas da rede privada pertencentes à zona urbana, sendo que a percepção dos professores quanto à seriedade do problema da falta de motivação do estudante tem efeito negativo no desempenho. No nível país, o PIB investido tem efeito positivo e significativo. Por fim, o resultado, no nível país, apresenta variação de 6,7% e, no nível escola, de 18,6%.

Agüero e Beleche (2013) investigam a alteração no número de dias letivos no México antes de um exame nacional e seu efeito sobre o desempenho dos estudantes. Demonstram que ter mais dias para se preparar para um teste pode estar relacionado a notas médias mais altas dos alunos em matemática e leitura. León e Valdivia (2015) apontam, para o caso do Peru, que deixar de levar em conta as grandes desigualdades geográficas na distribuição de recursos escolares e as restrições enfrentadas pelos alunos pobres leva a uma subestimação do efeito dos recursos escolares de aproximadamente 100%. Indicam ainda que as características da escola e do professor são determinantes no desempenho dos alunos. No entanto, esses atributos tornam-se empiricamente relevantes somente quando explicam as limitações nas escolhas escolares enfrentadas pelos pais.

Beuchert *et al.* (2018), a partir de um movimento de consolidação (fusões) de escolas na Dinamarca, estimam o efeito de curto prazo desse desenho institucional. Os autores ressaltam que o desempenho dos alunos no fechamento das escolas é adversamente afetado no curto prazo, com isso, ao agrupar os resultados em leitura e matemática, encontram um efeito geral de consolidação negativa e estatisticamente significativa ao nível de 5%.

Além disso, quando se observa a construção de indicadores, há alguns pontos a serem observados. Por exemplo, Minaya e Agasisti (2019), ao explorar o uso de indicadores de desempenho escolar, para fins de prestação de contas e escolha das unidades de ensino, estudam a estabilidade de desempenho de estimativas de

escolas primárias italianas, ao longo do tempo, usando dados do Comitê Nacional de Avaliação da Educação. Os autores apontam que não há diferenças substancialmente elevadas nas estimativas de desempenho escolar e que essas são instáveis, com dificuldades de se prever o desempenho futuro com base no desempenho passado e atual.

2.2.7 Método de análise utilizado nos estudos empíricos

A partir dos estudos empíricos apresentados de acordo com a Revisão Sistemática de Literatura (RSL), apêndice A, organizou-se o quadro 1 que descreve os autores, o método utilizado e o país. A maioria dos estudos emprega a regressão multinível como método de análise dos dados.

Quadro 1 - Método de pesquisa utilizado nos estudos empíricos

Autor(es)	Método	País
Agüero e Beleche (2013)	Análise de Regressão	México
Bartholo e Costa (2016)	Regressão Multinível e análise de regressão	Brasil (Rio de Janeiro)
Becker (2011)	Análise de Regressão	Reino Unido
Benevides e Soares (2020)	Análise de Regressão e Método de Pareamento CEM (<i>Coarsened Exact Matching</i>)	Brasil (Ceará)
Bernal, Mittag e Qureshi (2016)	Análise de Regressão	EUA
Beuchert <i>et al.</i> (2018)	Método de Diferença em Diferenças	Dinamarca
Blanco (2011)	Regressão Multinível	México
Brault, Janosz e Archambault (2014)	Regressão Multinível	Canadá (Quebec)
Castelló-Climent e Hidalgo-Cabrillana (2012)	Análise de Regressão	Ampla amostra de países
Choi, Moon e Ridder (2014)	Análise de Regressão	Coreia do Sul (Seul)
Galvão (2018)	Análise de Regressão	Brasil (São Bernardo)

		do Campo)
Griew <i>et al.</i> (2010)	Regressão Multinível	Reino Unido (Bristol)
Hægeland, Raaum e Salvanes (2012)	Análise de Regressão com procedimento de dois estágios	Noruega
Hobin <i>et al.</i> (2010)	Regressão Multinível	Canadá (Ontário)
Humlum e Smith (2015)	Análise de Regressão	Dinamarca
Jurado (2013)	Regressão Multinível	Colômbia
Jurado (2016)	Regressão Multinível	15 países da América Latina
Kim e Rhee (2019)	Análise de Regressão	Quênia
León e Valdivia (2015)	Análise de Regressão com procedimento de dois estágios e probit	Peru
Masci <i>et al.</i> (2017)	Regressão Multinível	Itália
McVicar, Moschion e Ryan (2018)	Análise de Regressão	Austrália (Victoria)
Melo e Moraes (2019)	Análise de Regressão	Brasil
Minaya e Agasisti (2019)	Regressão Multinível	Itália
Oswald e Backes-Gellner (2014)	Análise de Regressão com Estudo Experimental	Suíça
Rocha e Funchal (2019)	Análise de Regressão e Regressão Quantílica	Brasil
Rodríguez e Correa (2019)	Regressão Multinível	Colômbia
Sampaio <i>et al.</i> (2011)	Análise de Regressão com Probit	Brasil
Sansani (2011)	Análise de Regressão	EUA
Seabra, Carvalho e Ávila (2019)	Regressão Multinível	Portugal (Lisboa)
Vardardottir (2015)	Análise de Regressão Quantílica	Suíça
Vinha e Laros (2016)	Revisão Sistemática - Predominância de Regressão Multinível	Brasil, Chile e Argentina

Fonte: Elaborado pelo autor de acordo com os estudos empíricos (2021).

Dos 31 estudos analisados, 38,71% utilizaram regressão multinível conhecida também como modelos hierárquicos; 35,48% empregaram análise de regressão apenas; 22,58% utilizaram análise de regressão combinada com outros métodos; e apenas um estudo aplicou o método de diferença em diferenças, conforme exposto na Tabela 2. Tem-se, portanto, tradicionalmente a análise de regressão só ou combinada com outros modelos com ampla aplicação. Contudo, de forma específica, a análise multinível esteve presente na maioria dos casos.

Tabela 2 - Método de análise identificado na RSL por quantidade e percentual

Método	Quantidade	Percentual %
Regressão Multinível (<i>Hierarchical Linear Model</i> , HLM)	12	38,71%
Análise de Regressão	11	35,48%
Análise de Regressão combinado com outros métodos	7	22,58%
Método de Diferença em Diferenças	1	3,23%
Total	31	100,00%

Fonte: Elaborado pelo autor (2021).

Identificados os principais métodos de análise citados nos estudos empíricos, segue-se para as principais variáveis dependentes e independentes.

2.2.8 Principais variáveis dependentes e independentes dos estudos empíricos

As principais variáveis dependentes e independentes utilizadas nos estudos empíricos foram resumidas no quadro 2, a fim de destacar as mais frequentes. As variáveis dependentes comumente empregadas são as medidas de desempenho do estudante em matemática e linguagem (de acordo com a língua de cada país). As variáveis independentes, por outro lado, obedecem ao objetivo de cada estudo, destacando-se aquelas relacionadas ao estudante e seu nível socioeconômico e à escola.

Quadro 2 - Variáveis dependentes e independentes utilizadas nos estudos empíricos

Autor(es)	Variável Dependente	Variável Independente
Agüero e Beleche	Desempenho em	Dias de instrução (duração do calendário letivo)

(2013)	matemática e leitura	
Bartholo e Costa (2016)	Desempenho em matemática	Prova Rio 1ª avaliação, cor, educação dos pais, indicador de baixa renda (bolsa família), indicador de família biparental, sexo, proporção de pais com baixo nível de escolaridade, proporção de não brancos, proporção de pobres, qualificação dos professores, índice de complexidade escolar, qualificação do diretor
Becker (2011)	Testes de vocabulário para crianças de 3 a 5 anos	Educação dos pais, situação profissional da mãe, horários regulares de dormir, horários regulares das refeições, frequência à pré-escola, interação entre frequência à pré-escola e educação dos pais
Benevides e Soares (2020)	Desempenho em matemática	Variável <i>dummy</i> para escola militar, proficiência em matemática no 5º ano, sexo, raça, idade, índice socioeconômico, infraestrutura adequada, educação da mãe
Bernal, Mittag e Qureshi (2016)	Desempenho em leitura e matemática	<i>Proxies</i> para qualidade (qualidade para professores, certificação profissional avançada, certificação de ensino fundamental, anos de escolaridade, cursos universitários em matemática, cursos universitários em linguagem), <i>proxies</i> de qualidade dos recursos (tamanho médio da classe da primeira série, computadores instrucionais por aluno, equipe administrativa por aluno, proporção de pessoal especializado por aluno)
Beuchert <i>et al.</i> (2018)	Desempenho em leitura	Variável de consolidação da escola (dicotômica), variável que identifica se o aluno foi exposto a uma consolidação escolar (dicotômica) e outras variáveis de características dos alunos

Blanco (2011)	Desempenho em linguagem e matemática	Nível socioeconômico, sexo, condição étnica, trajetória escolar, estrutura do lar e tipo de apoio familiar em relação ao papel da escola, preferências e práticas escolares, percepções de respeito em relação ao clima escolar e à aula, entorno socioeconômico da escola, entorno institucional, estrutura escolar, infraestrutura escolar, experiência e capacitação docente, gestão, clima escolar, participação familiar e práticas docentes
Brault, Janosz e Archambault (2014)	Expectativa dos professores em relação aos estudantes	Composição étnica da escola, composição socioeconômica da escola, composição acadêmica da escola, tamanho da escola, clima educacional, localização da escola, sexo, idade, classe para alunos com necessidades especiais
Castelló-Climent e Hidalgo-Cabrillana (2012)	Parcela da população de 25 anos ou mais com ensino médio e superior como o nível mais alto de escolaridade	Qualidade da educação (testes de proficiência e Ranking acadêmico de universidades mundiais)
Choi, Moon e Ridder (2014)	<i>Korean CSAT score</i> (coreano, matemática, inglês e outros)	Fração de alunos que recebem almoço gratuito ou a preço reduzido, gastos por aluno anualmente, fração de professores, indicador de escola privada, idade da escola em 2008, tamanho da classe para veteranos
Galvão (2018)	Desempenho em matemática	Gasto total por aluno, professores, assistente de classe, coordenação, direção, nível socioeconômico, tempo e disciplina
Griew <i>et al.</i> (2010)	Atividade física escolar	Individual (gênero, merenda escolar gratuita, Índice de massa corporal, etnia, índice de privação múltipla), escolar (índice de privação

		múltipla escolar, percentual de minoria étnica, percentual de sobrepeso, luz do dia)
Hægeland, Raaum e Salvanes (2012)	Desempenho em matemática, norueguês ou inglês	Recursos escolares (gastos, horas de professor), características municipais, <i>background</i> familiar do aluno, receita tributária de energia elétrica
Hobin <i>et al.</i> (2010)	Matrícula em Educação Física	Nível estudante (variáveis demográficas de idade, série, sexo, altura e peso, variáveis comportamentais de tabagismo, de sedentarismo, variáveis psicossociais, incentivo dos pais à atividade física, influência dos amigos na atividade física); e variáveis no nível escola (variáveis ambientais de localização da escola, proporção aluno-professor de educação física, frequência da participação em aulas de educação física, alunos novos e veteranos, renda familiar média da localização da escola)
Humlum e Smith (2015)	Conclusão do ensino médio ou profissionalizante, matrícula no ensino médio ou profissionalizante, inatividade, ganhos aos 30 anos	Tamanho da escola, mudança no tamanho da escola (inauguração), mudança no tamanho da escola (fechamento) e variáveis de controle de nível socioeconômico
Jurado (2013)	Desempenho em matemática	Sexo, posse de computador, deficiência em informática e no estudo, gosto pela escola, gosto por matemática, materiais de instrução, características da aula, características do docente, tipo de escola (privada ou pública), zona (urbana ou rural)
Jurado (2016)	Desempenho em matemática	Sexo do estudante, taxa de reprovação, sentido de pertencimento, posse de recursos para

		estudar, sexo do professor, escolaridade do professor, percepção sobre a falta de motivação dos estudantes de matemática, anos de experiência docente, tipo de escola, zona escolar, serviços básicos da escola, recursos de consulta da escola, sala de aula, percentual do PIB investido em educação
Kim e Rhee (2019)	Taxa de frequência à escola	Sexo, proporção aluno-banheiro, variáveis de controle da escola, variáveis de localização
León e Valdivia (2015)	Desempenho em matemática e linguagem	Sexo, linguagem materna da criança, escolaridade, índice de ativos HH, tipo de escola, percentual de professores com curso superior, computadores por 1000 alunos, presença de biblioteca, proporção sem esgoto, percentual da população sem eletricidade no distrito
Masci <i>et al.</i> (2017)	Desempenho em matemática e leitura	Sexo, geração de imigrante, se o estudante foi matriculado antecipadamente, se o estudante está atrasado, se reside com ambos os pais, se é estudante com irmãos, indicador de <i>status</i> econômico e cultural
McVicar, Moschion e Ryan (2018)	Desempenho na alfabetização (gramática, leitura, ortografia e escrita)	Classificação de novos pares, ocupação dos pais, nível de escolaridade
Melo e Morais (2019)	Média de desempenho no ENEM 2015	Variáveis de clima escolar e contextuais (nível Socioeconômico do estudante)
Minaya e Agasisti (2019)	Desempenho em matemática e linguagem	Tipo de estudante (regular, acima da idade, abaixo da idade para o ano escolar), gênero, <i>status</i> de imigração, índice de <i>status</i> socioeconômico e cultural do estudante
Oswald e Backes-	Média padronizada	Programa de pagamento por desempenho

Gellner (2014)	das notas do primeiro e do segundo ano do programa analisado	(PPP), paciência do aluno, características pessoais, medidas de habilidade, características da empresa
Rocha e Funchal (2019)	Desempenho médio da escola no ENEM (Linguagens e Códigos e Matemática)	Custo direto, número médio de alunos por sala de aula, características do estudante (variáveis de nível socioeconômico)
Rodríguez e Correa (2019)	Desempenho em matemática	Índice socioeconômico do estudante (INSE), nível socioeconômico médio da escola (INSEC), interação entre o INSE e o INSEC (INSEX), taxa de mortalidade infantil, desempenho fiscal e de conflitos (variáveis municipais)
Sampaio <i>et al.</i> (2011)	Variável dicotômica para evasão (probabilidade de evasão no ensino superior)	Renda e educação dos pais, gênero, idade, estado civil, nota de entrada no vestibular, se o aluno estudou em escola pública (dicotômica)
Sansani (2011)	Taxa de mortalidade	Proporção professor-aluno, duração do período escolar, salário relativo do professor
Seabra, Carvalho e Ávila (2019)	Desempenho em matemática	Nível socioeconômico do estudante, composição étnica da escola, tamanho da escola, nacionalidade, sexo e estudantes estrangeiros de língua portuguesa
Vardardottir (2015)	Resultado no PISA para leitura, matemática, ciências e resolução de	Condição socioeconômica, <i>status</i> socioeconômico dos colegas de classe (pares), sexo, idade, alunos nativos ou não, quantidade de livros em casa, quantitativo de livros didáticos em casa, pai desempregado, pais trabalham em

	problemas	tempo integral, língua estrangeira falada em casa, proporção aluno professor. Em resumo: uma medida de qualidade, uma medida de qualidade média dos pares e características individuais, familiares e demográficas
Vinha e Laros (2016)	Não se aplica (revisão sistemática de literatura)	Não se aplica (revisão sistemática de literatura)

Fonte: Elaborado pelo autor de acordo com os estudos empíricos (2021).

A partir da análise do referencial teórico e da revisão dos estudos empíricos, parte-se para a síntese do capítulo e formulação da tese na próxima seção.

2.3 Síntese do capítulo e formulação da tese

A partir dos estudos teóricos e empíricos do capítulo 2, pode-se constatar que o desempenho dos estudantes é afetado por diversas variáveis dentro de vários contextos. A relação entre insumos e desempenho acadêmico por meio de testes de proficiência pode ser dada como uma função de produção a partir da análise de eficiência com foco no desempenho, ideia básica da microeconomia neoclássica com a abordagem da função de produção educacional. Contudo, a ideia de insumo, nesse contexto, não deve estar restrita apenas à instituição escolar. As informações de *background* familiar (COLEMAN, 1966), as habilidades inatas, o efeito dos pares e os recursos aplicados pela escola são importantes para explicar o resultado (ALBERNAZ; FERREIRA; FRANCO, 2002; AMMERMÜLLER; HEIJKE; WÖßMANN, 2005; CURI; SOUZA, 2015).

A dificuldade apresentada, desde o relatório Coleman (1966), quanto ao papel da escola no desempenho está associada ao peso significativo do *background* familiar. Qualquer estudo na busca de determinantes e de seus efeitos precisa levar esse elemento em consideração, contudo, sua importância deve ser observada de forma complementar, o que implica que a escola tem seu papel no resultado acadêmico dos estudantes. Alguns estudos corroboram esse entendimento, ao apontar que não apenas a escolaridade, mas, sobretudo, a qualidade da escola é

importante para explicar, inclusive, o desenvolvimento dos países (HANUSHEK; WOBMANN, 2007). Isso quer dizer que a tecnologia desenvolvida na escola e a maneira como ela utiliza seus recursos, combinando infraestrutura física com ambientes pedagógicos, docentes e equipe de apoio técnico-pedagógico, são fatores que impactam na explicação do desempenho.

As evidências empíricas permitem observar que os recursos escolares são importantes na determinação do desempenho acadêmico tanto de países desenvolvidos como de países em desenvolvimento, como o Brasil. Dentre os fatores significativos, destacam-se tamanho da escola, recursos escolares, recursos financeiros, qualidade da escola, número de dias letivos, relação entre escola e ambiente, gastos e custos escolares, clima, composição social da escola e variáveis do professor (GLEWWE *et al.* 2011; HUMLUM; SMITH, 2015; HÆGELAND; RAAUM; SALVANES, 2012; JURADO, 2013; KIM; RHEE, 2019; OSWALD; BACKES-GELLNER, 2014; BERNAL; MITTAG; QURESHI, 2016; AGÜERO; BELECHE, 2013; BLANCO, 2011; ROCHA; FUNCHAL, 2019; MELO; MORAIS, 2019; BARTHOLO; COSTA, 2016).

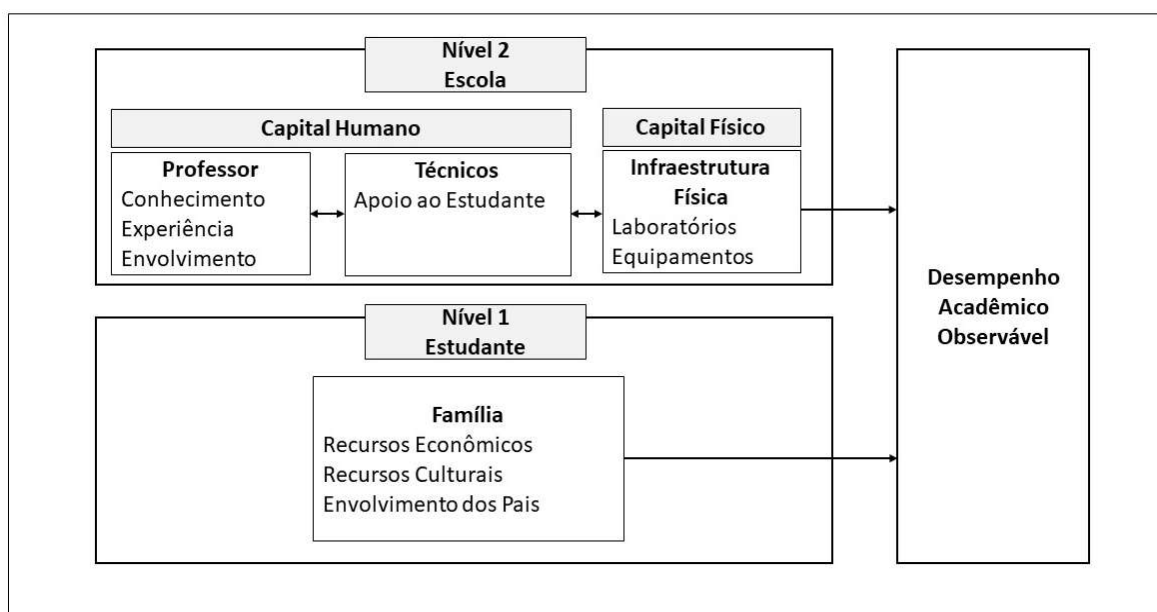
Dada a importância da escola a partir da organização de seus recursos, importa saber se o efeito da escola sobre o desempenho dos estudantes é superior ao esperado para alunos com mesmo nível socioeconômico. O efeito escola, nesse sentido, reforça a importância dos recursos escolares de acordo com a tecnologia desenvolvida, pois, do ponto de vista empírico, os resultados são positivos e significativos, com percentual líquido, controlado o nível socioeconômico, da escola de até 19% (CHOI; MOON; RIDDER, 2014; MASCI *et al.*, 2017; VINHA; LAROS, 2016; LEÓN; VALDIVIA, 2015; BENEVIDES; SOARES, 2020).

Esses efeitos ocorrem porque a escola é o espaço em que docentes, técnicos, recursos materiais e pedagógicos são organizados e geridos com o intuito de se ofertar um ensino que repercuta numa aprendizagem satisfatória. Portanto, percebe-se que os recursos escolares têm importância no desempenho acadêmico dos estudantes (HUMLUM; SMITH, 2015; HÆGELAND; RAAUM; SALVANES, 2012; JURADO, 2013; KIM; RHEE, 2019; OSWALD; BACKES-GELLNER, 2014; BERNAL; MITTAG; QURESHI, 2016; AGÜERO; BELECHE, 2013; BLANCO, 2011; ROCHA; FUNCHAL, 2019; MELO; MORAIS, 2019; BARTHOLO; COSTA, 2016); que esses recursos são relevantes para o efeito escola (CHOI; MOON; RIDDER, 2014; MASCI *et al.*, 2017; VINHA; LAROS, 2016; LEÓN; VALDIVIA, 2015; BENEVIDES; SOARES,

2020); e que a função de produção educacional permite verificar a relação desenvolvida entre recursos escolares e desempenho acadêmico (CURI; SOUZA, 2004; GYIMAH-BREMPPONG; GYAPONG, 1991; FIGLIO, 1999; COOPER; COHN, 1997; PRITCHETT; FILMER, 1999; TODD; WOLPIN, 2003; GLEWWE; KREMER, 2006; GLEWWE *et al.* 2011).

Com base no modelo conceitual de Soares (2007), apresenta-se como tese deste trabalho que o desempenho dos estudantes decorre tanto de fatores individuais como escolares. Nesse nível, há o que se define por função de produção escolar que tem como produto o desempenho dos alunos e como insumos um grupo de itens relacionados ao capital físico e outros ao capital humano. A tese que se propõe é de que ambos são determinantes do desempenho acadêmico, porém, o capital humano é mais relevante para o desempenho dos estudantes, conforme modelo conceitual da figura 4.

Figura 4 - Modelo conceitual da tese



Fonte: Adaptado de Soares (2007, p. 141).

A figura 4 possibilita identificar os dois níveis sob análise que influenciam o desempenho acadêmico. O nível 1 corresponde às características do estudante, o seu nível socioeconômico, com peso significativo no desempenho. O nível 2 representa as características da escola, dividida em capital humano e capital físico, sendo o

capital humano a força propulsora da aprendizagem. Em outras palavras, os recursos escolares são importantes para explicar o desempenho.

3 METODOLOGIA

Este capítulo aborda os procedimentos metodológicos para avaliar a contribuição dos insumos da função de produção educacional para o desempenho acadêmico dos estudantes no Exame Nacional do Ensino Médio (Enem). Dessa forma, inicia-se com a descrição do modelo e das variáveis utilizadas, dos dados da pesquisa, da amostra, da base de dados e do tratamento e das variáveis de controle.

3.1 Análise Envoltória de Dados - DEA

A função de produção normalmente utilizada em estudos para avaliar o desempenho dos estudantes é a análise multinível. No entanto, para avaliar a otimização na utilização dos recursos visando a maximização do desempenho, optou-se pela Análise Envoltória de Dados, conhecida como DEA (*Data Envelopment Analysis*), que leva em consideração essa relação. Essa é a função tradicionalmente utilizada na economia em tais situações.

O método DEA consiste num método não paramétrico que determina o nível de eficiência das unidades de análise através de um processo de otimização (PEÑA, 2008; MOREIRA; BENEDICTO; CARVALHO, 2019). Para a sua utilização, identifica-se primeiro as DMUs (*Decision Making Units*), as quais representam uma unidade mínima homogênea em que há a tomada de decisão (SUGUIY, 2017), que, neste caso, foram os *campi* dos Institutos Federais de Educação, Ciência e Tecnologia (IFs)³.

Na aplicação da DEA, pode-se escolher dentre duas opções: usar o modelo orientado a *outputs*, em que se obtém o nível máximo de produto, mantendo-se os insumos fixos, ou o modelo orientado a *inputs* em que se visa o menor uso de *inputs* dada a quantidade dos *outputs* (LINS; ALMEIDA; BARTHOLO JUNIOR, 2004). Nesta pesquisa, empregou-se a DEA orientada a produto e com retornos variáveis.

O modelo é orientado a produtos porque, dada uma estrutura de capital físico e humano em um determinado *campus*, o que se objetiva é ampliar o desempenho dos alunos, não sendo o intento manter o desempenho dos discentes reduzindo recursos. O emprego de retornos variáveis decorre primeiramente porque os ajustes

³ Uma DMU consiste em um Campus, mas estão referenciadas no texto de forma geral como IF, sendo possível ainda o uso de campi e unidade(s) como sinônimas.

entre os insumos e os produtos possuem melhores ajustes com equações polinomiais. Em segundo lugar, porque há diferença entre os resultados das estimações por retornos constantes e por retornos variáveis. Por retornos constantes, obteve-se apenas 104 DMUs eficientes enquanto que, por retornos variáveis, alcançou-se 167, indicando a existência de retornos variáveis em diversas unidades.

A aplicação do modelo BCC para análise de eficiência com dados do Enem vai ao encontro de outros estudos presentes na literatura (FLORES, 2019). Como se pretende maximizar a produção, dados os níveis de insumos (orientação ao produto), a formulação do Modelo BCC é dada da seguinte forma:

$$\text{Minimizar } h_k = \sum_{i=1}^n v_i x_{ki} + v_k \quad (11)$$

Sujeito a:

$$\sum_{r=1}^m u_r y_{rk} = 1 \quad (12)$$

$$\sum_{r=1}^m u_r y_{jr} - \sum_{i=1}^n v_i x_{jr} - v_k \leq 0 \quad (13)$$

$$u_r, v_i \geq 0 \quad (14)$$

Considerando:

$y = \text{outputs}; x = \text{inputs}$

$u, v = \text{pesos}$

$r = 1, \dots, m; \quad i = 1, \dots, n; e$

$j = 1, \dots, N$

Além disso, foi empregado também o modelo DEA com análise espacial dos dados, com a utilização de insumos no nível dos estados para caracterizar a condição socioeconômica das regiões. Dadas as definições teóricas do método DEA, o próximo tópico aborda o desenvolvimento das variáveis usadas no modelo.

3.2 Variáveis usadas no modelo DEA

Para constituição da Análise Envoltória de Dados (DEA) foram utilizados como insumos variáveis relacionadas à instituição (nível escola), aos discentes (nível individual) e aos estados (nível regional). No que se refere à escola, foram avaliados os insumos relacionados ao capital humano e ao capital físico. Para o capital humano, constituíram-se dois indicadores. Um para avaliar o efeito da titulação (indicador de

titulação do corpo docente - ITCD) e outro para avaliar a contribuição da experiência (indicador de experiência do corpo docente - IECD), sendo o de titulação dado por:

$$ITCD = \frac{G*1+A*2+E*3+M*4+D*5}{G+A+E+M+D} \quad (15)$$

Em que (*G*) identifica a qualificação docente com graduação com peso 1; (*A*) é o aperfeiçoamento com peso 2; (*E*) especialização com peso 3; (*M*) mestrado com peso 4; e (*D*) Doutorado com peso 5. O indicador representa o somatório de todos os docentes efetivos ou temporários da unidade em análise ponderado pela titulação dividido pelo somatório de todos os docentes da unidade em análise, independente de titulação e regime de trabalho (MEC, 2016). Logo, os valores variam de 1 a 5, os mais elevados indicam maior titulação dos professores.

O Indicador de Experiência do Corpo Docente (IECD) foi constituído no mesmo sentido do anterior, com a diferença de que os pesos são equivalentes à posição do servidor no plano de carreira, que representa aproximadamente a quantidade de anos na educação profissional e tecnológica. Logo, para avaliar o efeito da experiência, utilizou-se o seguinte indicador:

$$IECD = \frac{\sum_{i=1}^n p_i x_i}{\sum_{i=1}^n x_i} \quad (16)$$

Em que: p_i peso atribuído a classe e ao nível de carreira docente conforme tempo constante no plano de carreira representado em anos; e x_i quantidade de docentes na classe e nível de carreira identificado.

O IECD foi constituído de acordo com a progressão na Carreira de Magistério do Ensino Básico, Técnico e Tecnológico com base nos critérios gerais estabelecidos na Lei nº 12.772/2012. As progressões estão organizadas da seguinte forma: (i) Classe DI com dois níveis com interstício de 24 meses entre os níveis (101 e 102), (ii) classe DII com dois níveis com interstício de 24 meses entre os níveis (201 e 202), classe DIII com quatro níveis com interstício de 24 meses (301, 302, 303 e 304), (iii) classe DIV com quatro níveis com interstício de 24 meses (401, 402, 403 e 404) e classe de professor titular.

Com isso, quem está no início de carreira, por exemplo, entrará na classe DI nível 101 e só mudará para o próximo nível (DI nível 102) após o tempo decorrido de 24 meses. Nesse sentido, p_i é o respectivo peso dados a classe e o nível de carreira em anos, ou seja, para a classe DI nível 1 (DI 101), o peso equivale a 2, que são os

24 meses convertidos em anos, para o nível D I nível 2 (D 102), o peso equivale a 4 e assim sucessivamente; *xi* por sua vez representa a quantidade de docentes na classe e nível respectivamente. Nesse sentido, quanto maior o indicador, maior a experiência.

Quanto à qualidade da infraestrutura, foram utilizados quatro indicadores, conforme o quadro 3, adaptado de Vasconcelos *et al.* (2020).

Quadro 3 - Variáveis de recursos escolares agrupados por dimensão

Indicador	Variáveis
I - Infraestrutura Geral (INFRA_G)	1 - Existência de auditório 2 – Existência de biblioteca 3 – Existência de refeitório 4 – Existência de sala de professores 5 – Existência de banheiro ou vestiário com chuveiro
II - Infraestrutura de Aprendizagem (INFRA_AP)	1 – Existência de acesso à internet aos estudantes 2 – Existência de laboratório de informática 3 – Existência de laboratório de ciências 4 – Existência de quadra de esportes coberta 5 – Existência de piscina 6 – Existência de sala/ateliê de artes 7 – Existência de sala de música/coral 8 – Existência de sala/estúdio de dança 9 – Existência de sala multiuso (música, dança e artes) 10 – Existência de sala de leitura
III – Infraestrutura Audiovisual (INFRA_AUD)	1 - Quantidade de computadores de mesa em uso pelos estudantes 2 - Quantidade de computadores portáteis em uso pelos estudantes 3 - Quantidade de tablets em uso pelos estudantes 4 - Quantidade de aparelhos de som 5 - Quantidade de aparelhos de televisão 6 - Quantidade de Lousas digitais
IV – Infraestrutura Apoio Profissional ao Estudante (INFRA_PR)	1 – Quantidade de Bibliotecário(a), auxiliar de biblioteca ou monitor(a) da sala de leitura 2 – Quantidade de Bombeiro(a) brigadista, profissionais de assistência à saúde (urgência e emergência), enfermeiro(a), técnico(a) de enfermagem e socorrista 3 – Quantidade de Nutricionista 4 – Quantidade de Psicólogo(a) escolar 5 – Quantidade de Profissionais de apoio e supervisão pedagógica: pedagogo(a), coordenador(a) pedagógico(a), orientador(a) educacional, supervisor(a) escolar e coordenador(a) de área de ensino;

	6 – Quantidade de profissionais que atuam na escola - Técnicos(as), monitores(as) ou auxiliares de laboratório(s)
--	---

Fonte: Elaborado pelo autor de acordo com as variáveis do Censo Escolar (2019).

Os indicadores de infraestrutura geral (INFRA_G) e infraestrutura de aprendizagem (INFRA_AP) foram constituídos de acordo com a disponibilidade de recursos, sendo atribuído o valor 1 para a existência do atributo na escola e 0 para a sua ausência, representado pela equação conforme cada indicador:

$$INFRA_G \text{ ou } INFRA_AP = \frac{\sum_{i=1}^n E_{pi}}{\sum_{i=1}^n E_{maxi}} \quad (17)$$

Em que: (i) $INFRA_G$ = indicador de infraestrutura geral referente à p -ésima escola; (ii) E_{pi} = Escore obtido pela p -ésima escola na i -ésima variável (0 ou 1); (iii) E_{maxi} = Escore máximo da i -ésima variável (1); (iv) $p = 1, \dots, k$, número de escolas; (v) $i = 1, \dots, n$, número de variáveis. O mesmo procedimento foi adotado para o indicador de infraestrutura de aprendizagem.

Os indicadores de infraestrutura audiovisual (INFRA_AUD) e infraestrutura de apoio profissional (INFRA_PR) representam cada um o somatório do quantitativo de itens (quadro 3), dividido pelo total de estudantes, multiplicado pela base de mil estudantes para poder comparar as unidades, representado pela equação a seguir conforme cada indicador:

$$INFRA_AUD \text{ ou } INFRA_PR = \left(\frac{\sum_{i=1}^n QE_{pi}}{TE_p} \right) 1.000 \quad (18)$$

Sendo, (i) $INFRA_AUD$ = indicador de infraestrutura audiovisual por mil estudantes; (ii) QE_{pi} = Quantidade de equipamentos audiovisuais obtida pela p -ésima escola na i -ésima variável; (iii) TE_p = Total de estudantes na p -ésima escola. O mesmo procedimento foi adotado para o indicador de infraestrutura de apoio profissional (INFRA_PR).

As características individuais são elementos significativos para explicar o desempenho, como visto nas pesquisas de Hanushek (1989) e Albernaz, Ferreira e Franco (2002), tendo sido consideradas, neste estudo, e inseridas em termos de valor médio por escola, as seguintes variáveis: renda, escolaridade da mãe, número de computador por família, internet disponível às famílias e, por fim, o indicador de número de pessoas por família, avaliado apenas nas estatísticas descritivas, não sendo avaliado no modelo DEA.

Além disso, utilizou-se, no modelo DEA, variáveis no nível dos estados, sendo o Índice de Desenvolvimento da Educação Básica (Ideb) dos anos finais do ensino fundamental da rede pública, privada e estadual de ensino do ano de 2019; o percentual da população urbana residente em domicílios ligados à rede de esgotamento sanitário do ano de 2017; e o rendimento médio per capita do ano de 2019. De forma resumida, tem-se as seguintes variáveis utilizadas como insumos e produtos no modelo DEA, conforme quadro 4.

Quadro 4 - Variáveis utilizadas no modelo DEA

Variáveis	Tipo de Variável	Nível	Base de Dados
Proficiência em Linguagem, Matemática, Ciências Humanas, Ciências da Natureza e Redação	Produto	Estudante	ENEM
Renda familiar	Insumo	Estudante	ENEM
Escolaridade da mãe	Insumo	Estudante	ENEM
Número de Computador	Insumo	Estudante	ENEM
Disponibilidade de internet	Insumo	Estudante	ENEM
Indicador de Titulação do Corpo Docente	Insumo	Escola	Nilo Peçanha
Indicador de Experiência do Corpo Docente	Insumo	Escola	Nilo Peçanha
Indicador de Infraestrutura Geral	Insumo	Escola	Censo Escolar
Indicador de Infraestrutura de Aprendizagem	Insumo	Escola	Censo Escolar
Indicador de Infraestrutura Audiovisual	Insumo	Escola	Censo Escolar
Indicador de Infraestrutura de Apoio Profissional ao Estudante	Insumo	Escola	Censo Escolar
Índice de Desenvolvimento da Educação Básica (Ideb) dos anos finais do ensino fundamental da rede pública de ensino.	Insumo	Estado	INEP
Índice de Desenvolvimento da Educação Básica (Ideb) dos anos finais do ensino fundamental da rede privada de ensino.	Insumo	Estado	INEP
Índice de Desenvolvimento da Educação Básica (Ideb) dos anos finais do ensino fundamental da rede estadual de ensino.	Insumo	Estado	INEP
Percentual da população urbana residente em domicílios ligados à rede de esgotamento sanitário.	Insumo	Estado	Atlasbrasil

Rendimento médio per capita.	Insumo	Estado	IBGE (PNAD)
------------------------------	--------	--------	-------------

Fonte: Elaborado pelo autor (2023).

Com isso, o modelo DEA foi aplicado levando em consideração os insumos de capital humano e capital físico como variáveis no âmbito dos Institutos Federais, os quais são variáveis no nível escola e, portanto, passíveis de intervenção. Essas variáveis correspondem a média por Campus que pertence a um município. O modelo leva em consideração ainda as variáveis socioeconômicas no nível do estudante que também correspondem a média por Campus pertencente a um município, representativas do *background* familiar, que não são passíveis de intervenção. Além disso, aplicaram-se os insumos das condições socioeconômicas em âmbito estadual, os quais não são controláveis, mas que representam algumas das condições de vida da população.

3.3 Dados da pesquisa

Esta pesquisa foi organizada a partir de seis bases de dados, as quais contêm as variáveis de interesse no nível aluno, escola e estado. A necessidade de seis bases sustenta-se porque não há dados suficientes em apenas uma plataforma para a execução da pesquisa. Por isso, foram utilizadas as bases de dados do Exame Nacional do Ensino Médio (Enem), Censo da Educação Básica, Plataforma Nilo Peçanha, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), Atlasbrasil e Instituto Brasileiro de Geografia e Estatística (IBGE) com a utilização da Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD).

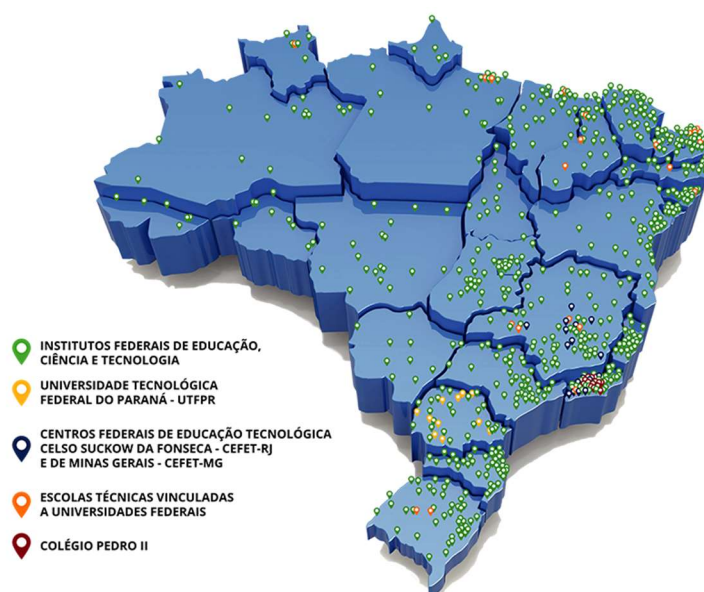
3.3.1 Amostra

O objeto deste estudo corresponde aos *campi* dos Institutos Federais de Educação, Ciência e Tecnologia, em 2019, os quais estão organizados em 38 IFs, com um total de 599 unidades entre *campi* e Centros de Referência em Educação à Distância, dos quais 460 Unidades de Tomada de Decisão (*DMUs*) constituem a amostra desta pesquisa referente apenas aos *campi*. Essas instituições estão distribuídas em todo o país e contemplam 949.831 matrículas em todos os cursos ofertados naquele ano, sendo 234.950 matrículas do ensino técnico de nível médio na

forma integrada (MEC, 2019). Desse total, a amostra contemplou 37.824 que realizaram a prova do Enem, das quais foram extraídas as variáveis relacionadas aos indivíduos e à proficiência.

A distribuição dos *campi* dos IFs, em termos geográficos, pode ser observada na figura 5, legenda em verde, em que é possível identificar que as regiões Sudeste e Nordeste possuem o maior número de unidades. Apesar de a Rede Federal de Educação Profissional e Tecnológica contar com outras instituições, este estudo concentra seus esforços apenas nos Institutos Federais por ser uma política pública com regras uniformes a todas as unidades do país. Esse pressuposto está de acordo com os pré-requisitos do modelo DEA de uniformidade entre as unidades de tomada de decisão.

Figura 5 - Localização dos Institutos Federais de Educação, Ciência e Tecnologia no Brasil – 2019



Fonte: Ministério da Educação (MEC, 2022).

Após a explanação sobre os procedimentos da amostra, as próximas seções abordam as bases de dados que possibilitaram os dados para a construção dos indicadores de insumos e produtos que foram aplicados no modelo DEA.

3.3.2 Base de dados

Utilizou-se o Exame Nacional do Ensino Médio (Enem) que foi instituído em 1998 com a finalidade de avaliar o desempenho dos estudantes ao final da educação

básica. A partir de 2009, o exame começou a ser utilizado como mecanismo de acesso à educação superior, além disso, os resultados podem ser utilizados para pesquisas e indicadores educacionais (INEP, 2021). Quanto à avaliação, o exame contém 180 questões objetivas, divididas em quatro áreas do conhecimento: (i) linguagens, códigos e suas tecnologias; (ii) ciências humanas e suas tecnologias; (iii) ciências da natureza e suas tecnologias; (iv) e matemática e suas tecnologias, além do desenvolvimento de uma redação a partir de uma situação-problema.

O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) disponibiliza os microdados do exame que é o menor nível de desagregação de dados. Os microdados contêm as notas das provas e do questionário socioeconômico dos estudantes inscritos no Enem. Essa base está organizada da seguinte forma: (i) dados do participante; (ii) dados da escola; (iii) dados do pedido de atendimento especializado; (iv) dados do pedido de atendimento específico; dados dos pedidos de recursos especializados e específicos para a realização das provas; (v) dados do local e aplicação da prova; (vi) dados da prova objetiva; dados da redação; e (vii) dados do questionário socioeconômico. O interesse nessa base de dados recai sobre as medidas de desempenho e o questionário socioeconômico.

Dessa forma, quanto às medidas de desempenho, foi utilizada a nota da prova de linguagens e códigos, nota da prova de matemática, nota da prova de ciências humanas, nota da prova de ciências da natureza e nota da prova de redação. Quanto ao questionário socioeconômico, foram empregadas a variável renda, a variável escolaridade da mãe ou responsável, número de computador por família e disponibilidade de internet pelas famílias. Além disso, a variável número de pessoas por família também foi utilizada, contudo, apenas na apresentação dos dados descritivos.

Utilizou-se também o Censo Escolar, que segundo o INEP (2021), é o principal instrumento de coleta de dados da educação básica brasileira, além de fonte significativa quanto às estatísticas educacionais no país. A pesquisa conta com a participação de todas as escolas públicas e privadas do país, abrangendo as diferentes etapas e modalidades da educação básica e profissional, sendo: (i) Ensino regular (educação infantil, ensino fundamental e médio); (ii) Educação especial (escolas e classes especiais); (iii) Educação de Jovens e Adultos (EJA); Educação profissional (cursos técnicos e cursos de formação inicial continuada ou qualificação profissional). No processo de coleta de dados, primeiramente, buscaram-se os dados

sobre os estabelecimentos de ensino, gestores, turmas, alunos e profissionais escolares em sala de aula. Após essa etapa, coletam-se os dados sobre a situação do aluno, movimento e rendimento escolar dos estudantes ao final do ano letivo.

Os microdados do Censo Escolar estão organizados em cinco grupos de dados: (i) dados da escola; (ii) dados do aluno; (iii) dados da turma; (iv) dados do profissional escolar em sala de aula; e (v) dados do gestor escolar. O interesse recai no grupo de dados da escola em que se encontram as variáveis relacionadas à infraestrutura das unidades de ensino. No questionário para os dados das escolas, as principais variáveis de interesse são aquelas já elencadas no quadro 3, as quais se encontram nos tópicos: (i) “Caracterização e Infraestrutura”; (ii) “Equipamentos”; (iii) “Recursos Humanos”; e (iv) “Organização Escolar”.

Para os dados dos Institutos Federais, utilizou-se a Plataforma Nilo Peçanha (PNP), que, segundo o Ministério da Educação (MEC, 2020), é um ambiente virtual de coleta, validação e disseminação das estatísticas oficiais da Rede Federal de Educação Profissional, Científica e Tecnológica (Rede Federal). A PNP tem como objetivo reunir e organizar dados para informações a respeito dos estudantes, servidores docentes, servidores técnicos-administrativos e de gastos das unidades. Esse conjunto de dados possibilita o cálculo de indicadores de gestão para acompanhamento da Secretaria de Educação Profissional e Tecnológica (SETEC) vinculada ao MEC. Além disso, os dados disponíveis podem ser utilizados para a elaboração de estudos e pesquisas, como é o caso desta tese.

Os arquivos com os microdados da PNP estão organizados em formato CSV, disponibilizados da seguinte forma: (i) Microdados Matrículas; (ii) Microdados Eficiência Acadêmica; (iii) Microdados Servidores; e (iv) Microdados Financeiros. Foram utilizados os microdados dos servidores, especificamente, os relacionados aos docentes, conforme o quadro 5.

Quadro 5 - Microdados servidores

Coluna	Descrição
Vínculo com a administração pública	Apresenta o vínculo funcional com a Administração Pública, podendo ser Efetivo ou Substituto/Temporário.
Vínculo da Carreira	Aponta a qual carreira o servidor pertence: EBTT, PCCTAE ou PROF. OUTROS.

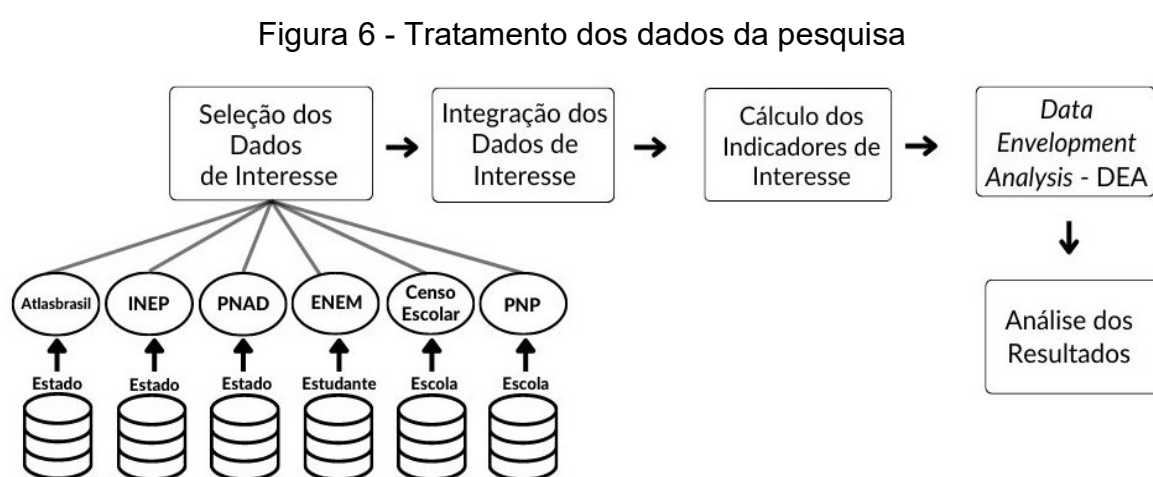
Nível da Carreira	Níveis funcionais do servidor. As opções dependem do vínculo Carreira, conforme tabela a seguir:	
	Vínculo Carreira	Nível da Carreira
	EBTT	'D'
PROF. OUTROS		'4'
		'5'
		'6'
		'7'
		'8'
PCCTAE		'-'
		'A'
		'B'
		'C'
		'D'
Jornada de Trabalho	Indica a jornada de trabalho cumprida pelo servidor, incluindo o regime de Dedicção Exclusiva (DE).	
RSC	Indica se o servidor possui RSC (Reconhecimentos dos Saberes e Competências) e o respectivo nível, podendo ser: RSC-I, RSC-II e RSC-III.	
Titulação	Apresenta a maior escolaridade do servidor, podendo ser: Ensino Fundamental, Ensino Médio, Técnico, Graduação, Aperfeiçoamento, Especialização, Mestrado e Doutorado.	
Código da Unidade de Ensino – SISTEC	Código gerado no SISTEC que identifica de maneira única cada unidade de ensino. As reitorias, as direções-gerais e os polos de inovação não possuem código por não serem cadastrados no SISTEC. Refere-se ao código da unidade de ensino a qual o servidor está lotado.	
Unidade de Lotação	Nome da unidade a qual o servidor está lotado	
Instituição	Sigla da Instituição na PNP. Cada escola técnica vinculada às Universidades Federais foi compreendida como uma Instituição.	

Fonte: Plataforma Nilo Peçanha (2020).

As bases de dados do Enem, da Plataforma Nilo Peçanha e do Censo da Educação Básica foram as principais fontes utilizadas nesta pesquisa pelo volume e complexidade dos dados. Além dessas três principais bases, para os dados no nível dos estados, foram utilizados o banco de dados do INEP relacionado ao Índice de Desenvolvimento da Educação Básica (Ideb), o AtlasBrasil e a Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE) para variáveis de condições socioeconômicas.

3.3.3 Tratamento dos dados

Após a identificação das bases de dados e das variáveis do modelo, faz-se necessário a explanação sobre o processo de tratamento de dados. Sendo assim, definiram-se as seguintes fases, de acordo com a figura 6: (i) seleção dos dados de interesse; (ii) integração dos dados de interesse com a especificação das variáveis; (iii) cálculo dos indicadores; (iv) aplicação do modelo DEA; e (v) análise dos resultados.



Fonte: Elaborado pelo autor (2021).

Como procedimento para organização da base de dados, geração das informações necessárias para a realização da modelagem DEA e integração do questionário do Enem, Censo da Educação Básica e Plataforma Nilo Peçanha e demais bases foi empregado o software *Knime*. Com essa ferramenta, foi possível realizar os procedimentos de filtragem, avaliação da existência de *outliers* e de informações faltantes, resultando na amostra de 460 unidades.

Para facilitar a comparação dos dados, utilizou-se do procedimento de normalização com z score, com média igual a zero e desvio padrão de um, o que permitiu a comparabilidade e a redução do efeito escala entre as variáveis. Além disso, foi adicionada uma constante para todas as variáveis para se evitar valores negativos.

3.4 Variáveis de controle

Quanto às variáveis de controle, a grande maioria dos artigos pesquisados, a partir da revisão sistemática de literatura, aponta que o nível socioeconômico do estudante (NSE) é amplamente utilizado, conforme o quadro 6. Além disso, variáveis de educação do pai e da mãe e renda também são citadas na pesquisa. O NSE é assim tido como um dos fatores essenciais na explicação das notas em avaliações externas (GALVÃO, 2018).

Quadro 6 - Variáveis de controle por autor conforme Revisão Sistemática de Literatura

Tipo de variável de controle	Autor(es)
Educação dos pais, indicador de baixa renda	Bartholo e Costa (2016)
Educação dos pais, situação profissional da mãe	Becker (2011)
Renda e educação dos pais	Sampaio <i>et al.</i> (2011)
Índice socioeconômico, educação da mãe	Benevides e Soares (2020)
<i>Background</i> familiar	Hægeland, Raaum e Salvanes (2012)
Nível socioeconômico	Blanco (2011)
	Galvão (2018)
	Humlum e Smith (2015)
	Masci <i>et al.</i> (2017)
	Melo e Morais (2019)
	Minaya e Agasisti (2019)
	Rocha e Funchal (2019)
	Rodríguez e Correa (2019)
	Seabra, Carvalho e Ávila (2019)
	Vardardottir (2015)

Fonte: Elaborador pelo autor (2022).

O NSE tem como componentes principais renda e escolaridade dos pais. Galvão (2018) utilizou em seu trabalho o Indicador de Nível Socioeconômico das Escolas (INSE) que consiste numa síntese de medidas relativas à condição social dos alunos, como, por exemplo, escolaridade dos pais e disponibilidade de determinados bens em suas casas. Para este estudo, optou-se por utilizar as variáveis de renda, escolaridade da mãe, número de computadores por família e acesso à internet para

se ter as características de *background* familiar. A variável número de pessoas na família foi utilizada apenas nas estatísticas descritivas para compreensão da realidade socioeconômica dos estudantes e não foi incluída no modelo DEA.

4 ANÁLISE DOS RESULTADOS

Este capítulo contém as análises realizadas com base na metodologia proposta e está organizado em duas seções, a primeira relacionada às estatísticas descritivas e a segunda à análise de eficiência por meio da Análise Envoltória de Dados (DEA).

4.1 Estatística descritiva

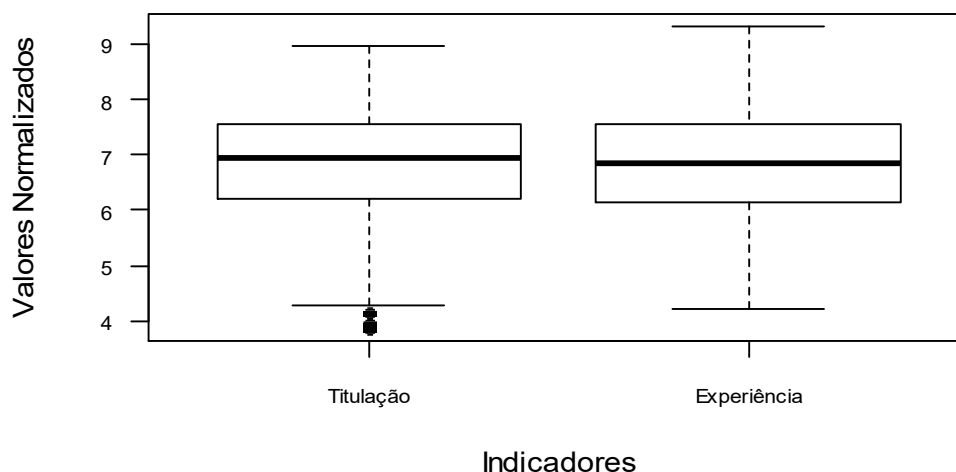
As estatísticas descritivas têm como fim descrever os dados utilizados na análise, permitindo com isso compreender se há transformação de dados ou eliminação de observações e sua compreensão auxilia no entendimento dos fenômenos observados. São analisados os indicadores de capital humano, capital físico, nível socioeconômico, proficiência no Exame Nacional de Ensino Médio (Enem), descrição dos dados por localidade e, por fim, é feita a síntese desses resultados.

4.1.1 Estatística descritiva dos indicadores de Capital Humano e Capital Físico

As estatísticas descritivas dos insumos (*inputs*) permitem evidenciar que os indicadores de capital humano, representados pela titulação e experiência docente possuem distribuição semelhante (gráfico 1). Apesar disso, é possível observar alguns valores extremos (*outliers*) no indicador de titulação docente. Os três pontos identificados são Campus Tarauca (IFAC), indicador de 3,33; Campus Labrea (IFAM) (3,35); e Campus São Gabriel da Cachoeira (IFAM) (3,40). Um dos fatores associados a menores indicadores de titulação para esses IFs⁴ pode estar relacionado à localização, os IFs da região Norte possuem maiores dificuldades para atrair e fixar docentes com maior qualificação, convivendo com alta rotatividade de servidores.

⁴ As análises foram realizadas no nível de cada campus, mas para facilitar a compreensão, ao se referir a essa unidade, será referenciado como IFs.

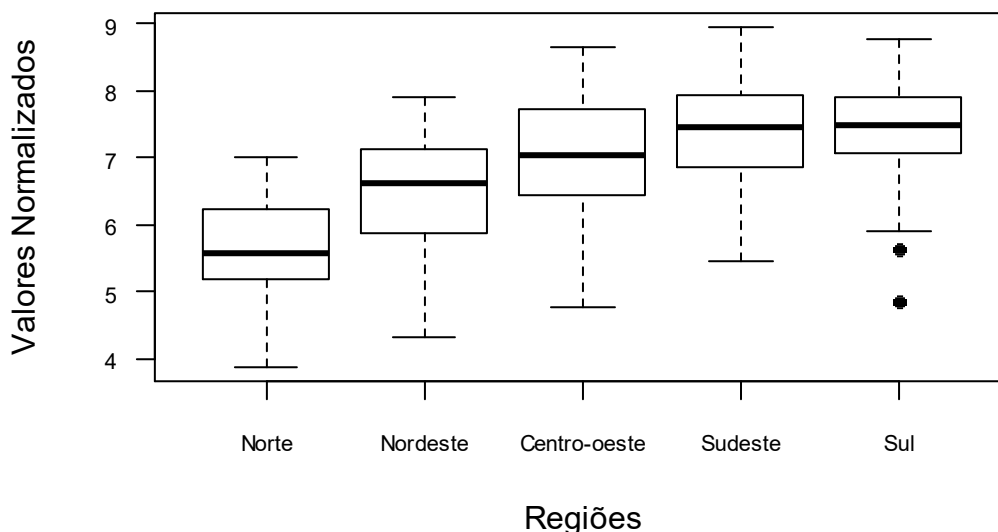
Gráfico 1 - Distribuição dos indicadores de Titulação e Experiência Docente dos IFs (2019).



Fonte: Elaborado pelo autor (2023).

Ao segmentar a distribuição dos indicadores de titulação por regiões, há mais evidências que corroboram as dificuldades de retenção de docentes nas regiões Norte e Nordeste com elevada titulação (gráfico 2). Isso pode ser explicado, em parte, pela maior concorrência nos processos de seleção via concurso público nos principais centros do país em relação às outras regiões. Apenas a região Sul aparece com dois *outliers*, contudo, dentro da distribuição de outras regiões e mesmo assim superiores às regiões Norte e Nordeste. Os valores discrepantes da região Sul pertencem ao Campus de Santana do Livramento do Instituto Federal sul-rio-grandense (IFSul) com indicador de 3,58 e ao Campus Avançado Goioere do Instituto Federal do Paraná (IFPR) com indicador de 3,78.

Gráfico 2 - Distribuição dos indicadores de Titulação Docente dos IFs por regiões do país (2019).

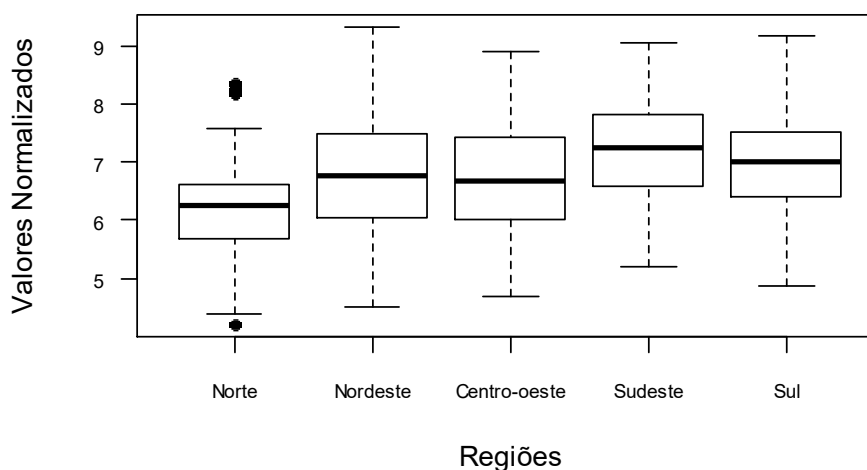


Fonte: Elaborado pelo autor (2023).

A mesma análise para o indicador de experiência docente pode ser verificada no gráfico 3, em que é possível notar uma distribuição próxima entre as regiões do país, com valores similares entre elas, com exceção da região Norte, que apresenta docentes com menor experiência em relação a outras regiões. Essa diferença pode ser explicada pela maior rotatividade de professores nesse território, o que leva a uma menor permanência nessas unidades. Isso pode ser um desafio para a eficiência técnica das instituições de ensino na região Norte, já que a experiência docente é um fator importante na qualidade do ensino.

A região Norte ainda apresenta alguns valores discrepantes, trata-se de três IFs com docentes com experiência acima dos 16 anos na Rede Federal de Educação Profissional e Tecnológica, localizadas nas capitais dos estados, sendo o Campus Boa Vista do Instituto Federal Roraima (IFRR), o Campus Belém do Instituto Federal do Pará (IFPA) e o Campus Manaus Centro do Instituto Federal do Amazonas (IFAM). Com pouca experiência, há apenas o Campus Guajará-mirim do Instituto Federal de Rondônia com média de 2,8 anos de experiência de seus docentes.

Gráfico 3 - Distribuição dos indicadores de Experiência Docente dos IFs por regiões do país (2019)



Fonte: Elaborado pelo autor (2023).

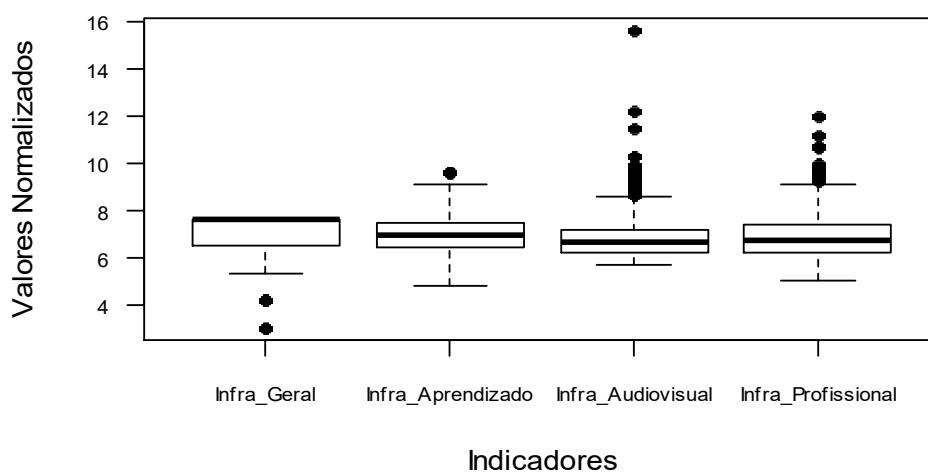
Portanto, ao se avaliar os indicadores de capital humano através das variáveis de titulação e experiência, observa-se que a região Norte apresenta os menores níveis de qualificação dos professores, com uma diferença acentuada em comparação com as demais regiões. A região Nordeste também apresenta baixo nível de titulação em relação às demais, porém, com bom nível de experiência. Isso sugere que a performance dos IFs da região Norte, com base nos indicadores de capital humano, deve apresentar menor desempenho⁵ nas áreas avaliadas pelo Enem em comparação com as demais regiões devido à menor qualificação e experiência docente.

Já os indicadores de capital físico estão divididos em quatro categorias, conforme o gráfico 4. O indicador de infraestrutura geral abrange itens básicos, como biblioteca e auditório. O indicador de infraestrutura de aprendizagem refere-se aos ambientes de ensino, como laboratórios de ciências e acesso à internet para os estudantes. O indicador de infraestrutura audiovisual quantifica o acesso dos alunos a equipamentos de informática e tecnologia. Por fim, o indicador de infraestrutura

⁵ Desempenho aqui é avaliado conjuntamente com os insumos de acordo com a realidade dos IFs em cada região. A partir da análise de eficiência, será possível analisar a relação entre recursos e desempenho, por meio do modelo DEA orientado ao produto (desempenho).

profissional mede a disponibilidade de profissionais de apoio, como psicólogos, técnicos em laboratório e nutricionistas para cada unidade.

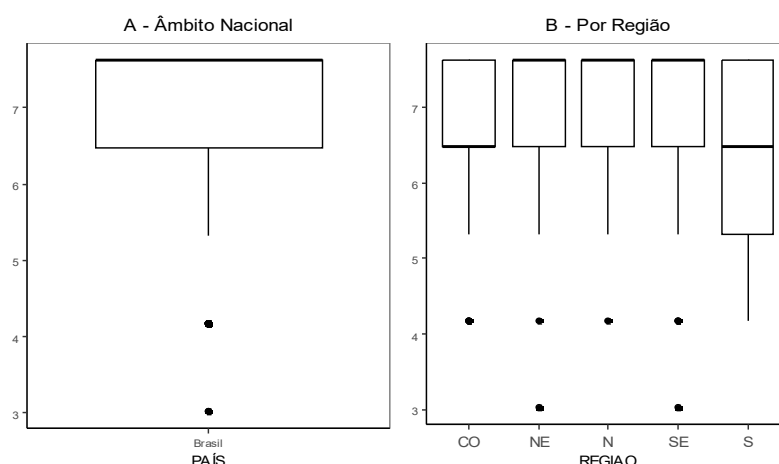
Gráfico 4 - Distribuição dos indicadores de Infraestrutura Física dos IFs (2019).



Fonte: Elaborado pelo autor (2023).

Abrindo os indicadores, verifica-se que o Indicador de Infraestrutura Geral (gráfico 5) apresenta uma distribuição concentrada nos valores mais elevados, com 75% dos dados tendo a mediana no limite superior do boxplot. Isso indica um elevado ativo disponibilizado para os alunos e com pouca variação na infraestrutura básica dos IFs. A maioria dos IFs possui estruturas como auditórios, bibliotecas, refeitórios, salas de professores e banheiros/vestiários com chuveiros. A distribuição entre as regiões do país (gráfico 5) é semelhante, exceto na região Sul. Dois *outliers* chamam a atenção: o primeiro, localizado na região Nordeste (Campus Valência do Piauí/IFPI), possui apenas uma biblioteca; e o segundo, na região Sudeste (Campus Avançado São João da Barra/IFF), possui somente uma sala de professores, de acordo com dados do Censo Escolar de 2019.

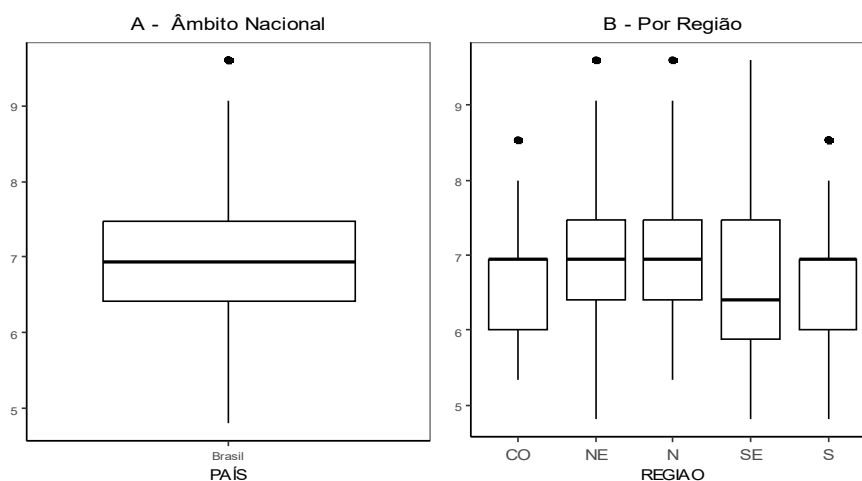
Gráfico 5 - Distribuição do Indicador de Infraestrutura Geral em âmbito nacional e por regiões do país.



Fonte: Elaborado pelo autor (2023).

O Indicador de Infraestrutura de Aprendizagem (gráfico 6) apresenta alguns valores atípicos, porém, esses valores são justificados pela presença de todos os recursos que compõem o indicador, com um ambiente de ensino mais propício para a aprendizagem dos estudantes. Quando avaliados considerando as regiões do país, os dados estão bem dispersos, com exceção para as regiões Norte e Nordeste, de distribuição semelhante. Assemelham-se ainda as regiões Centro-Oeste e Sul, sendo o Sudeste o mais disperso ao mesmo tempo que possui a menor mediana do conjunto de dados.

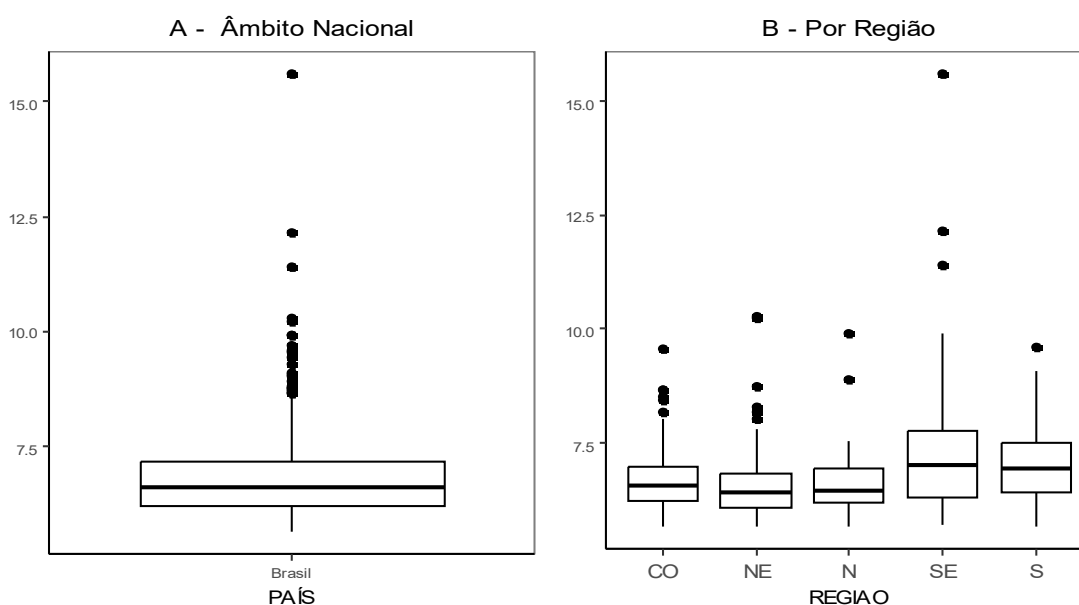
Gráfico 6 - Distribuição do Indicador de Infraestrutura de Aprendizagem dos IFs (2019) em âmbito nacional e por regiões do país.



Fonte: Elaborado pelo autor (2023).

O Indicador de Infraestrutura Audiovisual, por outro lado, apresenta extremos, especialmente nos IFs de Congonhas (IFMG), com indicador de 15,59; Campus Serra (IFES), com indicador de 12,16; e Campus Caraguatatuba (IFSP), com indicador de 11,42. O gráfico 7 permite verificar que esses valores extremos estão mais presentes na região Sudeste, indicando que a quantidade de computadores e equipamentos disponíveis para os estudantes é significativamente maior do que a média dos demais IFs, proporcionando aos estudantes melhores condições para o aprendizado. Em suma, a existência desses *outliers* indica que existem campus com infraestrutura de audiovisual significativamente superior aos demais e, portanto, esses IFs deveriam ter um nível de desempenho superior à média.

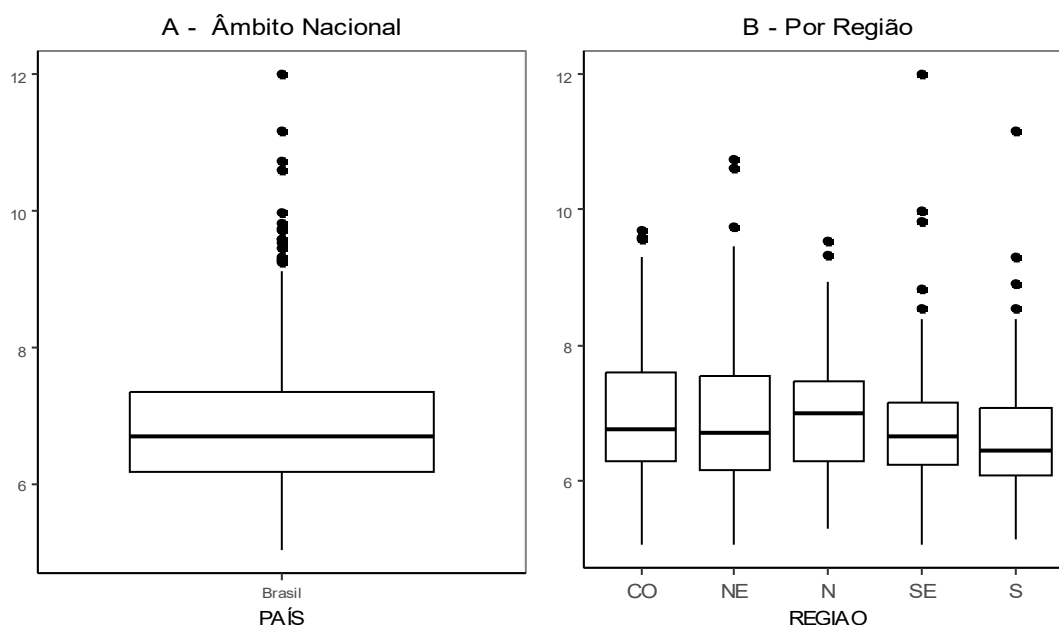
Gráfico 7 - Distribuição do Indicador de Infraestrutura Audiovisual em âmbito nacional e por regiões do país.



O Indicador de Infraestrutura Profissional (gráfico 8) apresenta alguns *outliers*, concentrados principalmente nas regiões Sudeste e Sul. É importante observar que esses *outliers* não estão relacionados à falta de profissionais, mas sim à presença de um número elevado de técnicos que oferecem suporte aos estudantes. Por exemplo, o Campus Caraguatatuba (IFSP), na região sudeste, possui 12 profissionais, incluindo três bibliotecários, um psicólogo, quatro profissionais de segurança e 14 profissionais de pedagogia. Já o Campus de Jaguari (IF Farroupilha), na região Sul, possui 11 profissionais, incluindo dois bibliotecários, um profissional de saúde, um nutricionista,

16 profissionais de pedagogia e um profissional de segurança. É possível notar que ambas as unidades possuem muitos profissionais de pedagogia, indicando priorização das atividades de ensino.

Gráfico 8 - Distribuição do Indicador de Infraestrutura Profissional em âmbito nacional e por regiões do país.



Fonte: Elaborado pelo autor (2023).

Ao se avaliar as correlações entre os insumos inerentes aos IFs, tanto os relacionados ao capital humano quanto os ligados ao capital físico, observa-se que há correlação significativa ($p < 0,01$) entre a titulação dos docentes e a experiência, com um coeficiente de correlação de 0,59, tabela 3. Isso sugere que, com o tempo, os docentes vão completando as suas qualificações a partir das políticas internas dos institutos federais de incentivo à qualificação, como a possibilidade de afastamento. Além disso, há correlação significativa ($p < 0,01$) entre a titulação dos docentes e a infraestrutura audiovisual, com um coeficiente de correlação de 0,20, indicando possivelmente que os docentes mais qualificados estão alocados em IF com melhor oferta de equipamentos audiovisuais (tecnologia).

Verifica-se também correlação positiva entre o indicador de experiência do professor e a infraestrutura geral ($r = 0,18$, $p < 0,01$), a experiência e infraestrutura de aprendizagem ($r = 0,28$, $p < 0,01$), e a experiência com a infraestrutura audiovisual ($r = 0,12$, $p < 0,01$). Isso sugere que os professores mais experientes tendem a proporcionar melhores condições para o desenvolvimento das atividades de ensino.

Contudo, é possível que os professores mais experientes estejam alocados em unidades com infraestrutura mais adequada, devido ao critério de mudança entre os servidores no âmbito de cada Instituto Federal, que é baseado no tempo de serviço.

A análise dos dados revela ainda correlação positiva entre o indicador de infraestrutura geral e o indicador de infraestrutura de aprendizagem ($r = 0,34$, $p < 0,01$), indicando que as unidades com melhores condições de infraestrutura física básica, como auditório e biblioteca, estão correlacionadas com melhores espaços de aprendizagem para os estudantes, como laboratórios. As demais correlações analisadas não foram significantes ao nível de 1%.

Tabela 3 - Correlação de Pearson das variáveis relacionadas ao Capital Humano e ao Capital Físico dos IFs (2019).

Variável	1	2	3	4	5
1. Titulação	—				
2. Experiência	0.593 ***	—			
3. Infraestrutura Geral	0.022	0.180 ***	—		
4. Infraestrutura Aprendizagem	0.071	0.288 ***	0.349 ***	—	
5. Infraestrutura Audiovisual	0.200 ***	0.129 **	0.009	-0.032	—
6. Infraestrutura Profissional	-0.02	0.008	0.057	-0.027	0.233 ***

* $p < .05$, ** $p < .01$, *** $p < .001$

Fonte: Elaborada pelo autor (2023).

As estatísticas descritivas das distribuições dos dados permitem verificar que mesmo as unidades dos IFs, tendo uma política uniforme de pessoal, orçamento e outros recursos com legislação comum, ainda apresentam diferenças significativas por regiões do país, principalmente, na região Norte, com menores níveis de titulação e experiência docente. A região Nordeste possui indicadores de capital humano similar a região Norte, porém, com a experiência nos mesmos níveis das demais regiões do país.

É possível verificar ainda que os indicadores de capital físico, apesar de algumas diferenças, principalmente para o indicador de infraestrutura de aprendizagem, evidenciam condições semelhantes por regiões, não havendo destaque significativo de um indicador para uma região específica. Já as correlações entre os indicadores apontam para a importância da experiência nos insumos de capital físico, assim, possuir docentes em um ambiente estável que possibilite a sua

permanência acaba reverberando para outras áreas relacionadas ao ambiente de aprendizagem.

Essas estatísticas reforçam o papel do capital humano por meio da atuação do professor como fundamental na política pública dos Institutos Federais. Um docente experiente, o que está intrinsecamente ligado com melhor nível de qualificação, além de possibilitar melhor desempenho, contribui para um ambiente de aprendizagem mais rico. Esses achados reforçam a importância do professor, principalmente, no que concerne a que a experiência de ensino tem efeitos positivos no desempenho dos estudantes, ou seja, que a qualidade do docente é importante (ROCKOFF, 2004; RAVKIN; HANUSHEK; KAIN, 2005).

Nesse sentido, a região Norte apresenta um quadro com docentes com menor nível de experiência e titulação, impactando diretamente na qualidade do ensino, pois, como visto, o professor tem papel fundamental não apenas no desempenho, mas também na articulação de um melhor ambiente de aprendizagem. Há, assim, um descompasso na política dos IFs ligado à política de incentivo e permanência dos docentes nessa região, principalmente, para municípios do interior, fora das capitais.

Após a realização da análise descritiva dos indicadores de insumo relacionados ao capital humano e à infraestrutura física dos IFs, a próxima seção trata dos indicadores socioeconômicos.

4.1.2 Estatística descritiva dos indicadores de nível socioeconômico

A avaliação socioeconômica está dividida em cinco indicadores que representam a média por campus: (i) número de pessoas na família; (ii) renda; (iii) escolaridade da mãe; (iv) quantidade de computador por unidade familiar; (v) disponibilidade de internet na residência (gráfico 9). Esses indicadores possibilitam relacionar as unidades com as condições socioeconômicas da região.

O indicador do número de pessoas na família tem relação inversa com o nível socioeconômico da região, ou seja, quanto maior o número de pessoas na família, menor a condição socioeconômica. A medida consta de alguns *outliers* com destaque para duas unidades com valores extremos nas regiões Norte e Nordeste com uma média de sete pessoas por unidade familiar. Apesar disso, entende-se que é possível famílias com sete pessoas, sendo uma realidade no contexto brasileiro.

O indicador de renda é o mais comumente utilizado em estudos com o fim de identificar a realidade social para os diversos fins. Na educação, permite verificar como se dá, por exemplo, a proficiência dada a renda. Na distribuição dos dados, identificou-se um *outlier* apenas, Campus de Campinas (IFSP), com renda bem superior aos demais. Apesar de ser um valor extremo, faz parte da própria realidade nacional, pois o município de Campinas possui um IDHM de 0,816 e PIB per capita bem acima da média brasileira.

O indicador de escolaridade materna é utilizado para compreender o contexto socioeconômico do estudante. A literatura aponta que a maior escolaridade da mãe tende a influenciar nas escolhas quanto à educação dos filhos, especialmente, em situações de litígio conjugal, em que a responsabilidade pela condução da vida familiar é na maioria das vezes da mãe. Ao analisar os dados, observam-se alguns valores extremos, com destaque para o Campus de Santa Maria da Boa Vista (IF Sertão – PE), com o maior valor, 6,5; e o Campus de Viçosa (IFAL) com o menor valor, 3,6. No entanto, não há uma discrepância que justifique a exclusão desses valores na análise.

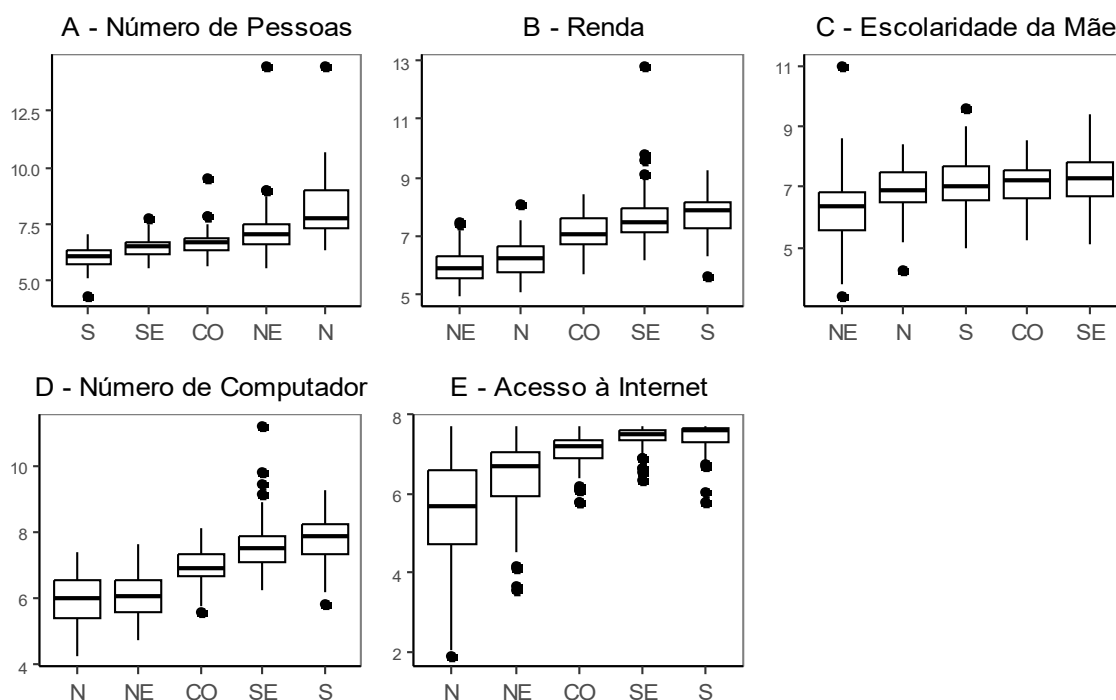
O indicador do número de computadores por família tem relação direta, quanto maior o número, maior o nível socioeconômico da unidade em análise. Dessa maneira, se for vultoso o indicador de computadores por unidade significa que o campus em análise possui em média estudantes com bom nível socioeconômico. Feita a análise dos dados, dois *outliers* foram identificados, o do Campus de Itapetininga (IFSP), com indicador de 2,5; e o do Campus de Campinas, do mesmo Instituto.

O indicador de disponibilidade de internet permite verificar indiretamente a renda familiar e diretamente o acesso a um serviço essencial para o desenvolvimento dos estudos. A distribuição dos dados difere dos demais indicadores, havendo a concentração da disponibilidade para a maioria, contudo, com alguns *outliers* representando baixa disponibilidade desse recurso para algumas unidades. O Campus de Boa Vista Zona Oeste, nesse sentido, apareceu como ausência, o que indica não a falta do serviço no campus, mas que, em média, os seus estudantes não possuem acesso ao serviço de internet em suas residências.

Quando se analisa os cinco indicadores por regiões do país (gráfico 9), a região Norte lidera com os três menores indicadores dos cinco (número de pessoas, número de computador e disponibilidade de internet). O acesso a serviços de internet, nessa região, está com uma distribuição diferente das demais, com boa parte dos estudantes

sem acesso a esse serviço em seus lares. Os três IFs que apresentam os estudantes com as piores condições de acesso é o Campus Boa Vista Zona Oeste (IFRR), o Campus São Gabriel da Cachoeira (IFAM) e o Campus Amajari (IFRR). A região Nordeste segue com o menor nível de renda e escolaridade da mãe. Por outro lado, as regiões Sul e Sudeste continuam com os melhores indicadores.

Gráfico 9 - Indicadores Socioeconômicos por regiões de localização dos IFs (2019).



Fonte: Elaborado pelo autor (2023).

As análises estatísticas das correlações entre os indicadores, apresentadas na tabela 4, permitem verificar a associação com as variáveis socioeconômicas. Os resultados evidenciam que todas as correlações são estatisticamente significativas ao nível de 1%. A menor correlação observada foi entre o número de pessoas por unidade familiar e a escolaridade da mãe (-0,16), indicando uma relação inversa, ou seja, quanto menor a escolaridade da mãe, maior o número de pessoas por unidade familiar.

A correlação mais forte aparece entre a renda e a quantidade de computadores, sendo estatisticamente significativa a 1%, com um coeficiente de correlação de 0,89, indicando uma relação positiva e forte. Isso sugere que as famílias com maior poder aquisitivo possuem melhor infraestrutura doméstica para o estudo dos filhos. Além disso, verificou-se que a variável quantidade de computadores também está

correlacionada com acesso à internet (coeficiente de correlação de 0,75), o que já era esperado, pois esses bens e serviços são complementares e dependentes.

Tabela 4 - Correlação dos Indicadores Socioeconômicos por regiões de localização dos IFs (2019)

Variável	1	2	3	4
1. Número de Pessoas	—			
2. Renda	-0.453 ***	—		
3. Escolaridade da Mãe	-0.164 ***	0.647 ***	—	
4. Quantidade de Computador	-0.534 ***	0.889 ***	0.564 ***	—
5. Acesso à Internet	-0.674 ***	0.662 ***	0.393 ***	0.75 ***

* $p < .05$, ** $p < .01$, *** $p < .001$

Fonte: Elaborada pelo autor (2023).

Assim como visto que a taxa de analfabetismo das pessoas de 15 anos de idade ou mais é superior para as regiões Norte e Nordeste, segundo o IBGE (2020), os dados são categóricos ao apontar que essas mesmas regiões apresentam os menores indicadores socioeconômicos, impactando diretamente no desempenho. Esse impacto ocorre de várias formas, como no poder de escolha dos estudantes ao longo do tempo (SAMPAIO *et al.*, 2011). São discentes cujo ambiente familiar apresenta menores níveis de acesso a bens tecnológicos e culturais, sendo, portanto, um contexto que demanda maior esforço institucional para o alcance de um aprendizado satisfatório.

Levando em consideração os estudos desde o relatório Coleman até as pesquisas mais recentes, tem-se consolidado que o nível socioeconômico impacta de forma significativa o desempenho estudantil (CURI; MENEZES-FILHO, 2013; SOUZA; OLIVEIRA; ANNEGUES, 2018; ASSUNÇÃO; ARAÚJO; ALMEIDA, 2019). Ele consiste em um dos fatores que deve ser levado em consideração em qualquer política pública de educação. Com isso, ao se observar que há uma disparidade entre regiões do país, é conveniente aprimorar os instrumentos que dão suporte à execução de políticas públicas, como o Decreto n.º 7.234, que dispõe sobre o Programa Nacional de Assistência Estudantil (PNAES) (BRASIL, 2010).

Feitas as análises dos indicadores socioeconômicos, a próxima seção trata dos indicadores de proficiência dos estudantes no Exame Nacional do Ensino Médio (Enem).

4.1.3 Estatística descritiva dos indicadores de proficiência no ENEM

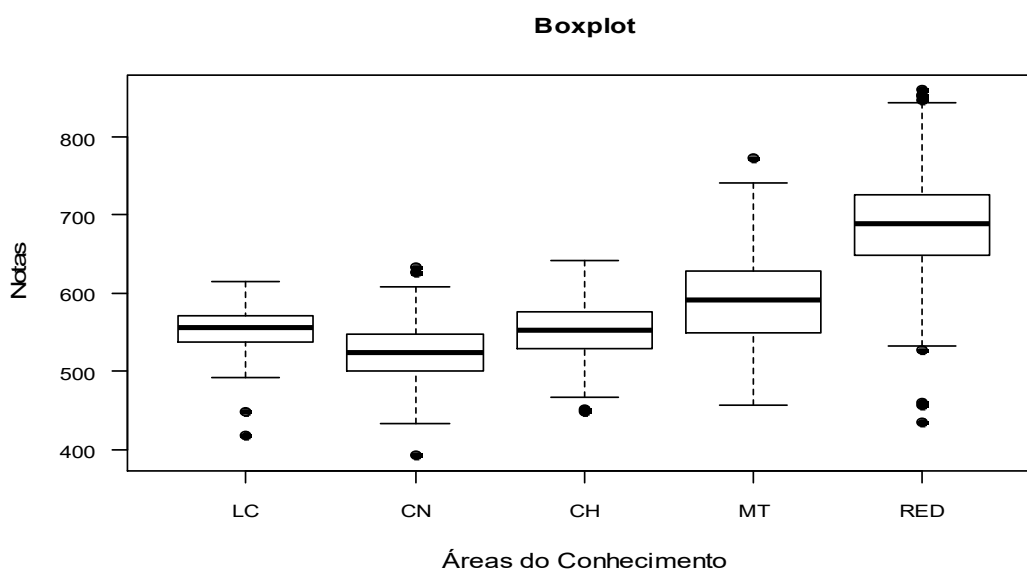
Os resultados do Exame Nacional do Ensino Médio (Enem) foram utilizados como indicadores de desempenho para este estudo e, assim, como produto na Análise Envoltória de Dados (DEA). A análise das notas, nas diferentes áreas avaliadas, revelou a presença de alguns *outliers* concentrados mais na área de redação. A área de Linguagens e Códigos (LC) apresentou uma média geral de 554 pontos, com dois valores discrepantes observados no Campus de Boa Vista Zona Oeste (IFRR) com média de 418; e Campus Amajari (IFRR) com média de 448 (gráfico 10).

A área de Ciências da Natureza (CN), que inclui as disciplinas de Física, Química e Biologia, foi a que apresentou a menor média (524), indicando ser a de maior complexidade para os alunos, inclusive, em relação à Matemática. Um *outlier* foi observado, referente ao Campus de Boa Vista Zona Oeste (IFRR), com média de 393. Por outro lado, *outliers* referentes a notas mais elevadas foram observados nos IFs de Vitória (IFES) (629) e Vila Velha (IFES) (626).

A área de Ciências Humanas (CH) ficou na média entre as demais áreas com 552 pontos, com apenas um *outlier* observado, referente ao Campus de Boa Vista Zona Oeste (IFRR) (448). Já a área de Matemática foi identificada com uma média geral de 591 pontos, com um *outlier* observado no Campus de Vitória (IFES), com média de 772, destacando-se entre as unidades avaliadas.

Por fim, as notas de redação diferem-se das demais pela dispersão, cuja menor pontuação está entre as menores notas observadas no Enem e as maiores são significativamente superiores as de matemática, área com a segunda maior nota entre as avaliadas. A média geral foi de 688 pontos. Os menores valores considerados como *outliers* foram observados no Campus de Marabá Rural (IFPA), com média de 435 pontos; Campus Amajari (IFRR) (456); e Campus Boa Vista Zona Oeste (IFRR) (460). Por outro lado, os *outliers* com os maiores valores foram identificados no Campus de Jaru (IFRO) (860); Campus Formiga (IFMG) (853); e Campus Avançado Conselheiro Lafaiete (IFMG) (851). Esses dados podem ser observados no gráfico 10 que segue.

Gráfico 10 - Distribuição das notas no Enem por área do conhecimento dos IFs (2019).

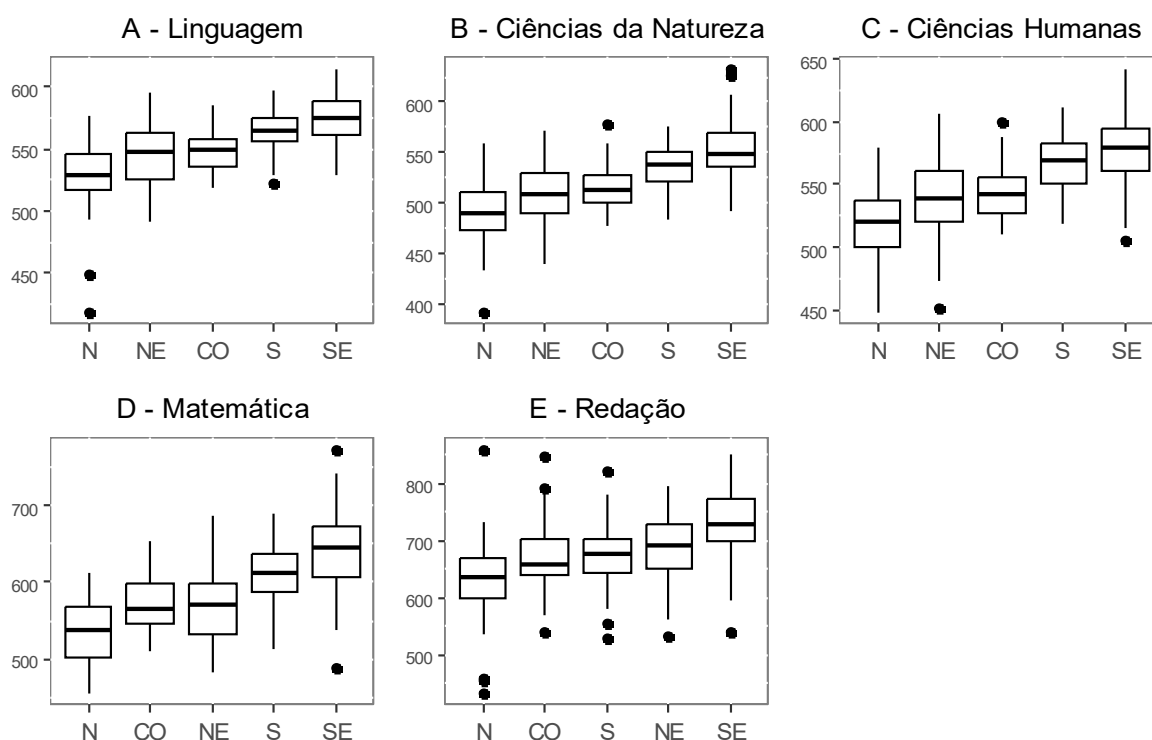


Fonte: Elaborado pelo autor (2023).

A partir da análise das notas por regiões do país, conforme o gráfico 11, tem-se que a região Sudeste apresenta os melhores resultados médios. Essa região é a mais populosa, com maior PIB per capita juntamente com melhores indicadores de vida da população. Além disso, há maior oferta de profissionais das várias áreas demandadas pelos IFs para a carreira docente somada ao melhor nível de renda dos estudantes, propiciando, assim, tendência no acesso a bens básicos, como serviços de internet, bens tecnológicos e socioculturais, como museus, teatros, cinemas e esportes.

A análise dos resultados revela que a região Nordeste evidencia desempenho superior em relação às regiões Norte e Centro-Oeste, em Matemática, e no que diz respeito às regiões Norte, Centro-Oeste e Sul, no que tange à redação. Apesar de apresentar a menor renda e escolaridade da mãe entre as regiões, esses resultados indicam um desempenho significativo. Por outro lado, a região Norte evidencia as menores médias em todas as áreas, o que pode ser atribuído às dificuldades estruturais da população e à dificuldade de atrair e manter profissionais qualificados.

Gráfico 11 - Distribuição da média de notas dos alunos por IF, por área de conhecimento e regiões do país na prova de 2019.



Fonte: Elaborado pelo autor (2023).

Quando se avalia os indicadores de proficiência no Enem considerando os indicadores socioeconômicos, como apontado na tabela 5, há correlação positiva e significativa com todas as áreas ($p < 0,01$). De todos os avaliados, o indicador de quantidade de computador por família apresenta a maior correlação com o desempenho estudantil ($r > 0,7$), enquanto o indicador da nota em redação dispõe da menor correlação ($r = 0,41$). Esses resultados sugerem que o desempenho acadêmico está relacionado à condição socioeconômica, sendo que os estudantes com famílias mais estruturadas, com renda elevada, acesso à internet, maior escolaridade da mãe e menor número de pessoas na residência tendem a apresentar melhor desempenho no Enem.

A correlação com a redação sugere um impacto menor da condição socioeconômica dos estudantes nessa área do conhecimento. Por exemplo, enquanto a correlação entre a quantidade de computadores por família e a proficiência em linguagem é de ($r = 0,75$), a correlação entre redação e essa mesma variável é de apenas ($r = 0,41$), evidência observada para as demais variáveis socioeconômicas. Esses resultados sugerem que, embora a condição socioeconômica ainda tenha um

efeito no aprendizado em redação, esse impacto é menor do que o observado em outras áreas, o que pode ser explorado pelos Institutos Federais como uma oportunidade de melhoria no curto prazo.

Verifica-se também que a titulação e a experiência dos professores apresentam correlação positiva e significativa com a proficiência dos estudantes nas áreas avaliadas, com nível de significância menor que 1%. A titulação apresenta uma correlação superior à experiência. A análise das notas por área, conforme ilustrado no gráfico 11, revela que a área de ciências da natureza apresenta as menores notas entre os IFs, indicando ser esse o componente de maior dificuldade para os estudantes. Contudo, observa-se que essa área apresenta maior correlação com a qualificação docente (tabela 5). Isso pode estar relacionado ao fato de que professores com melhor nível de formação tendem a dominar novas tecnologias e metodologias em suas práticas pedagógicas, o que pode contribuir para o desempenho dos estudantes.

A análise dos indicadores de infraestrutura física revela que o indicador de Infraestrutura Audiovisual apresenta significância estatística ao nível de 1% ($p < 0,01$). Esse indicador refere-se à disponibilidade de equipamentos de informática e/ou tecnologia em uso pelos estudantes. A partir desse resultado, pode-se inferir que os IFs com melhores condições de infraestrutura, que proporcionam acesso a computadores e equipamentos de tecnologia, também apresentam melhores desempenhos.

Apesar de resultados positivos nas correlações dos indicadores, infraestrutura geral exibiu uma correlação negativa (-0,10, $p < 0,05$) com linguagem; e infraestrutura profissional apresentou uma correlação negativa (-0,12, $p < 0,05$) para a mesma área. Essas informações a priori indicam que mais desses recursos não levam a melhor desempenho, o que pode estar associado a unidades que, mesmo com elevados indicadores, não apresentam melhores resultados, ou ainda, por serem recursos de atividade-meio podem tornar mais difícil a captação de alguma associação com o desempenho.

Tabela 5 - Correlação de Pearson das variáveis de proficiência no Enem com as variáveis socioeconômicas, capital humano e infraestrutura física dos IFs (2019).

Variável	LC	CN	CH	MT	RED
1. Renda	0.671 ***	0.651 ***	0.695 ***	0.648 ***	0.347 ***
2. Escolaridade da Mãe	0.467 ***	0.427 ***	0.447 ***	0.441 ***	0.284 ***
3. Quantidade de Computador	0.758 ***	0.731 ***	0.767 ***	0.740 ***	0.412 ***
4. Acesso à Internet	0.733 ***	0.704 ***	0.713 ***	0.669 ***	0.482 ***
5. Número de Pessoas	-0.524 ***	-0.522 ***	-0.523 ***	-0.486 ***	-0.335 ***
6. Titulação	0.430 ***	0.506 ***	0.456 ***	0.447 ***	0.231 ***
7. Experiência	0.370 ***	0.403 ***	0.361 ***	0.401 ***	0.244 ***
8. Infraestrutura Geral	-0.106 *	-0.081	-0.103 *	-0.086	-0.064
9. Infraestrutura Aprendizagem	0.022	0.062	0.023	0.034	0.042
10. Infraestrutura Audiovisual	0.284 ***	0.268 ***	0.267 ***	0.288 ***	0.141 **
11. Infraestrutura Profissional	-0.120 *	-0.097 *	-0.126 **	-0.122 **	-0.096 *

* p < .05, ** p < .01, *** p < .001

Fonte: Elaborada pelo autor (2023).

Quando se correlaciona os indicadores de nível socioeconômico com os indicadores de proficiência no Enem, os resultados são expressivos, pois existe correlação elevada e significativa entre todos os indicadores socioeconômicos, apontando para as evidências já consolidadas na literatura a partir do Relatório Coleman. A região Sudeste, como era esperado, apresenta os melhores resultados médios no Enem, principalmente, por seu contexto socioeconômico ser superior ao das demais regiões do país, em boa parte com melhores indicadores de condições de vida da população.

A região Nordeste, nesse sentido, apresenta resultados que destoam de sua condição socioeconômica, para as áreas de Matemática e Redação, apesar de apresentar a menor renda e escolaridade da mãe entre as regiões. Essa é uma evidência importante sobre a capacidade dos IFs em transformar a realidade na qual estão inseridos, pois de fato há a comprovação de resultados expressivos frente a um contexto desfavorável, corroborando assim a eficácia dos recursos institucionais, a exemplo dos estudos de Gaviria, Martínez-Arias e Castro (2004), para países em desenvolvimento.

Além disso, outro importante estudo, realizado pelos autores Lockheed e Bruns (1990), no contexto brasileiro, apontou que as escolas têm um efeito significativo no desempenho dos estudantes em Matemática e Português, ao se controlar fatores

individuais como o nível socioeconômico, evidenciando que a qualidade do ensino por meio do desempenho dos estudantes está relacionada à gestão escolar, especificamente, à qualificação e motivação dos professores, à utilização dos recursos didáticos adequados e à concomitância de um ambiente de aprendizagem positivo.

Após a análise dos indicadores de proficiência no Enem, a próxima seção aborda a análise estatística dos dados por localidade, se urbana ou rural.

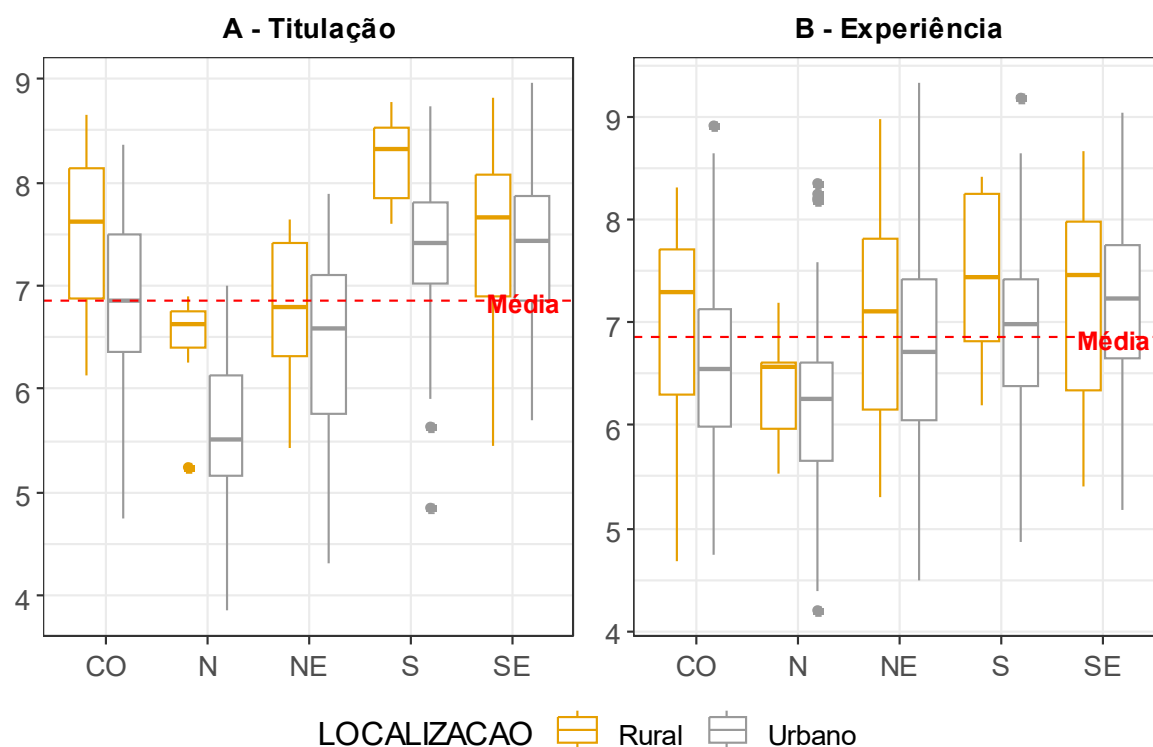
4.1.4 Descrição das variáveis de capital humano e físico por região e localidade

Ao se avaliar os indicadores de titulação e experiência docente por regiões do país e por localidade (gráfico 12), a titulação é superior nas localidades rurais em relação às unidades urbanas. A região Norte evidencia o menor indicador nessa área com a especificidade de que os IFs rurais possuem docentes com titulação bem superior aos IFs urbanos, no entanto, bem abaixo da média nacional. Além disso, há dissimilaridade na distribuição dos dados para essa região, com as unidades urbanas bem distantes umas das outras, enquanto nas unidades rurais os dados estão mais concentrados.

Na região Nordeste, encontra-se o segundo menor indicador com a mediana das unidades rurais superior às urbanas. Já as regiões Sul e Sudeste apresentam os maiores indicadores de titulação, porém, com as unidades rurais superiores às urbanas. O indicador de experiência segue a mesma tendência, com as unidades rurais superiores às urbanas, no entanto, com menor variação. A região Norte mantém a menor posição, porém, com a distribuição dos dados de forma não assimétrica entre as localidades.

Esses resultados evidenciam, assim, que existe diferenciação não apenas entre as regiões do país, mas também entre localidades urbanas e rurais. Esses achados podem estar associados ao fato de que, antes de 2008, data de criação dos Institutos Federais, por meio da Lei 11.892, de 29 de dezembro de 2008, existiam as Escolas Agrotécnicas Federais (EAFA), unidades agrícolas já consolidadas no território brasileiro que foram integradas aos IFs. Com isso, as unidades mais antigas podem estar contribuindo para o aumento da média dos indicadores de titulação e experiência.

Gráfico 12 - Distribuição dos indicadores de titulação e experiência docente por região e localidade (2019).

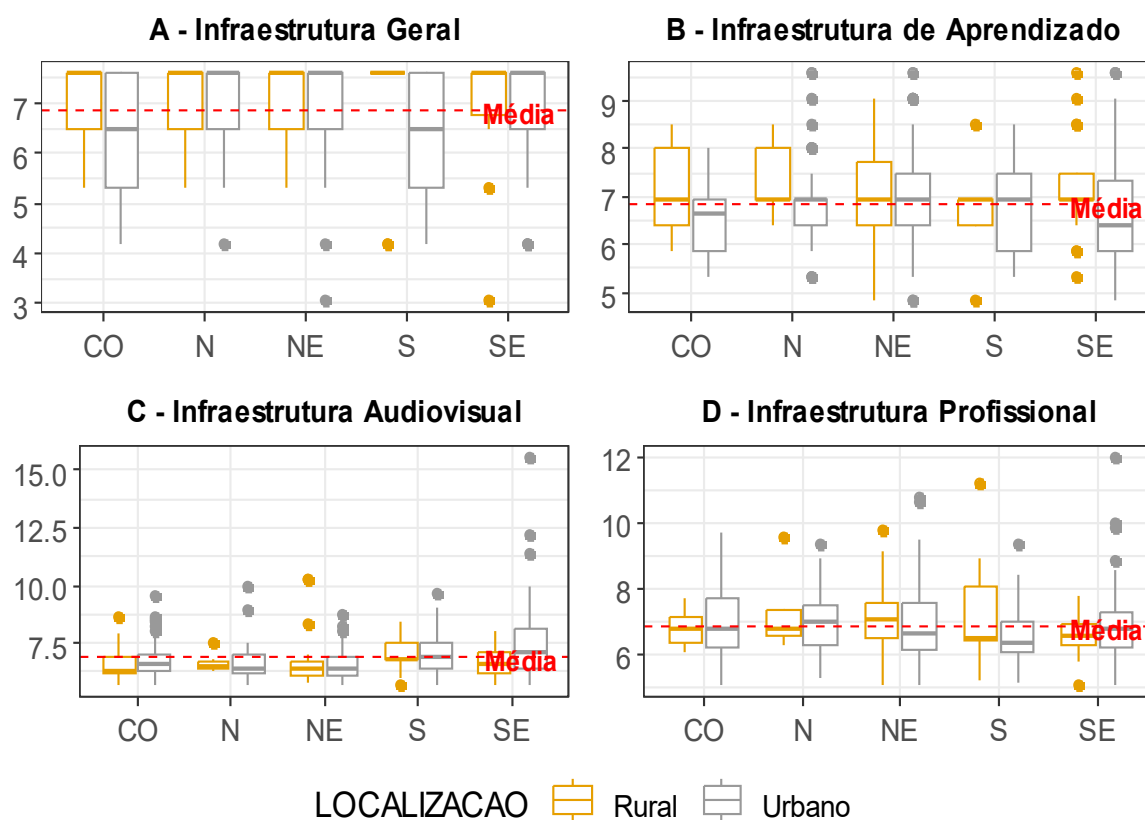


Fonte: Elaborado pelo autor (2023).

Em relação aos indicadores de capital físico (gráfico 13), há diferença mais acentuada entre as localidades para o indicador de infraestrutura de aprendizado com as unidades rurais superiores às urbanas. Nos demais indicadores, as diferenças são menores, o indicador de infraestrutura geral segue essa tendência, com exceção para as regiões Centro-Oeste e Sudeste com as unidades urbanas com indicadores menores. Situação similar ocorre para o indicador de infraestrutura profissional com diferença mais acentuada para a região Sudeste.

Ao se avaliar os indicadores de capital físico considerando os indicadores de capital humano, em relação às regiões e às localidades, as diferenças daqueles são menores, no entanto, os IFs rurais, mesmo em menor proporção, possuem melhores condições gerais. Essas evidências ajudam a compreender que o fato de alguns dos Institutos Federais estarem sediados na zona rural não guarda relação com menores níveis de capital físico e capital humano. Pelo contrário, tanto a qualificação docente como a infraestrutura estão mais bem distribuídas nas localidades rurais.

Gráfico 13 - Distribuição dos indicadores de capital físico classificados por região e localidade dos IFs (2019).



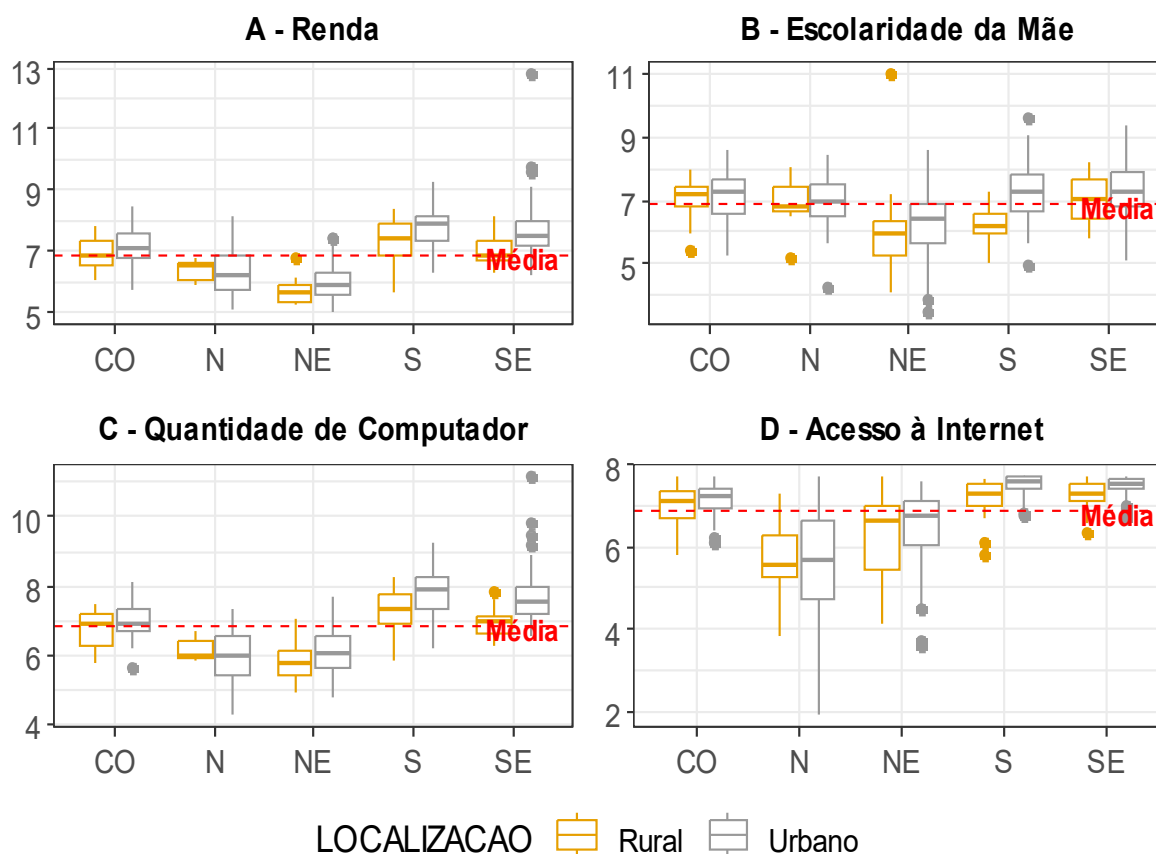
Fonte: Elaborado pelo autor (2023).

Contrariamente aos indicadores de capital humano e capital físico, as localidades rurais apresentam menores níveis socioeconômicos em comparação às localidades urbanas (gráfico 14). A região Nordeste, assim como na avaliação geral, continua com a menor renda seguida da região Norte. O indicador de internet é o mais sensível para as regiões Norte e Nordeste. Já o indicador de renda para a zona rural da região Norte possui dados bem concentrados e com mediana superior a zona urbana, sendo assim exceção no âmbito das regiões analisadas, situação similar aparece para os indicadores de escolaridade da mãe e quantidade de computadores.

Com algumas exceções, portanto, as localidades rurais possuem indicadores de nível socioeconômico menores em relação às localidades urbanas. Essa evidência aponta que esses IFs estão em regiões com um público com menor poder aquisitivo e com maior necessidade de suporte de políticas públicas, principalmente, porque é uma realidade detectada em todas as regiões do país. Essa diferença é mais acentuada para as regiões Nordeste, Sul e Sudeste. Além disso, o indicador de

internet aponta para essa diferença entre as duas localidades, contudo, com maior dispersão para as regiões Norte e Nordeste.

Gráfico 14 - Distribuição dos indicadores socioeconômicos classificados por região e localidade dos IFs (2019).



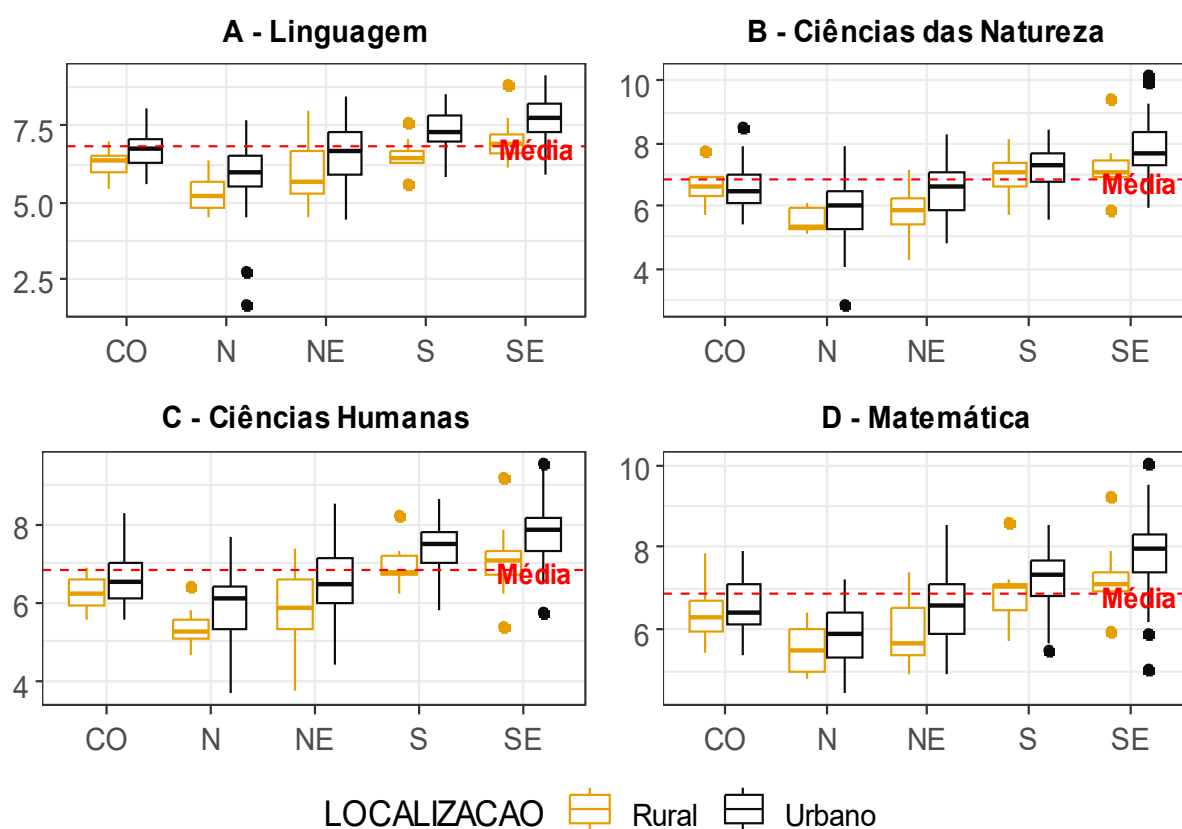
Fonte: Elaborado pelo autor (2023).

Já quando se avalia as notas no Enem, as pontuações das unidades localizadas na zona urbana são superiores às unidades localizadas na zona rural, para todas as áreas do conhecimento avaliadas (gráfico 15), diferentemente dos indicadores de capital humano e capital físico em que as localidades rurais possuem em sua grande maioria melhores indicadores. Além disso, as diferenças nas pontuações são bem acentuadas, evidenciando ainda mais a discrepância entre os IFs dessas duas localidades. As menores diferenças identificadas estão localizadas na região Centro-Oeste que praticamente possui o mesmo padrão para todas as áreas do conhecimento.

Ao se avaliar os indicadores de capital humano, capital físico, nível socioeconômico e de resultado no Enem, com base nas regiões e localidades,

observa-se que as localidades rurais apresentam indicadores melhores em relação à titulação e experiência dos professores. Além disso, os indicadores de capital físico oscilam entre equivalentes a superiores para a localidade rural, no entanto, apesar dos aspectos positivos, a pontuação no Enem, para as regiões urbanas, ainda é superior em relação às regiões rurais. Assim, observa-se que as unidades rurais possuem bom nível de indicadores de capital humano e de capital físico, entretanto, com menor nível socioeconômico, o que pode impactar no desempenho.

Gráfico 15 - Distribuição das notas no Enem classificadas por região e localidade dos IFs (2019).



Fonte: Elaborado pelo autor (2023).

Os resultados por localidade reforçam os estudos quanto ao menor desempenho das escolas da zona rural. Rodrigues (2017) apresenta a disparidade entre os alunos de escolas rurais e urbanas, indicando uma vantagem para os estudantes de escolas urbanas, com a maior parte da discrepância entre os grupos atribuída ao impacto das características das escolas e do ambiente familiar dos alunos. Rocha (2022), ao analisar o desempenho dos estudantes no Enem de 2017 a 2019, para o Nordeste, identificou diferença de desempenho entre os estudantes das

escolas rurais com pontuações médias inferiores. No contexto colombiano, Jurado (2013) evidencia que as escolas rurais seguem essa mesma tendência com menor desempenho em relação às urbanas.

De forma diferente do estudo de Rodrigues (2017), com maior discrepância nas características das escolas rurais em relação às urbanas, os IFs apresentam no nível escolar melhores condições de qualificação docente e capital físico de suas unidades rurais. Apesar disso, no nível dos estudantes, os dados são categóricos ao apresentar uma realidade socioeconômica inferior à das unidades urbanas, reverberando em menor desempenho para as unidades rurais.

Além dessas características, deve-se levar em consideração os níveis de coordenação das políticas de educação envolvidas, pois a educação básica de nível fundamental e médio é de competência dos municípios e estados, cabendo a cada ente federativo, nas diversas regiões, desempenhar suas funções. Isso leva também a diferentes implementações de acordo com as políticas educacionais de cada estado e município, ao passo que a política de educação profissional e tecnológica está vinculada ao Ministério da Educação por meio da Secretaria de Educação Profissional e Tecnológica (Setec), com arcabouço legal uniforme entre as unidades.

Essa realidade acaba permitindo a implementação de uma política que, apesar de todos os seus pontos de melhoria, permite a distribuição de recursos entre os IFs, com coordenação em nível nacional, embora haja outros pontos que devem ser levados em conta como o tempo de funcionamento das unidades. Com isso, uma unidade rural tem direito à contratação de servidores com os mesmos planos de carreira, recursos para funcionamento e recursos para assistência estudantil.

4.1.5 Síntese da seção da análise das estatísticas descritivas

Dos resultados apresentados, destaca-se que, apesar de a Lei 11.892 trazer uma política pública com critérios comuns aos Institutos Federais, ainda assim existem diferenças na aplicação de recursos e no desempenho dos estudantes entre as regiões do país. Essa diferença é mais evidente para o capital humano com os níveis de titulação e experiência inferiores para a região Nordeste e, principalmente, para a região Norte. Esse é um resultado que evidencia a maior fragilidade dessas regiões em virtude de suas unidades atuarem com docentes com menor titulação e experiência.

A preocupação de fixação de servidores fez com que o governo federal emitisse instrumentos de controle, como a Portaria nº 10.723, de 19 de dezembro de 2022, do Ministério da Economia/Secretaria Especial de Desburocratização, Gestão e Governo Digital/Secretaria de Gestão e Desempenho de Pessoal, ao vedar a redistribuição de cargo efetivo ocupado por servidor em estágio probatório bem como estipular como requisito não ter sido o servidor redistribuído nos últimos cinco anos. Em 2023, esse instrumento foi alterado pela Portaria nº 619, de 9 de março de 2023, do Ministério da Gestão e da Inovação em Serviços Públicos/Secretaria de Gestão de Pessoas e Relações do Trabalho com a mudança do requisito de redistribuição anterior de cinco para três anos.

Mesmo com esses instrumentos de controle, as regiões Nordeste e Norte vivenciam um maior fluxo de servidores, em boa parte porque, havendo os requisitos necessários, os processos de redistribuição são efetivados com os servidores retornando às suas regiões de origem ou mesmo mais próximas a elas. É comum, nesse contexto, que os candidatos a uma vaga de professor saiam das regiões Sul e Sudeste especialmente para tentarem uma vaga nas regiões menos populosas e com maior chance de aprovação. Os IFs da região Norte, nesse sentido, são os que mais recebem servidores de outras regiões que demandam retornar para as suas localidades de origem.

Esse ponto do capital humano é significativo porque a qualidade docente é fundamental numa instituição de ensino, principalmente, porque a experiência de ensino reverbera positivamente no desempenho dos estudantes (ROCKOFF, 2004; RAVKIN; HANUSHEK; KAIN, 2005), e é justamente nesse aspecto que as regiões Norte e Nordeste são afetadas. Os dados ainda evidenciam que a experiência possibilita um ambiente de aprendizagem mais rico, isso porque o docente, ao ter mais tempo de serviço, acaba tendo também maior qualificação em nível de pós-graduação e por conseguinte domínio de processos e metodologias de pesquisa que contribuem para a identidade dos IFs dentro de suas missões institucionais.

A análise de correlação, nesse sentido, reforça a importância do professor ao apontar significativa associação dos indicadores de capital humano com o desempenho dos estudantes, principalmente, com o indicador de experiência, que apresenta um efeito positivo tanto num cenário de menor quanto de maior nível socioeconômico, enquanto a titulação se faz mais efetiva num contexto de menor condição socioeconômica. Esse resultado evidencia a importância do professor no

processo de ensino-aprendizagem ao relacionar que a qualidade da educação depende tanto de docentes experientes como qualificados, sendo que o aprendizado é mais importante do que anos de escolaridade (HANUSHEK; WOBMANN, 2007).

Dado esse contexto, há que se pensar em políticas que possam amenizar os impactos oriundos de menores níveis de capital humano, pois os resultados dos estudantes são diretamente afetados, em especial, os dos estudantes das regiões Nordeste e Norte. Isso não quer dizer que os docentes dessas regiões não conseguem resultados equivalentes a outras unidades de outras regiões, mas que, se eles estivessem nas mesmas condições de qualificação dos demais, poderiam ser mais eficientes em suas atuações. Por isso, nota-se a importância de investimento na qualidade dos professores, seja por meio de qualificação na pós-graduação, seja pela valorização da experiência de ensino.

Além dos indicadores de capital humano das regiões Nordeste e Norte estarem abaixo das demais regiões, há a confluência dos indicadores de nível socioeconômico na mesma condição. O Nordeste apresenta a menor renda e a menor escolaridade da mãe, ao passo que o Norte revela o maior número de pessoas na família, o menor número de computadores por família e o menor nível de internet. O IBGE (2020) aponta, para essa realidade, indicando que a taxa de analfabetismo das pessoas de 15 anos ou mais de idade é liderada pelo Nordeste, com aproximadamente 14%, seguido de Norte, com 8%, Centro-Oeste, 5%, Sul e Sudeste com as menores taxas de 4%.

O nível socioeconômico é uma das condições que mais afeta os estudantes, pois situações de pobreza e mesmo de menores níveis de renda diminuem drasticamente as oportunidades, a partir da tenra infância, impactando desde a nutrição até o consumo de bens econômicos e culturais. Um menor *background* familiar torna-se assim uma barreira invisível e dura que vai subtraindo do estudante as oportunidades de um ambiente familiar que incentiva a curiosidade e o estudo. Um estudante com uma família, por exemplo, que possui internet, computador, TV, acesso a serviços de *streaming*, moradia adequada, pais com bom estudo formal e que vez ou outra viaja nas férias já está diante de um ambiente de melhores oportunidades.

Esses aspectos estão consolidados na literatura, com forte impacto no desempenho dos estudantes (COLEMAN *et al.* 1966; CURI; MENEZES-FILHO, 2013; SOUZA; OLIVEIRA; ANNEGUES, 2018; ASSUNÇÃO; ARAÚJO; ALMEIDA, 2019), tanto que todas as áreas avaliadas no Enem apresentam correlações elevadas,

significativas a 1% e com coeficiente de correlação próximo a 0,77 entre o indicador de quantidade de computadores por família e a área de Ciências Humanas. A região Norte apresenta o menor desempenho, em todas as áreas do conhecimento, enquanto a região Nordeste mostra a segunda menor nota em Linguagem, Ciências da Natureza e Ciências Humanas, com resultado expressivo em Matemática, a frente das regiões Norte e Centro-Oeste; e Redação, atrás apenas da região Sudeste.

O desempenho por regiões já foi documentado por Viggiano e Matos (2013), ao identificar, no contexto nacional do Enem de 2010, três grupos de performance: (i) grupo superior, com Sul e Sudeste; (ii) grupo médio, com Centro-Oeste; e (iii) grupo inferior, com Norte e Nordeste. Contudo, há uma diferença que merece destaque, os dados indicam que a região Nordeste dispõe de desempenho expressivo em Matemática e Redação, superando com isso indicadores de menor nível de capital humano e socioeconômico quando comparados aos IFs das demais regiões do país.

Além disso, as localidades rurais, como era esperado, demonstram um desempenho inferior em comparação às localidades urbanas, revelando assim mais um grupo com características próprias. Apesar disso, há algumas particularidades a serem consideradas, como o fato de as unidades rurais apresentarem níveis mais elevados de capital humano e capital físico, porém, com menor nível socioeconômico. Assim, como o nível socioeconômico é menor para Nordeste e Norte, as unidades rurais seguem a mesma tendência das regiões, com a evidência de estarem numa posição inferior à das unidades urbanas.

Esse é um dos resultados importantes, pois mesmo as unidades rurais estando com melhores níveis de capital humano e capital físico, há a confluência dos menores níveis socioeconômicos da amostra, resultando em menores desempenhos. Fica evidente que a necessidade de esforço institucional para menores níveis socioeconômicos deve ser significativamente superior quanto menor for a condição socioeconômica do estudante. Observa-se, portanto, que esse critério é atendido pelos IFs, propiciando melhores condições de ensino, em média, para as unidades rurais, o que leva a uma melhora na performance dos estudantes, caso contrário, o desempenho poderia ser menor.

Do exposto, sobressai a existência de desigualdade entre as regiões em âmbito nacional, com fragilidade para as regiões Norte e Nordeste, as quais acumulam menor capital humano, menor nível socioeconômico e menor desempenho. Em especial, a região Norte sofre mais com menor titulação e experiência de seus professores em

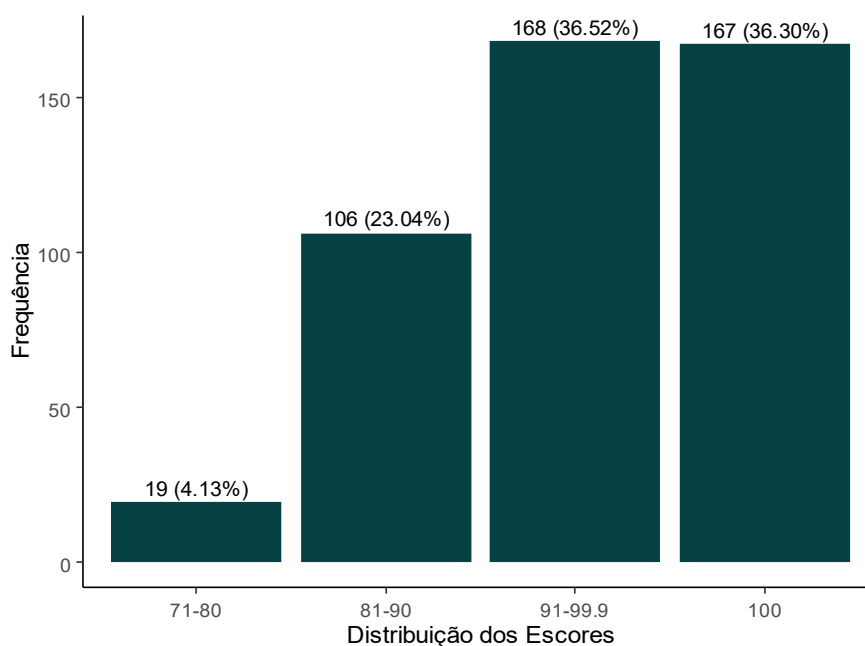
virtude de maior fluxo de servidores para outras regiões, acumulando menor nível de acesso à internet, maior número de pessoas por família e menor desempenho no Enem. A região Nordeste, por seu turno, apresenta a segunda menor titulação docente, acumulando os menores níveis de renda e escolaridade da mãe, além do segundo menor desempenho no Enem, com a exceção para as áreas de Matemática e Redação.

4.2 Análise dos determinantes da função de produção educacional

A contribuição de cada insumo, tendo como produto o desempenho dos estudantes dos cursos da educação técnica de nível médio na forma integrada dos Institutos Federais de Educação (IFs) e o nível de eficiência de cada um dos institutos, foi determinada com o emprego da Análise Envoltória de Dados (DEA) orientado para a maximização do desempenho (output). Foram analisados 460 IFs como DMUs do modelo DEA e, dessas, 167 foram classificadas como eficientes (36,30%), sendo que a distribuição dos níveis de eficiência pode ser observada no gráfico 16.

Os 19 IFs com menor performance apresentaram escores de eficiência entre 71 e 80 do obtido pelos IFs eficientes. Isso aponta que muitos dos IFs, considerando os insumos de capital humano, capital físico e nível socioeconômico, operam abaixo de suas capacidades. No total, são 293 IFs, 63,70% que não alcançaram a eficiência, o que implica que muitas dessas unidades estão com folga em seus insumos, ou seja, não aproveitando totalmente seus insumos de capital humano e capital físico no nível escolar, ou mesmo não atuando de forma efetiva num cenário de condições socioeconômicas favoráveis.

Gráfico 16 - Frequência dos escores de eficiência técnica dos IFs (2019) por quantidade e percentual analisadas.



Fonte: Elaborado pelo autor (2023).

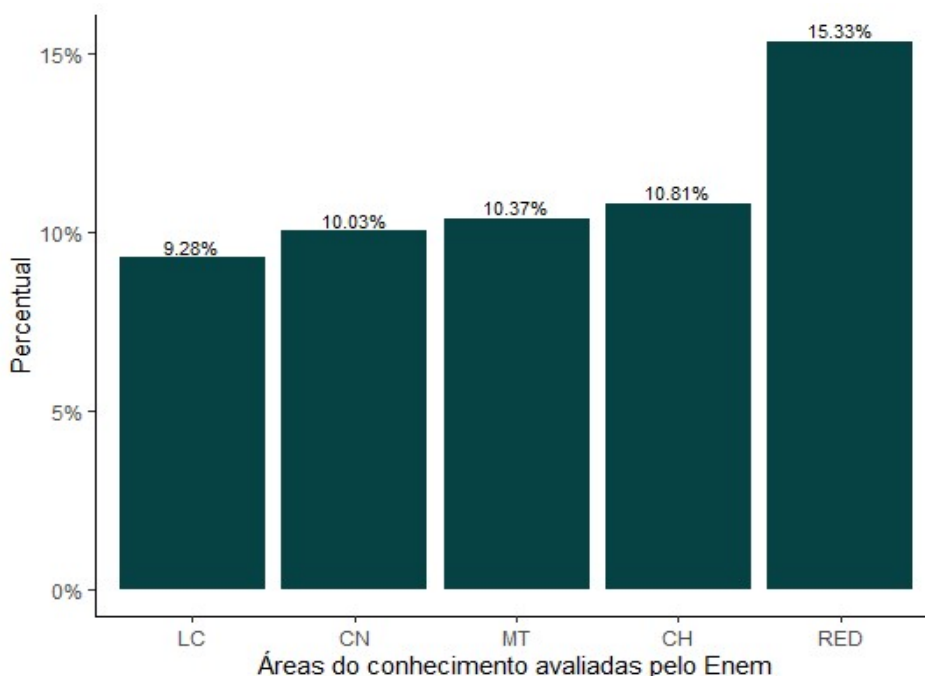
Considerando os insumos disponibilizados para os IFs, dada as condições individuais dos alunos, os percentuais dos potenciais de melhoria no desempenho dos alunos são similares entre as áreas de Linguagem (9,28%), de Ciências da Natureza (10,03%), Matemática (10,37%) e Ciências Humanas (10,81%), conforme mostra o gráfico 17. A fim de exemplificar o que esses dados indicam, imagine-se que, se os IFs utilizassem os recursos da mesma forma que os IFs eficientes, então, sendo a nota média de Matemática, hipoteticamente, de 500, o desempenho médio dos estudantes dos IFs passaria para 551 pontos.

Ainda de acordo com os potenciais de crescimento, mantidos os insumos na mesma proporção, a área de Redação foi a que apresentou a maior oportunidade de melhoria, com um percentual de 15,33% (gráfico 17). Entretanto, como visto na apresentação das estatísticas descritivas, essa área apresenta uma dispersão significativamente superior às demais e se atribui essa dispersão a questões inerentes ao conhecimento dos alunos e à complexidade da avaliação, que engloba cinco competências, a saber: (i) demonstrar domínio da modalidade escrita formal da língua portuguesa; (ii) compreender a proposta de redação e aplicar conceitos das várias áreas de conhecimento para desenvolver o tema, dentro dos limites estruturais do texto dissertativo-argumentativo em prosa; (iii) selecionar, relacionar, organizar e

interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista; (iv) demonstrar conhecimento dos mecanismos linguísticos necessários para a construção da argumentação; (v) elaborar proposta de intervenção para o problema abordado, respeitando os direitos humanos (BRASIL, 2022). Logo, obter ganhos nessa área é tão complexo como é a sua avaliação.

Apenas a título de exemplificação da complexidade em se estabelecer uma pontuação para a redação no Enem, para cada competência, há a atribuição de uma nota entre 0 e 200, por cada avaliador, que, ao final, poderá chegar a 1.000 pontos. A nota final será composta pela média aritmética das notas totais indicadas pelos dois avaliadores. Caso a nota atribuída por um dos avaliadores difira em mais de 100 pontos no total ou ainda seja superior a 80 pontos em qualquer uma das competências, a redação será avaliada, de forma independente, por um terceiro avaliador, sendo a nota final a média das duas notas totais que mais se aproximarem. Caso a discrepância ainda continue, após a terceira avaliação, a redação passará por uma banca de três professores que indicará a nota final (BRASIL, 2022).

Gráfico 17 - Percentuais dos potenciais de melhoria dos IFs (2019) com eficiência menor que 100 por área de conhecimento (outputs) do Modelo DEA-BCC-Output.



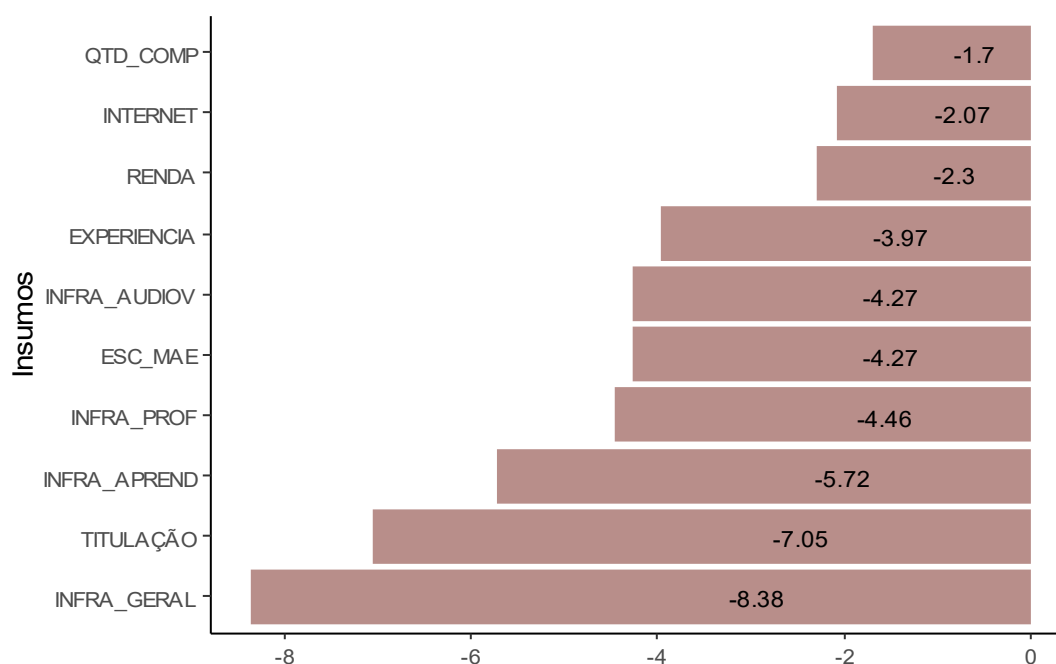
Fonte: Elaborado pelo autor (2023).

Em relação ao emprego dos insumos dos IFs que não alcançaram a eficiência, conforme o gráfico 18, a infraestrutura geral é o indicador com a maior folga entre os

IFs analisados. Ou seja, é o recurso com menor aproveitamento pelos IFs, de forma que poderiam ser reduzidos em 8,38% que não afetaria o desempenho dos estudantes, se todos os IFs fossem eficientes no seu uso. A titulação é o segundo recurso com maior folga, indicando que, em média, os IFs aproveitam somente 92,95% desse recurso. Já no que tange aos estudantes, verificou-se que o nível de escolaridade da mãe teve uma ociosidade de 4,27% e o indicador de acesso à internet de 2,07%, de maneira que alguns IFs não estão conseguindo aproveitar todas as competências individuais dos seus alunos.

A identificação dessas folgas é importante para estabelecer políticas públicas como capacitação das escolas no uso de recursos de laboratórios e outros recursos pedagógicos para que possibilitem otimizar o uso dos diferentes insumos disponibilizados para as instituições e os potenciais de seus alunos. Dessa forma, as análises realizadas possibilitam avaliar o nível potencial de melhoria no desempenho dos alunos mantida a estrutura atual de recursos e quais desses estão tendo dificuldades de extração de todo o potencial que ele possibilita.

Gráfico 18 - Percentuais de folgas por indicadores de capital físico, humano e nível socioeconômico do Modelo DEA-BCC-Output para 293 IFs com eficiência menor que 100% dos IFs (2019).



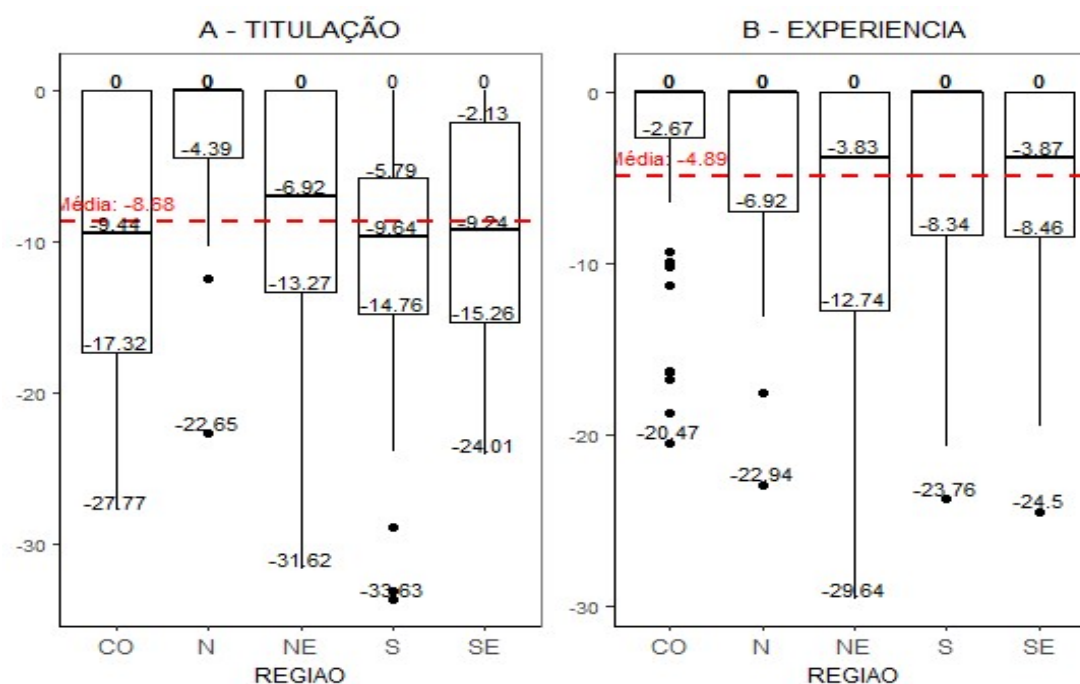
Fonte: Elaborado pelo autor (2023).

Quando se avalia as folgas dos insumos dos 293 IFs que não alcançaram a eficiência por regiões do país (gráfico 19), observa-se que a mediana do indicador de titulação, gráfico 19 A, é mais elevada para as regiões Sul (-9,44%), Centro-Oeste (-9,44%) e Sudeste (-9,24%), representando assim 50% dos IFs com uso desse recurso em níveis mais elevados, acima da média geral de -8,68%. Destaca-se a região Norte em que a mediana para esse insumo está próxima a zero, indicando que a metade (17) dos IFs utilizaram todo esse recurso.

Quanto ao indicador de experiência, gráfico 19 B, destacam-se as regiões Centro-Oeste, Norte e Sul com as medianas próximas a zero, com 50% dos IFs com consumo efetivo desse insumo. Por outro lado, as regiões Nordeste e Sudeste evidenciam folgas maiores, com os IFs do Nordeste com a maior variabilidade. Essas evidências, portanto, apontam que as folgas dos indicadores de capital humano divergem quanto à sua efetividade entre os IFs por regiões, com a titulação sendo menos aproveitada do que a experiência, o que está relacionado a menor eficiência dessas unidades quando combinado ao conjunto de variáveis de capital físico.

Essa realidade pode estar relacionada à própria forma como os IFs estão organizados, com a possibilidade da oferta de cursos de graduação e pós-graduação e a verticalização do ensino. Com isso, quanto mais doutores uma unidade possui, maior a tendência da oferta de especializações, mestrados e doutorados, com a realocação dos docentes em cursos superiores e de pós-graduação. Esse contexto pode ser mais evidente para regiões com maior oferta de professores qualificados. Há, assim, já na contratação, via concurso público, maior chance da entrada de professores doutores nos IFs das regiões centrais do país. Por outro lado, as regiões com menor oferta de profissionais com maior titulação têm dificuldades na contratação e permanência desses servidores.

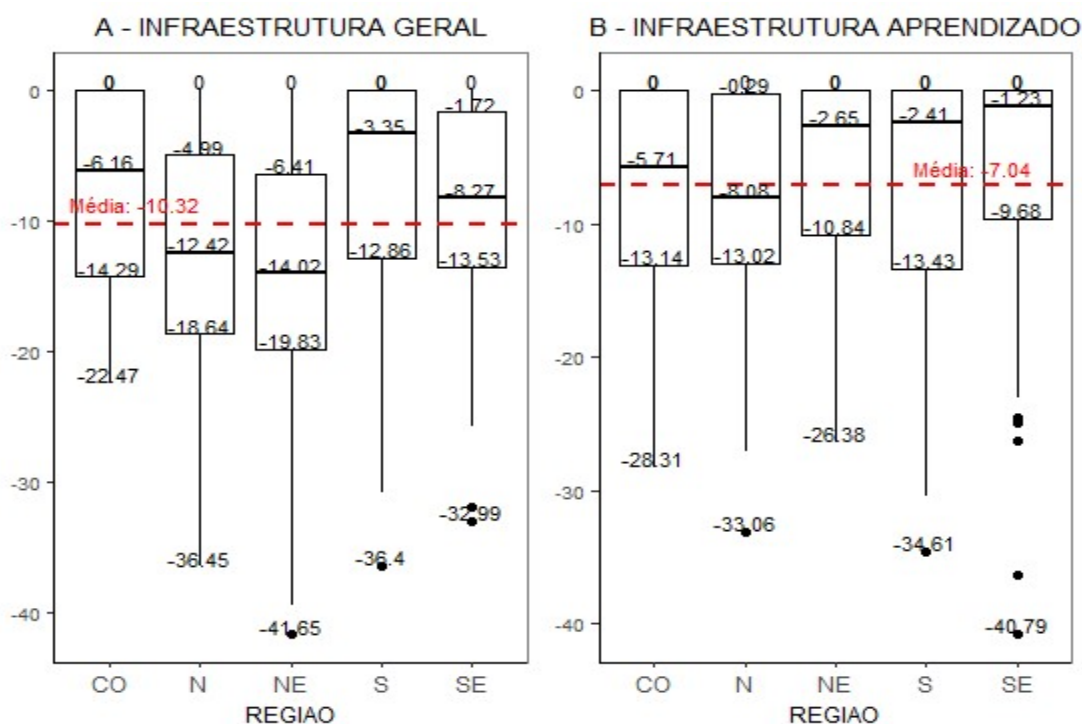
Gráfico 19 - Distribuição percentual das folgas dos insumos de Titulação e Experiência dos 293 IFs (2019) com escores ineficientes por regiões do país.



Fonte: Elaborado pelo autor (2023).

Quantos aos indicadores de capital físico, a infraestrutura geral (gráfico 20 A) é menos aproveitada pelos IFs da região Norte, com mediana de -12,42% e Nordeste com mediana de -14,02%, superiores à média geral de -10,32%. Além disso, essas duas regiões respondem pela maior variabilidade entre os IFs, chegando a -36,45% e -41,65%, respectivamente. Com a exceção da região Sul, com mediana de -3,35%, as demais regiões são equivalentes quanto às folgas de recursos de infraestrutura geral. Em relação à infraestrutura de aprendizado (gráfico 20 B), as regiões Norte e Centro-Oeste apresentam as maiores medianas, com -8,08% e -5,71%, respectivamente. De forma contrária, as regiões Sudeste (-1,23), Sul (-2,41) e Nordeste (-2,65) apresentam medianas próximas e com as menores folgas.

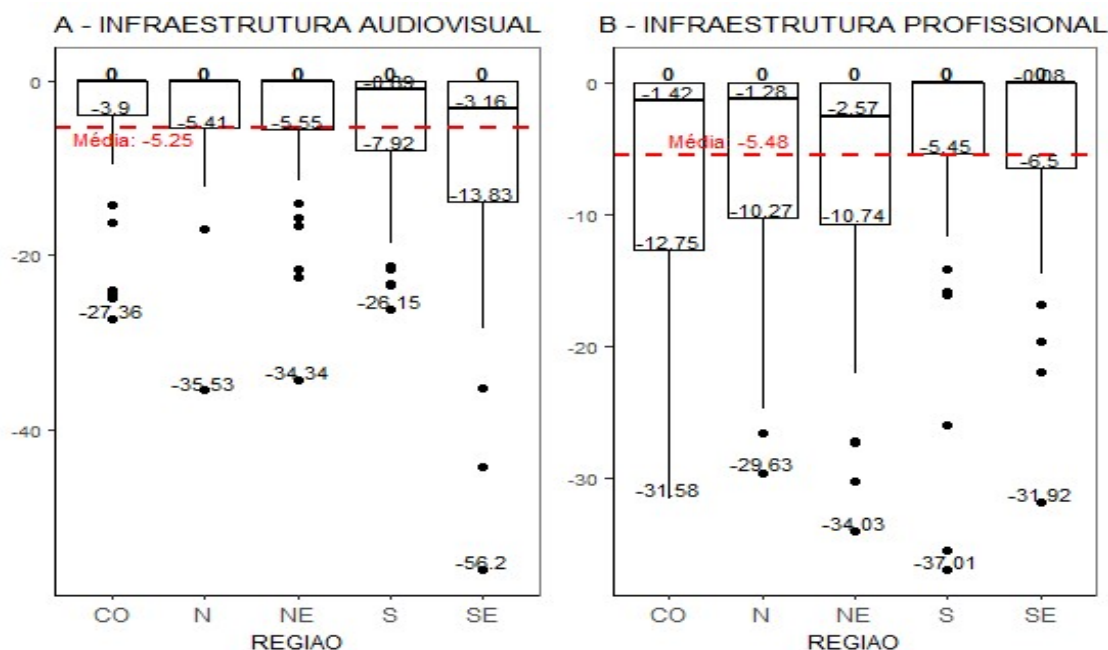
Gráfico 20 - Distribuição percentual das folgas dos insumos de Infraestrutura Geral e de Aprendizado dos IFs (2019) com escores ineficientes por regiões do país.



Fonte: Elaborado pelo autor (2023).

Quando se avalia o indicador de infraestrutura audiovisual (gráfico 21 A), ainda dentro dos indicadores de capital físico, apenas a região Sul (-0,89) e Sudeste (-3,16) não evidenciam uma mediana com valor zero, porém, próximos. Praticamente para 50% dos IFs, houve o aproveitamento total desse recurso, enquanto para os demais, os valores foram inferiores ao indicador de infraestrutura geral e infraestrutura de aprendizado. A infraestrutura profissional, gráfico 21 B, Sul (0%) e Sudeste (0%) evidenciam medianas com valores zero ou próximo, enquanto as demais regiões estão com as medianas bem próximas, porém, com variação maior entre os IFs.

Gráfico 21 - Distribuição percentual das folgas dos insumos de Infraestrutura audiovisual e Profissional dos IFs (2019) com escores ineficientes por regiões do país.



Fonte: Elaborado pelo autor (2023).

A tabela 6 possibilita identificar as médias das folgas de todos os indicadores de capital físico e capital humano por regiões do país, evidenciando de forma geral que as regiões diferem entre si quanto à utilização de recursos. Enquanto algumas estão com maiores folgas em titulação, outras estão sobretudo com maiores folgas em infraestrutura geral e infraestrutura de aprendizado. A região Norte, nesse sentido, evidencia a menor média de folga em titulação (-2,67%), mas com folga média elevada em infraestrutura geral (-13,06%). Já para a região Nordeste, há folgas elevadas tanto na titulação (-7,76%) quanto na infraestrutura geral (-14,43%).

Quando se analisam as maiores folgas nos indicadores para todo o conjunto de dados, a infraestrutura geral evidencia a média mais elevada das regiões com -10,42%, seguida da titulação com -8,17% e da infraestrutura de aprendizado com -7,34%. Especificamente, as regiões apresentam folgas maiores para infraestrutura geral e titulação, apontando para esses dois indicadores como os que mais contribuem para os IFs que não alcançaram a eficiência, impactando na aprendizagem dos estudantes.

Tabela 6 - Descrição das folgas percentuais dos indicadores de capital humano e capital físico dos IFs (2019) por regiões do país.

Região	Capital Humano %		Capital Físico %			
	Titulação	Experiência	Geral	Aprendizado	Audiovisual	Profissional
CO	-10.15	-3.09	-7.70	-7.58	-3.97	-7.46
N	-2.67	-3.55	-13.06	-9.45	-3.95	-6.80
NE	-7.76	-6.84	-14.43	-5.89	-3.63	-6.64
S	-10.82	-4.37	-7.71	-7.69	-5.30	-3.79
SE	-9.44	-5.30	-9.21	-6.10	-8.15	-3.98
Média	-8.17	-4.63	-10.42	-7.34	-5.00	-5.73
Máximo	-10.82	-6.84	-14.43	-9.45	-8.15	-7.46
Mínimo	-2.67	-3.09	-7.70	-5.89	-3.63	-3.79
Desvio Padrão	3.28	1.49	3.13	1.44	1.87	1.72

Fonte: Elaborada pelo autor (2023).

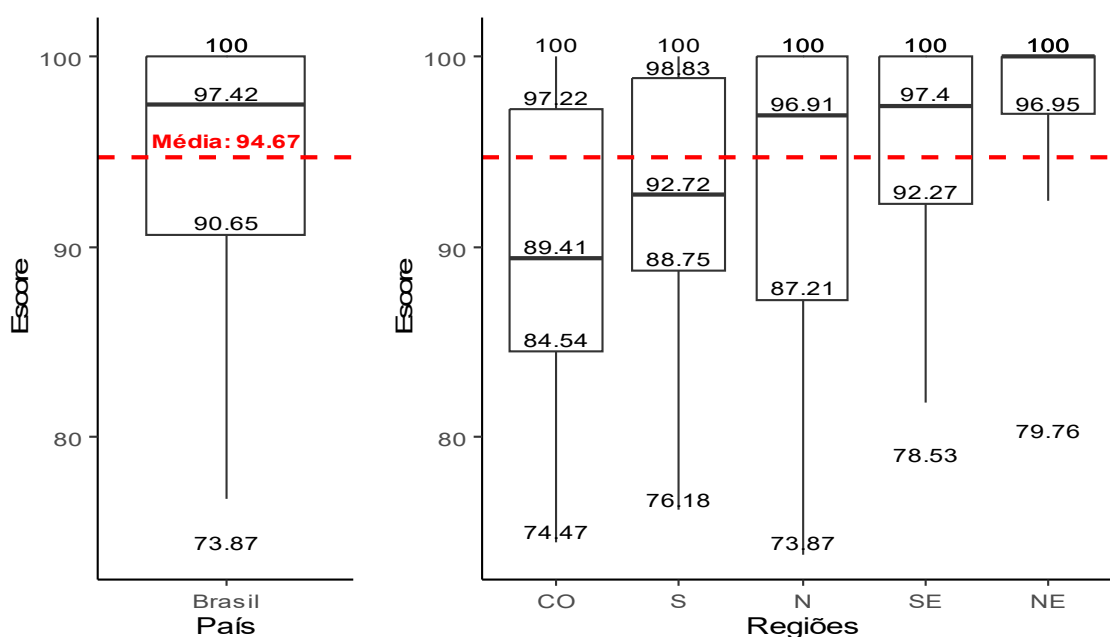
Quanto aos escores de eficiência, gráfico 22, a região com menor mediana pertence à região Centro-Oeste, com 50% dos IFs daquela região contendo escores de eficiência abaixo de 89,41. A região Norte apresenta a maior dispersão da amostra com o menor escore de eficiência (73,87) com 25% de seus IFs abaixo do escore de 87,21, por outro lado, no entanto, a metade de suas unidades estão concentradas com escores acima de 96,91 bem próximos à eficiência. Essas evidências revelam que a região Norte dispõe da maior disparidade entre os IFs, com unidades presentes em todos os intervalos de eficiência da amostra.

A região Sul apresenta escores de eficiência próximos aos menores do conjunto de dados, com a segunda menor mediana, porém, com 25% de seus IFs com escores superiores a 98,83. Já a região Sudeste possui escores elevados, variando de 92,27 a 100 para 75% dos seus IFs, sendo que 50% deles estão com escores maiores ou iguais a 97,4. De todas, porém, a que se destaca é a região Nordeste, que, apesar de apresentar alguns valores discrepantes, possui 75% de seus IFs com escores superiores a 96,95, sendo que mais de 50% são eficientes. Acrescenta-se ainda que o menor escore dessa região, considerado *outlier*, ainda é superior aos das demais regiões do país.

Quanto aos escores de eficiência em âmbito nacional (gráfico 22), 25% dos IFs estão com escores abaixo de 90,65, com 50% de eficiência igual ou superior a 97,42, ou ainda, com 75% de escores superiores a 90,65. Quando comparada às regiões Centro-Oeste, Sul e Norte, com a distribuição nacional, apresentam escores de

eficiência menores em relação ao primeiro quartil, situação similar para as medianas dessas três regiões. Contrariamente, a região Sudeste (97,4) evidencia uma distribuição similar a nacional (97,42), com os valores em cada quartil próximos, o que difere da região Nordeste, com todos os valores bem superiores à nacional, conforme evidenciado no gráfico.

Gráfico 22 - Distribuição dos escores de eficiência dos IFs (2019) em âmbito nacional e por regiões do país.



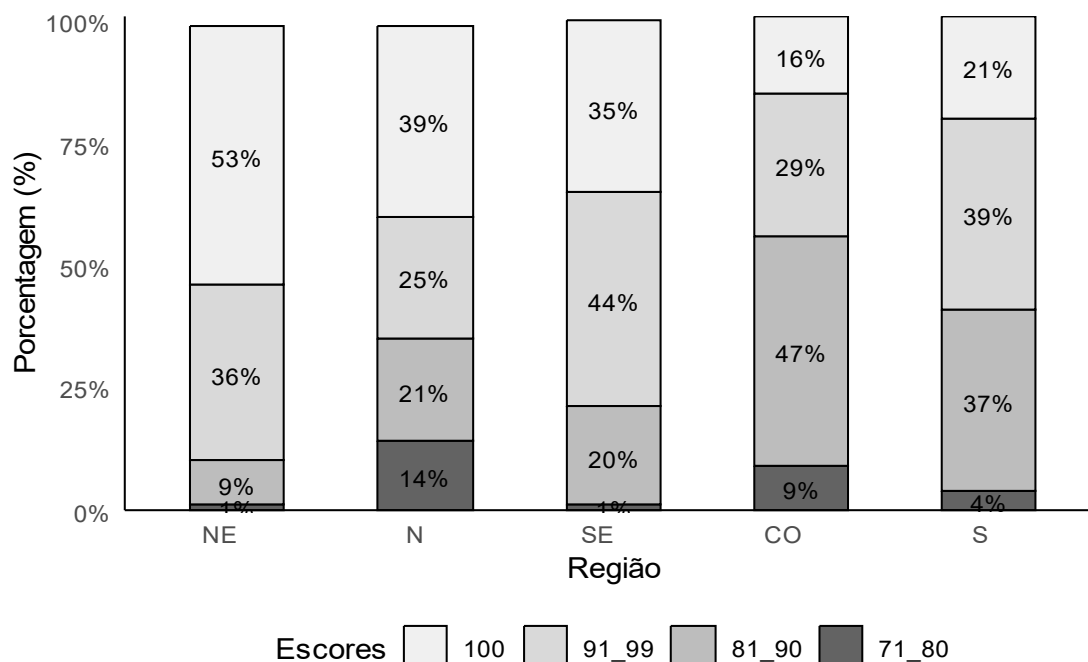
Fonte: Elaborado pelo autor (2023).

O gráfico 23 permite verificar a distribuição geral da eficiência dos IFs. Sobressai a região Nordeste, com a maior quantidade de IFs (148) ao mesmo tempo em que possui mais da metade de suas unidades eficientes (53%), com 89% delas compreendidas no intervalo de escore de 91 a 100. Apenas duas unidades estão no menor intervalo de eficiência (71 a 81). Apesar de haver indicadores socioeconômicos baixos, a região consegue aplicar os seus recursos de maneira eficiente, obtendo bons resultados, considerando seu capital humano, físico e nível socioeconômico.

A região Norte também se destaca, apesar de suas dificuldades, possui 39% de IFs eficientes, ficando atrás apenas da região Nordeste, com 64% das unidades na faixa de 91 a 100. Contudo, possui a maior proporção de unidades com escore de eficiência de 71 a 81 (14%), incluindo o campus com a menor pontuação da amostra, indicando que muitas de suas unidades não estão operando em sua capacidade total. Esses resultados evidenciam, assim, que mesmo as regiões com maiores

disparidades socioeconômicas conseguem IFs que aplicam de forma eficiente seus recursos.

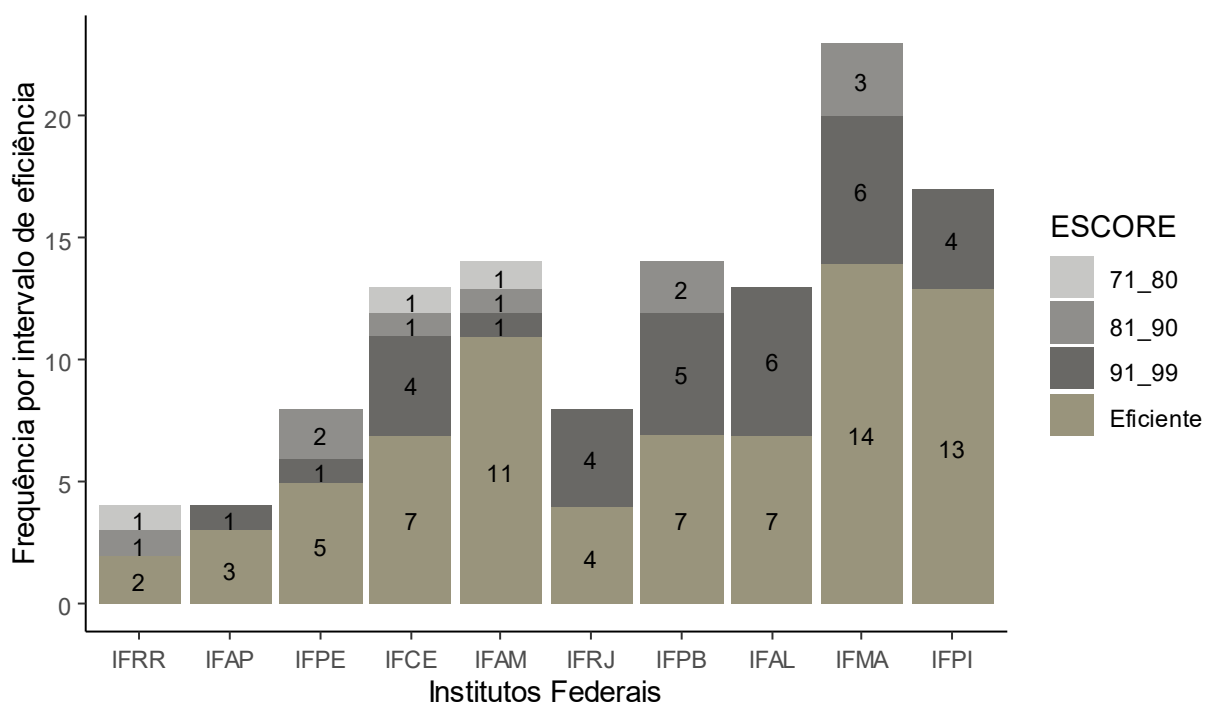
Gráfico 23 - Escores de eficiência dos IFs (2019) por regiões do país das 460 unidades avaliadas.



Fonte: Elaborado pelo autor (2023).

Quando se analisa os escores dos Institutos Federais por intervalos de eficiência, alguns se sobressaem. A partir do gráfico 24, é possível identificar dez institutos com mais de 50% de seus IFs classificados como eficientes. Destacam-se o Instituto Federal do Amazonas (IFAM), com o maior percentual de unidades eficientes (79%); o Instituto Federal do Piauí (IFPI), com 76% de suas 13 unidades classificadas como eficientes; e o Instituto Federal do Amapá (IFAP), com 75% de suas três unidades classificadas como eficientes de um total de quatro. Dos dez institutos com mais IFs eficientes, seis são da região Nordeste, três da região Norte e um da região Sudeste, reforçando com isso o realce das duas primeiras regiões.

Gráfico 24 - Número de Institutos Federais com 50% ou mais de unidades eficientes (2019).



Fonte: Elaborado pelo autor (2023).

A partir da análise dos resultados, evidencia-se que a eficiência dos IFs é impactada de forma desigual entre as regiões, com diferenças em todos os insumos, no entanto, de forma mais acentuada, nos insumos de infraestrutura geral e titulação. A ineficiência apontada, nesse sentido, pode estar associada aos IFs que, mesmo com elevada infraestrutura geral, apresentam desempenho relativo inferior aos IFs nas mesmas condições. De forma contrária, aqueles com baixa infraestrutura geral podem estar com desempenho médio superior, situação similar para a titulação.

No que se refere às regiões, o Nordeste se destaca pelo robusto número de IFs eficientes e reduzido número no menor intervalo de eficiência. Já a região Norte apresenta resultados mistos, com IFs eficientes, mas também com a maior proporção de unidades, no menor intervalo de eficiência, além disso, com o menor escore da amostra. Por fim, a região Centro-Oeste tem a menor quantidade de IFs eficientes, indicando a necessidade de melhoria na utilização de seus recursos de capital humano e capital físico.

Os resultados, desse modo, reforçam os achados sobre a eficiência dos IFs da região Nordeste e revelam que a região Norte, mesmo com todas as dificuldades oriundas de um baixo nível socioeconômico e, principalmente, pela menor proporção

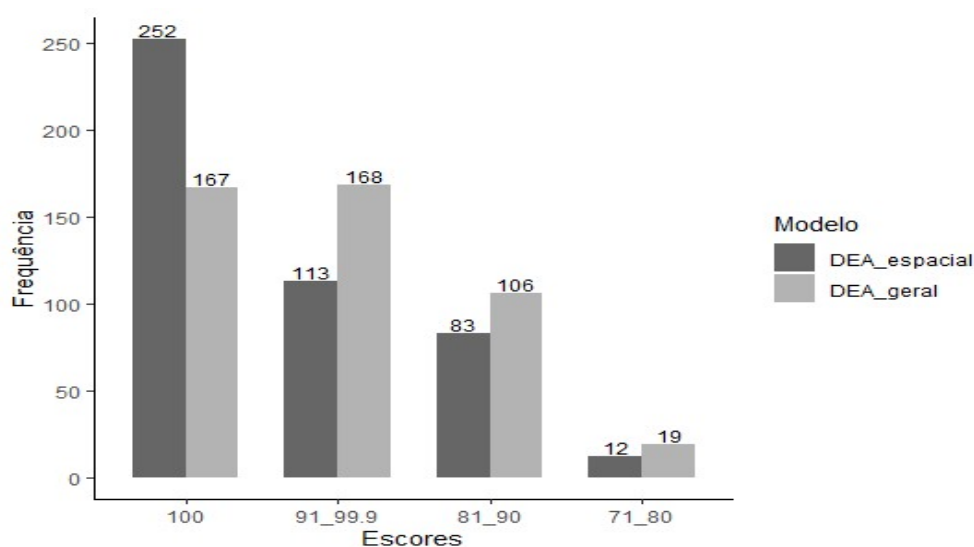
de docentes doutores e com menor experiência, consegue um desempenho significativo. No entanto, as limitações apontadas fazem também com que a região tenha o maior percentual de unidades na menor faixa de eficiência. As disparidades, em termos de desempenho, entre a região Sudeste e Norte, por exemplo, são compensadas porque os IFs da região Norte eficientes conseguem otimizar os seus insumos e assim alcançar resultados significativos dadas as suas condições.

4.3 Análise DEA espacial

Tendo em vista as disparidades de eficiência entre os IFs, nas diversas regiões do país, com diferenças no aproveitamento dos insumos, esta seção aborda a Análise Envoltória de Dados (DEA) com insumos variando espacialmente, no nível dos estados para indicadores de contexto socioeconômico.

Para isso, foram incluídas as variáveis do Ideb (Índice de Desenvolvimento da Educação Básica) dos anos finais do ensino fundamental, o rendimento médio per capita e o percentual da população urbana residente em domicílios ligados à rede de esgotamento sanitário no nível dos estados. O gráfico 25 permite identificar que houve uma melhora significativa na eficiência, passando de 167 IFs eficientes no modelo DEA geral para 252 no modelo DEA espacial, aumento de 50,89%. Por causa desse acréscimo, os demais intervalos de eficiência diminuíram, com isso, o número de IFs no menor intervalo de eficiência (71_80) saiu de 19 para 12 unidades no modelo DEA espacial.

Gráfico 25 - Frequência comparativa dos escores de eficiência dos IFs (2019) entre os modelos DEA geral e DEA espacial.



Fonte: Elaborado pelo autor (2023).

Dentro das regiões, no modelo DEA Espacial, houve um crescimento de 19 IFs com escores eficientes na região Norte, acréscimo de 86%, conforme pode ser observado no gráfico 26. Para o Nordeste, também houve mudança, saindo de 78 para 129 IFs eficientes, 65%, mesmo percentual da região Sul, que partiu de 17 para 28 unidades eficientes. Os IFs da região Sudeste pouco se alteraram, com um acréscimo de apenas quatro unidades. De forma contrária aos IFs das regiões Norte e Nordeste, o Centro-Oeste não alterou a quantidade de unidades, mantendo quatro unidades eficientes nos dois modelos incrementados.

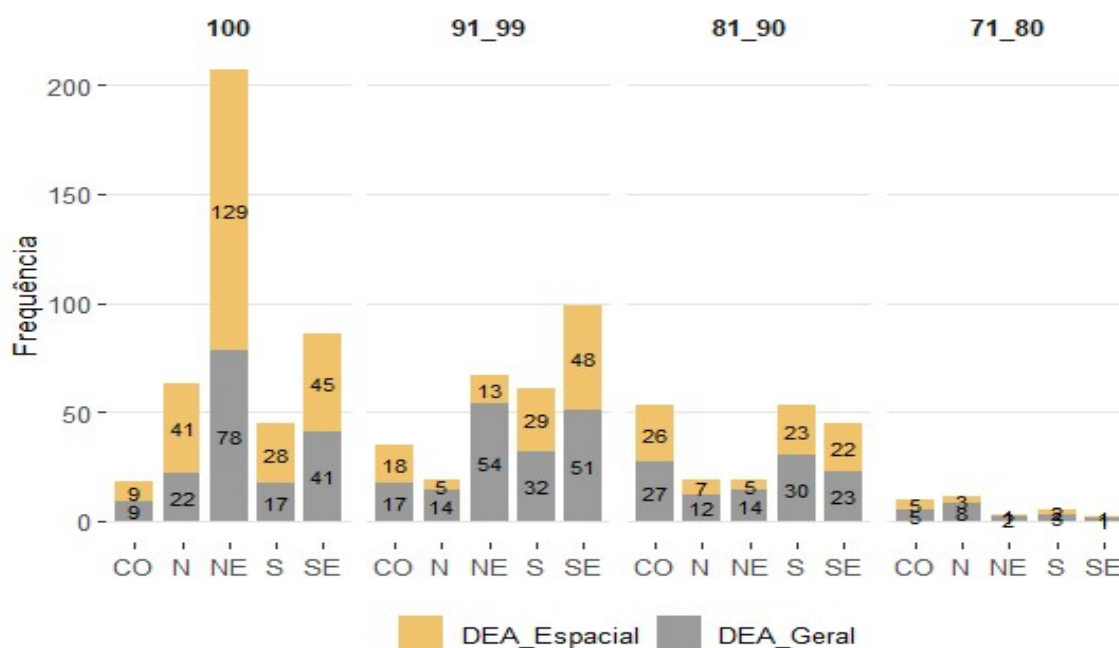
Como as regiões Norte e Nordeste apresentam os menores indicadores socioeconômicos, pode-se compreender que, mesmo diante de um cenário desfavorável, os IFs dessas regiões conseguem bons resultados relativos. Os indicadores de contexto socioeconômico confirmam esse cenário, tendo a região Nordeste o menor rendimento médio mensal real domiciliar per capita de R\$ 984,24 juntamente com a região Norte, com R\$ 1.036,28. Os rendimentos das demais regiões são superiores, com R\$ 1.714,53 para a região Centro-Oeste, R\$ 1.780,62 para região Sudeste e R\$ 1.965,75 para a região Sul.

Essa melhora na eficiência pode estar relacionada ao fato de as unidades localizadas em regiões com menores indicadores socioeconômicos conseguirem boas notas no Enem, alocando de forma eficiente os seus recursos de capital humano e

capital físico, frente a um cenário mais desafiador em comparação a outras regiões do país. De outra forma, a região Centro-Oeste, mesmo com renda superior às regiões Norte e Nordeste, não aumentou o número de IFs eficientes, indicando que não aloca seus recursos de maneira a impactar a aprendizagem tão fortemente como as regiões Norte e Nordeste num cenário com melhores condições sociais.

Já a região Sul apresenta uma realidade diferente, com a maior renda média das regiões, com R\$ 1.965,75 per capita, conseguindo aumentar o número de unidades eficientes, passando de 17 para 28, um acréscimo de 65%. As boas condições, nesse sentido, foram importantes para impulsionar o resultado de seus estudantes de forma relativa. O fato de estar comparativamente numa situação melhor não penalizou os seus resultados, pelo contrário, viabilizou o aumento de IFs eficientes, com a alocação otimizada de seus recursos.

Gráfico 26 - Comparação da distribuição de frequência dos escores de eficiência dos IFs (2019) por regiões do país entre os modelos DEA geral e DEA espacial.



Fonte: Elaborado pelo autor (2023).

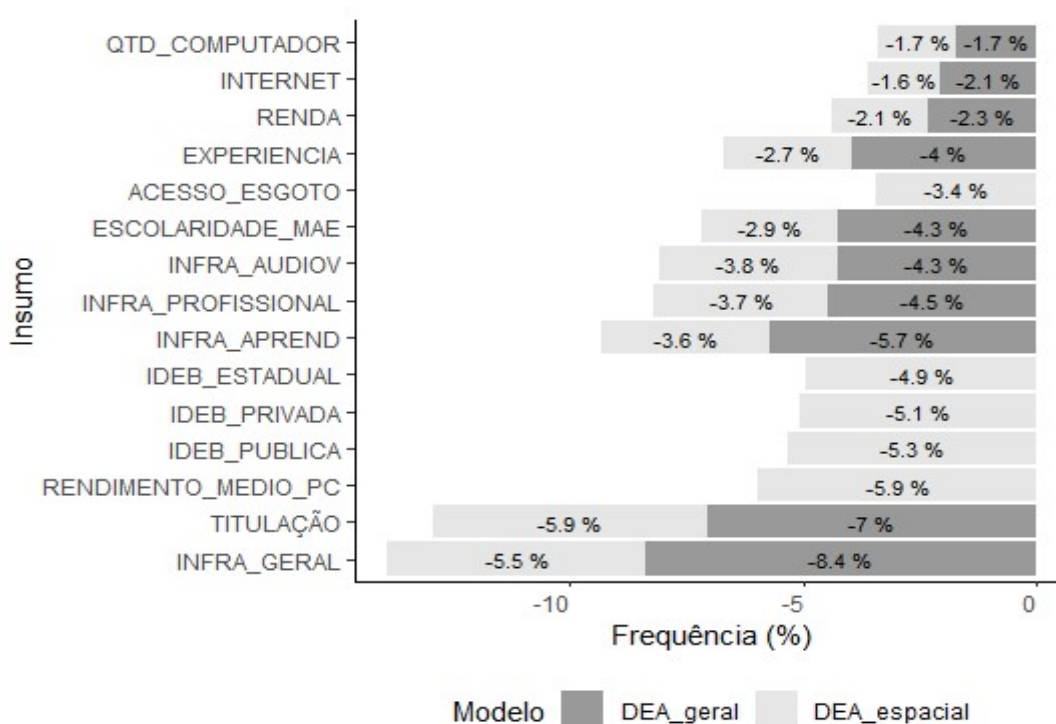
Quando se avalia as folgas dos insumos apenas para o modelo DEA espacial, para os IFs que não alcançaram a eficiência, conforme o gráfico 27, há a prevalência das variáveis de nível estadual com os maiores percentuais. O rendimento médio per capita juntamente com a titulação aparece com a maior folga (-5,9%), seguido da infraestrutura geral (-5,52%). Ao se comparar os dois modelos, há uma diminuição nas

folgas dos insumos pelo modelo DEA espacial, com a titulação saindo de uma folga de -7% para -5,9%, no entanto, ainda com o maior percentual.

Além disso, o gráfico 27 evidencia que, ao se incluir dados no nível estadual da condição socioeconômica, cria-se um bloco com as maiores folgas nesse nível, com a exceção da variável de esgotamento sanitário. Esse bloco insere-se no conjunto de insumos, totalizando quatro grupos com folgas similares: (i) titulação e infraestrutura geral; (ii) rendimento médio per capita, Ideb da rede pública, Ideb da rede privada e Ideb da rede estadual; (iii) Infraestrutura de Aprendizagem, Infraestrutura Profissional e Infraestrutura Audiovisual; (iv) renda, internet e quantidade de computadores.

Com isso, os recursos de titulação e infraestrutura geral são os que melhor poderiam ser aproveitados no nível escolar, seguidos das variáveis de contexto socioeconômico no nível regional. No nível individual, existem folgas, contudo, são as menores do conjunto de dados. Essas diferenças no uso dos insumos indicam que os IFs são suscetíveis a diferenças regionais e que as unidades que não são eficientes nem sempre aproveitam bem as condições favoráveis de contexto regional.

Gráfico 27 - Percentuais de folgas dos insumos do Modelo DEA geral e DEA espacial com dados regionais para 208 IFs com eficiência menor que 100% (2019).



Fonte: Elaborado pelo autor (2023).

No nível escolar, conforme a tabela 7, os IFs não conseguem a eficiência porque apresentam folgas elevadas para esse conjunto de indicadores. O indicador de titulação poderia ser mais bem aproveitado entre as unidades, por exemplo, a região Centro-Oeste apresenta a maior folga com -10,37%, próxima às regiões Nordeste, Sul e Sudeste. Por outro lado, a região Norte evidencia uma folga de apenas -1,44% nesse indicador, o que pode estar relacionado à escassez desse recurso e menor quantidade de docentes qualificados como mestres e doutores, fazendo com que a folga seja menor.

De forma geral, no nível escolar, as regiões se equivalem quanto à utilização dos recursos de capital humano e capital físico, com uma pequena diferença no conjunto de indicadores para a região Norte, a qual tem na Infraestrutura Geral (-10,86%) e na Profissional (-11,07%) as maiores folgas. Contrariamente, essa região evidencia menores folgas para as características docentes, com a titulação de -1,44% e a experiência de -3,54%. Esse é um dos desafios dos IFs, isto é, aproveitar de forma mais efetiva os seus recursos controláveis. A utilização da estrutura física como laboratórios, espaços pedagógicos, equipe técnica-pedagógica e docentes qualificados deve estar conectada com a aprendizagem dos estudantes nas unidades que não alcançaram a eficiência.

No nível regional, há maior disparidade entre as regiões quanto aos indicadores de nível estadual, a região Nordeste, com uma condição socioeconômica, nas últimas colocações, em relação às demais regiões, consegue menores folgas para esse conjunto de indicadores, situação semelhante a da região Norte. Enquanto o indicador de rendimento médio per capita evidencia uma folga de -2,15% para o Nordeste e de -2,6% para o Norte, esse mesmo indicador é de -12,37% para a região Centro-Oeste, -15,96% para o Sul e de -8,47% para o Sudeste. Com isso, os IFs localizados nas regiões Nordeste e Norte necessitam alcançar bons resultados frente a uma realidade mais desafiadora do ponto de vista socioeconômico no nível regional.

Por outro lado, os IFs das regiões Centro-Oeste, Sul e Sudeste evidenciam melhores condições de nível socioeconômico, no nível regional, o que não está sendo aproveitado de forma efetiva por algumas unidades. Isso pode estar relacionado à dificuldade de melhorar o desempenho dos estudantes num cenário mais favorável, em que a margem para o crescimento médio do desempenho no Enem seja menor quando comparado à média de crescimento para IFs em condições contrárias.

No nível individual, é possível verificar que a região Norte evidencia uma média de folga para o indicador de escolaridade da mãe de -14,15% bem superior às demais, impactando de forma significativa nos IFs de forma que não alcançaram a eficiência. De forma contrária, a região Nordeste evidencia uma folga de -0,17% para o indicador de renda, bem diferente das outras regiões analisadas, indicando que o fato de receber estudantes em condições de menor renda não impacta tão fortemente na eficiência de seus IFs. De forma geral, ao acrescentar os indicadores de nível estadual, há uma diminuição nos indicadores de nível individual, com o Sudeste apresentando as menores folgas.

Tabela 7 - Distribuição percentual das médias das folgas por região em relação aos indicadores de nível escolar, regional e individual dos Institutos Federais (IFs) ineficientes no ano de 2019 – Modelo DEA Espacial.

Indicadores	Regiões				
	Centro-Oeste	Norte	Nordeste	Sul	Sudeste
Nível Escolar - Soma	-41.5	-41.11	-47.06	-46.62	-43.07
Titulação	-10.37	-1.44	-8.13	-13.46	-10.09
Experiência	-2.93	-3.54	-8.87	-4.35	-5.32
Infraestrutura Geral	-8.58	-10.86	-8.3	-11.43	-8.94
Infraestrutura de Aprendizado	-6.91	-9.26	-6.79	-5.3	-5.75
Infraestrutura Audiovisual	-4.37	-4.94	-4.27	-6.78	-8.94
Infraestrutura Profissional	-8.34	-11.07	-10.7	-5.3	-4.03
Nível Regional - Soma	-48.16	-28.1	-20.3	-48.94	-43.85
Ideb anos finais rede Pública	-9.61	-7.71	-9.01	-9.63	-9.07
Ideb anos finais rede privada	-7.33	-8.92	-0.48	-13.89	-8.1
Ideb anos finais rede estadual	-11.98	-7.34	-7.18	-6.59	-8.34
Rendimento Médio Per Capita	-12.37	-2.6	-2.15	-15.96	-8.47
Percentual de Residências esgoto	-6.87	-1.53	-1.48	-2.87	-9.87
Nível Individual - Soma	-12.87	-21.51	-16.01	-17.94	-10.47
Renda	-3.51	-3.57	-0.17	-5.88	-2.77
Escolaridade da Mãe	-5.45	-14.15	-7.1	-3.33	-3.66
Quantidade de Computadores	-1.88	-1.22	-1.79	-5.82	-2.2
Acesso à Internet	-2.03	-2.57	-6.95	-2.91	-1.84

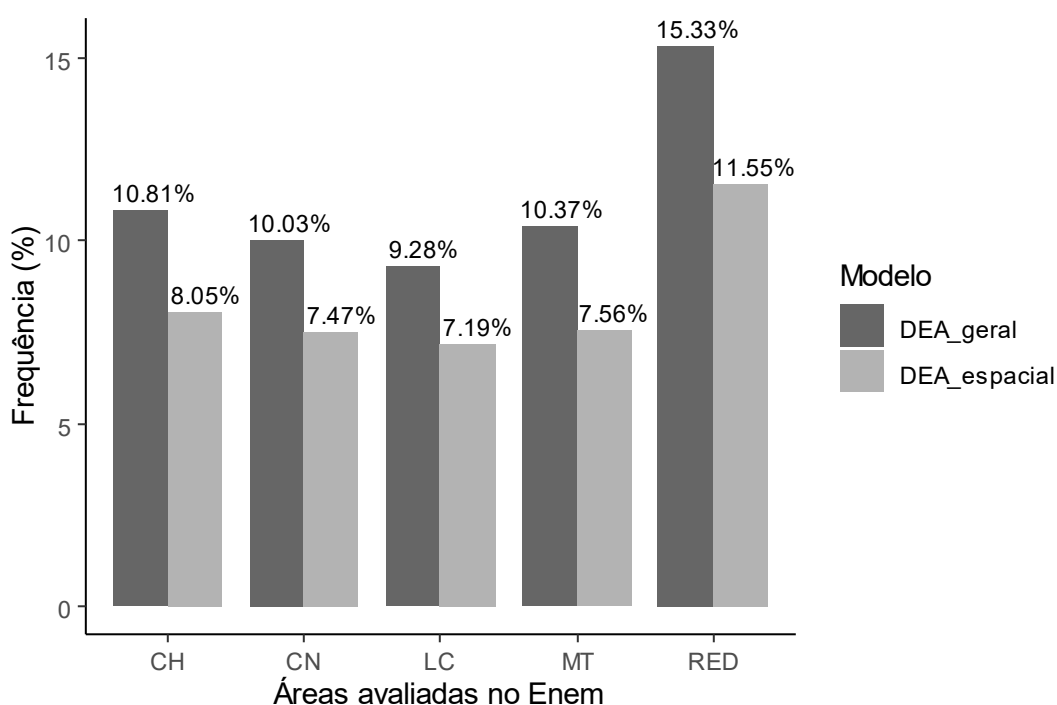
Fonte: Elaborada pelo autor (2023).

Reforçando os achados da avaliação dos insumos, os potenciais de melhoria das áreas avaliadas no Enem também são menores para o modelo DEA espacial, no entanto, mantendo-se as proporções do modelo DEA geral, conforme gráfico 28. A

área de Redação continua com o maior percentual de melhoria (11,55%), enquanto o segundo maior percentual corresponde à área de Ciências Humanas com 8,05%. Como as folgas foram menores, indicando menor quantidade de insumo para um nível eficiente, os potenciais de melhoria também o são.

Com a redução dos potenciais de melhoria, há mais uma evidência de que mesmo os IFs que não alcançaram a eficiência conseguiram diminuir o impacto em seus insumos com a inclusão de variáveis de contexto regional. As diferenças regionais impactam as unidades de forma que há um ganho de eficiência no consumo dos insumos. Num contexto de maior vulnerabilidade social, o desafio é elevar a nota média no Enem para além do esperado para aquele conjunto de estudantes, de forma contrária, num contexto favorável de nível socioeconômico no nível dos estados, o desafio é no mínimo manter e avançar no desempenho dos estudantes.

Gráfico 28 - Distribuição dos potenciais de melhoria das áreas avaliadas no Enem de acordo com o Modelo DEA geral e DEA espacial com dados regionais para 208 IFs (2019) com eficiência menor que 100.



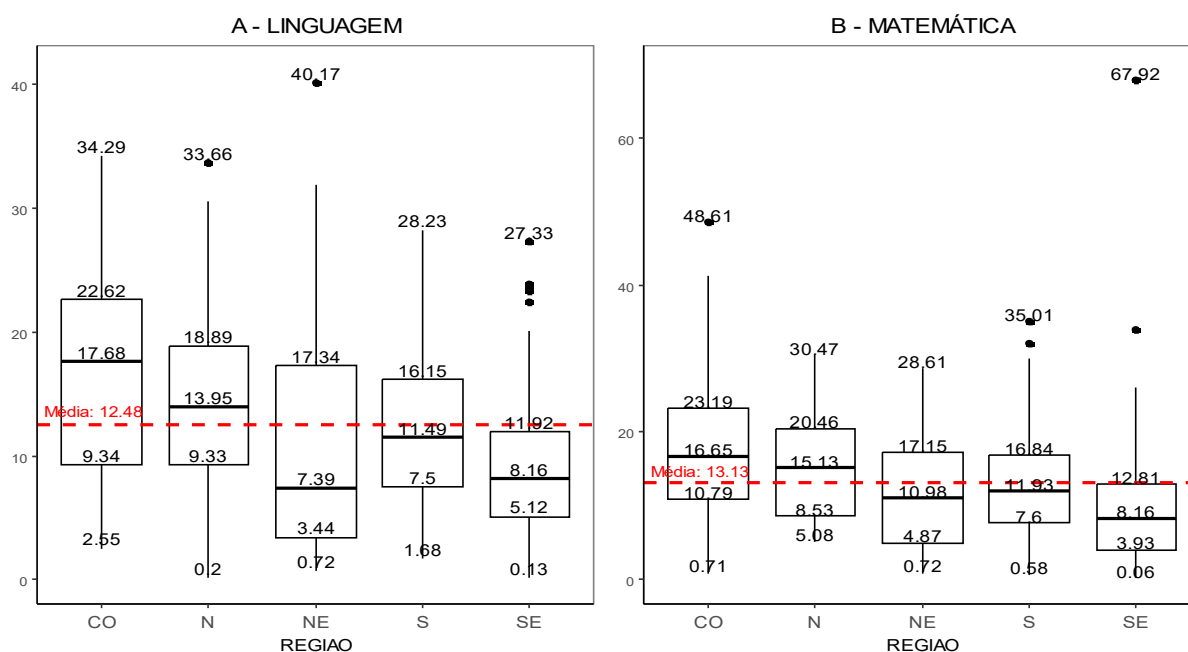
Fonte: Elaborado pelo autor (2023).

Ao se observar a distribuição dos potenciais de melhoria das áreas de Linguagem e Matemática, nas regiões do país, vê-se bastante disparidade entre os

IFs nas localidades. A partir do gráfico 29, é possível verificar que a região Centro-Oeste, além de ter vários IFs com disparidade elevada de potenciais de melhoria, apresenta a maior mediana com 17,68%, com isso, a metade das unidades poderia melhorar suas notas variando no intervalo de 17,68% a 34,29%. A região Nordeste evidencia uma mediana de 7,39% para a área de Linguagem, com a metade de seus IFs com potencial até esse limite.

Para a área de Matemática, a distribuição entre os IFs nas regiões é semelhante, com a menor mediana para a região Sudeste de 8,16%, com 75% dos IFs abaixo da média geral de 13,13%, seguindo-se a região Nordeste com mediana de 10,98%. Os potenciais de melhoria, dentro das regiões, apesar de serem menores quando comparados com o modelo DEA geral, ainda evidenciam IFs com potenciais superiores, chegando a 67,62%, indicando que as diferenças no acesso a insumos e na sua aplicação fazem com que várias unidades sejam significativamente impactadas nos resultados.

Gráfico 29 - Distribuição dos potenciais de melhoria das áreas de Linguagem e Matemática por região de acordo com o modelo DEA espacial para 208 IFs (2019) com eficiência menor que 100.



Fonte: Elaborado pelo autor (2023).

Diante desses resultados, ao se avaliar a eficiência com a inclusão de variáveis de contexto socioeconômico no nível dos estados, ocorreu um ganho significativo no

número de unidades eficientes, acréscimo de 50,89% com a entrada de vários IFs que antes eram ineficientes. Os IFs das regiões Norte, com acréscimo de 86%, Nordeste com 65%, e Sul com 65% foram os que mais se beneficiaram com a mudança. Principalmente, as regiões Norte e Nordeste se sobressaíram, tendo em vista que, apesar de se caracterizarem com menores indicadores socioeconômicos no nível regional, conseguiram aumentar significativamente o número de IFs eficientes.

Esses resultados apontam que, ao se incluir indicadores de vida da população, de seu contexto social, a curva de eficiência muda com o peso dessas variáveis. Além disso, ao se avaliar os insumos dos 208 IFs que não alcançaram a eficiência, evidencia-se que houve uma diminuição nas folgas de todos os indicadores quando comparados com o modelo DEA geral. No modelo DEA espacial, a configuração dos insumos se alterou, com os indicadores de titulação, infraestrutura geral e de nível estadual com as folgas mais elevadas do conjunto de insumos.

Explicita-se que os IFs são suscetíveis não somente ao nível socioeconômico dos estudantes, mas também a diferenças regionais. Um campus localizado numa região com melhores indicadores de renda média per capita estará diante de um cenário mais propício ao desenvolvimento de suas atividades, contudo, não quer dizer que haverá necessariamente melhor desempenho de seus estudantes. A exemplo disso, as regiões que mais se beneficiaram com a inclusão dos indicadores socioeconômicos dos estados (Norte e Nordeste) são as aquelas com os menores indicadores das condições de vida da população.

Com isso, pode ocorrer de os IFs com menores indicadores, nesse nível, alocarem de forma eficiente os seus recursos e terem mais espaço para avançar no desempenho dos estudantes do que os IFs em melhores condições. Portanto, a partir dos achados nesta seção, fica evidente a importância de se considerar fatores regionais na avaliação de eficiência dos IFs.

4.4 Síntese da análise de eficiência

A análise de eficiência dos Institutos Federais de Educação, Ciência e Tecnologia (IFs), em âmbito nacional, considerou a estrutura produtiva levando em conta os insumos de capital humano, insumos de capital físico, insumos de nível socioeconômico no nível dos estudantes e insumos de nível socioeconômico no nível dos estados, tendo como produto o desempenho no Enem. A partir da Análise DEA,

alguns pontos foram evidenciados como importantes na eficiência a partir da compreensão dos insumos que compõem a função de produção educacional dos IFs, como o impacto do nível socioeconômico dos estudantes, das diferenças regionais, dos recursos de capital humano e capital físico no nível escolar e da heterogeneidade da eficiência entre as regiões.

Os resultados permitem compreender a função de produção educacional dos IFs que foi desenvolvida, neste trabalho, de forma semelhante à função especificada por Albernaz, Ferreira e Franco (2002), no entanto, com foco nas características das famílias no nível individual dos estudantes; características dos professores e recursos escolares, no nível escolar; e das características de nível regional. Com isso, foram incluídas também as características dos professores dentro do ambiente escolar, porque há interferência das instituições de ensino nas políticas de qualificação.

No que tange às características socioeconômicas, no nível dos estudantes, há variação significativa desse insumo entre as regiões do país, impactando diretamente no desempenho dos alunos. Com indicadores bem abaixo, as regiões Norte e Nordeste divergem principalmente em termos de renda das demais regiões do país.

Por um lado, há o desafio das instituições de ensino de lidar com menores níveis socioeconômicos, em que os recursos escolares são mais exigidos. Nesse cenário, professores qualificados e, principalmente, experientes, somados a um ambiente escolar com boa infraestrutura física e pedagógica, são fundamentais para o aprendizado.

Por outro lado, há o desafio das instituições de ensino de tratarem com estudantes com bom *background* familiar. Nesse cenário, uma das metas é manter no mínimo o desempenho médio esperado, para, a partir disso, alcançar melhores resultados, o que também requer professores com bom nível de qualificação e experiência e recursos escolares a contento. Todavia, a margem de crescimento no desempenho para esse cenário tende a ser menor do que num cenário com baixo nível socioeconômico.

Outro ponto quanto ao nível socioeconômico está relacionado ao fato de as diferenças entre as regiões Norte e Nordeste serem significativas em relação às demais. Isso indica que existem peculiaridades no contexto regional que são inerentes a cada população, as quais acabam exercendo influência na atuação das instituições de ensino. Nesse sentido, o *background* familiar está inserido em um contexto

populacional que é influenciado pelas características regionais que impacta nas instituições de ensino.

Esse contexto foi evidenciado ao se incluir indicadores socioeconômicos no nível dos estados, o que fez com que houvesse um aumento no número de IFs eficientes, especialmente, para as regiões Norte, Nordeste e Sul. Há assim o efeito das condições regionais na capacidade dos IFs em lidar com os seus resultados. Os IFs são afetados de modo a potencializar seus resultados, recursos que antes não eram aproveitados na sua totalidade, quando inserido o contexto regional, passaram a ser utilizados, tornando a unidade eficiente. Essa análise torna-se significativa porque leva em consideração tanto o *background* familiar do estudante como as características da população.

Essas evidências são acentuadas ao se analisar as folgas nos insumos dos IFs das unidades que não alcançaram a eficiência, após a inclusão dos indicadores regionais. Todas as folgas foram menores, indicando melhor uso dos recursos. Além disso, o conjunto de indicadores regionais, ao lado dos indicadores de infraestrutura geral e da titulação, apresentaram as maiores folgas. Consolida-se assim as dificuldades desse conjunto de IFs em alcançar a eficiência, principalmente, porque a realidade regional teve impacto significativo. Isso pode estar associado a IFs tanto com boas condições regionais quanto em condições regionais relativamente inferiores.

No que tange aos insumos no nível escolar, a eficiência das instituições de ensino é impactada pela diferença de recursos e pela forma como são aplicados. A titulação e a experiência docentes são menores para Norte e Nordeste, fazendo com que os IFs dessas regiões tenham uma composição de capital humano diferente das demais. A maneira como são alocados os docentes também pode fazer diferença, tendo em vista que há nos IFs a possibilidade de verticalização do ensino, muitos dos docentes mestres e doutores podem assim não estar alocados em cursos técnicos de nível médio ou mesmo acumularem a docência em outros níveis de ensino como a graduação e a pós-graduação.

No entanto, o indicador de experiência foi evidenciado com menor folga do que a titulação. Nesse sentido, apesar de os docentes mais qualificados poder estarem em outros níveis de ensino ou em outras atividades, não é suficiente um IF apenas com docentes doutores do ponto de vista do desempenho, a experiência também é fundamental para uma unidade mais eficiente. Ela consiste em uma das

características de qualidade do professor que é significativa no desempenho, sendo mais importante ainda para estudantes com baixo nível de habilidades (ROCKOFF, 2004; RAVKIN; HANUSHEK; KAIN, 2005; AARONSON; BARROW; SANDER, 2007).

Quanto aos recursos de capital físico, não há diferença significativa entre as regiões como o indicador de titulação e os indicadores de nível regional. A infraestrutura geral, um indicador básico, está presente para a maioria das unidades distribuídas entre as regiões, bem como os recursos de infraestrutura audiovisual, infraestrutura de aprendizado e infraestrutura profissional. A sua presença, para a maioria das unidades, significa que nem todos estão aproveitando bem esses recursos, o que é mais evidente para o indicador de infraestrutura geral, gerando ineficiência.

Embora o indicador de infraestrutura geral não apresente variabilidade significativa entre as regiões, há variabilidade na sua aplicação, o que pode estar relacionado ao fato de as unidades terem diferenças de desempenho mantida a equivalência desse recurso. Há documentado, na literatura, a dificuldade de se identificar a associação de recursos dessa natureza com o desempenho dos estudantes, pois nem sempre a correlação é positiva (HANUSHEK, 1995). Apesar disso, o resultado determina que pode se elevar o desempenho frente ao nível de recurso disponível para se alcançar a eficiência.

Essas diferenças na aplicação dos recursos de capital físico podem estar associadas ao processo de gestão dos IFs, tendo em vista que é mais acessível recursos de ordem orçamentário-financeira do que recursos de pessoal, pois a estruturação de um laboratório, por exemplo, pode ser realizada mais rapidamente do que a contratação de servidores. Já os recursos destinados às políticas de contratação de pessoal são mais escassos, nem sempre há o profissional para os ambientes adequados no tempo exigido e com a qualificação necessária.

Se há o profissional, pode ocorrer de as folgas estarem associadas ao processo de ensino, pois um ambiente pedagógico, como um laboratório coordenado por um técnico da área, pode não estar necessariamente no planejamento docente ou mesmo estar sendo subutilizado. Além disso, pode ocorrer de os profissionais envolvidos nesse processo de utilização da infraestrutura das instituições de ensino não estarem qualificados como deveriam ou não estarem em processo constante de atualização. Tudo isso indica que há uma complexidade no processo de gestão quando estão

envolvidos os recursos de capital físico, intrinsicamente associados aos recursos de capital humano.

A análise DEA ainda permite apontar que há forte variação de eficiência entre as regiões do país, com realce para a região Nordeste com mais de 50% de suas unidades eficientes, o que foi comprovado com a análise do modelo DEA espacial, com acréscimo de 65%. Esse é um dos principais resultados deste estudo, tendo em vista que essa região possui o menor nível de renda e escolaridade da mãe no nível individual e os menores indicadores socioeconômicos no nível regional combinados com baixo nível relativo de capital humano, consistindo, portanto, em IFs que conseguem superar o contexto e obter excelentes resultados de seus estudantes. O achado é ainda mais significativo porque supera o número de IFs eficientes das regiões mais ricas e com estudantes com melhor *background* familiar.

A região Norte também evidencia satisfatória proporção de unidades eficientes confirmada com o aumento de IFs eficientes com a inclusão de indicadores de contexto regional, com acréscimo de 86%, contudo, com elevada dispersão entre suas unidades, com o maior percentual no menor intervalo do conjunto de dados. Esses resultados acabam por confirmar as dificuldades inerentes ao contexto regional, como permanência de docentes com maior titulação e experiência, infraestrutura de acesso das famílias a serviços de internet, nível socioeconômico dos estudantes e contexto regional.

Depreende-se, assim, a importância de se considerar os vários fatores na análise de eficiência dos IFs, principalmente, o melhor aproveitamento da infraestrutura geral e da titulação docente, além da relevância do professor na qualidade do ensino a partir da experiência. Além disso, há evidências robustas da eficácia da atuação dos IFs a partir dos resultados das unidades localizadas na região Norte e Nordeste, em que a maioria dos IFs são eficientes mesmo diante de um contexto com diferença em seus insumos de capital humano, nível socioeconômico dos estudantes e condições da população no nível regional.

5 CONCLUSÃO

Esta pesquisa teve como objetivo geral avaliar a contribuição dos insumos da função de produção educacional para o desempenho acadêmico dos estudantes no Exame Nacional do Ensino Médio (Enem), tendo como tese que tanto os recursos de capital físico quanto os recursos de capital humano são determinantes do desempenho acadêmico, porém, que os recursos de capital humano são mais relevantes para o desempenho dos estudantes.

Ao se avaliar os componentes da função de produção educacional, evidenciou-se que há diferença na aplicação de recursos de capital humano entre os Institutos Federais (IFs). Há desigualdade na distribuição de professores com maior titulação entre as regiões do país, em maiores proporções para a região Nordeste e a região Norte, que ainda apresenta IFs com menor experiência docente, diferentemente das demais regiões, com níveis mais elevados de capital humano. Por outro lado, não há diferença significativa na aplicação dos recursos de capital físico entre os IFs por regiões, demonstrando que as disparidades estão associadas a fatores não regionalizados.

As disparidades regionais são também acentuadas no nível socioeconômico dos estudantes, que, nesta tese, foram representadas por variáveis como renda, escolaridade da mãe, número de computador por família e acesso à internet. As regiões Norte e Nordeste acumulam também os IFs com o menor nível socioeconômico, a exemplo das disparidades de recursos de capital humano. Além disso, há diferença entre os IFs localizados na zona rural e na zona urbana, com as unidades da zona rural com menor desempenho no Enem e menor nível socioeconômico, entretanto, com maior nível de capital humano.

Tem-se assim que os insumos de capital humano e socioeconômico dos estudantes estão sujeitos a variações no nível das regiões, com maior escassez sobretudo para os IFs do Norte e Nordeste, os quais combinam menor desempenho. Isso significa que, para esses IFs, há um desafio duplo, atuar com estudantes com menor *background* familiar e com docentes com menor nível de titulação e experiência.

Através da análise de eficiência, por meio do modelo DEA geral, foi possível evidenciar que apenas um terço dos IFs são eficientes (36,3%). Ao se incluir variáveis de contexto regional, modelo DEA espacial, passou-se a ter 54,78% de IFs eficientes,

com maior proporção de aumento para a região Norte, seguida das regiões Nordeste e Sul. Esse resultado permite alcançar o objetivo específico da investigação de determinar os componentes da função de produção escolar, a partir de insumos de capital humano, capital físico, socioeconômico dos estudantes e, principalmente, insumos regionais como determinantes na eficiência dos IFs que impactam no desempenho.

Considerando a análise de eficiência, é possível verificar que, apesar das diferenças regionais entre os IFs, a região Nordeste, mesmo com os menores indicadores socioeconômicos no nível dos estudantes e dos estados, apresentou a maior proporção de unidades eficientes. Essa situação é similar a dos IFs da região Norte, que, além disso, atuam com menor nível de titulação e experiência docente e, ainda assim, conseguem a segunda maior proporção de unidades eficientes.

Dessa forma, conclui-se que os IFs dessas regiões conseguem alocar seus recursos de forma a mitigar os efeitos de contexto no nível dos estudantes e no nível regional. Contudo, o ideal seria que os IFs dessas regiões contassem com o mesmo nível de recursos de capital humano das demais regiões para assim poderem aumentar o desempenho de seus estudantes dadas as condições socioeconômicas em que estão inseridos.

Ao se avaliar as folgas dos IFs ineficientes, é possível verificar que essas unidades apresentam folgas em todos os insumos, contudo, de forma mais intensa para os insumos de infraestrutura geral e titulação. Isso significa que muitas unidades não estão alocando esses recursos de forma a impactar no desempenho. Pode ocorrer de unidades já consolidadas nesses recursos estarem com baixo desempenho, em boa medida porque a infraestrutura geral está presente para a maioria das unidades, e pelo fato de docentes com maior titulação estarem alocados em atividades diferentes do ensino técnico de nível médio.

Ao serem incluídas as variáveis de contexto regional dos estados, houve uma mudança na estrutura dos insumos em que se mantiveram as folgas de infraestrutura geral e titulação com os maiores percentuais ao lado dos insumos dos estados. Verificou-se que os IFs sofrem o efeito do contexto regional de forma a potencializar os seus resultados, com o aumento de unidades eficientes e mudança na composição dos insumos, os quais passaram a contar com menores folgas.

Esses resultados expressam o objetivo específico de determinar as folgas e restrições dos determinantes da eficiência escolar na função de produção. As maiores

folgas correspondem aos insumos de infraestrutura geral e titulação no nível escolar. Por não serem utilizadas de modo a aumentar a eficiência acabam sendo limitadores. As folgas dos insumos no nível regional também são elevadas, contudo, esse recurso atua minimizando as folgas dos demais insumos, o que torna mais IFs eficientes. Isso porque as características gerais da população amenizam o peso das características de *background* familiar e de menor capital humano.

Diante disso, é possível retomar o problema de pesquisa e apontar que, no nível escolar, o insumo de titulação, quando avaliado pela sua distribuição, é o mais escasso para as regiões Nordeste e Norte. Ainda cabe salientar que a experiência para essa região difere dos recursos de capital físico que seguem um padrão mais uniforme entre as regiões. Além disso, evidenciou-se que a escassez dos insumos de capital humano interfere na eficiência dos IFs de forma desigual entre as regiões impactando relativamente em mais unidades eficientes para as regiões Norte e Nordeste e, assim, reforçando o resultado dessas regiões, dado o contexto socioeconômico.

Após os resultados, levando em consideração a tese desta pesquisa, não há como precisar se o capital humano é mais relevante do que o capital físico para explicar o desempenho, porque ambos os insumos contribuem tanto para a eficiência dos IFs, quanto para a sua ineficiência, quando não são utilizados na sua plenitude. O que se pode afirmar é que, dentre esses insumos, os recursos de infraestrutura geral e titulação, que apresentaram as maiores folgas, ao não serem utilizados na sua totalidade, reduzem a eficiência dos IFs de maneira mais frequente do que os demais indicadores. Portanto, os dois insumos, acabam se tornando uma oportunidade de melhoria para os IFs ineficientes alcançarem a eficiência com o aumento do desempenho.

Por fim, é evidente que a eficiência como resultado do uso adequado dos recursos para aumentar o desempenho não está associada ao uso isolado de um recurso. Pelo contrário, existem diversas combinações que podem gerar resultados superiores. Além disso, ao considerar as variações de eficiência entre as regiões do país dado o contexto socioeconômico, mostra-se relevante avançar em pesquisas que busquem identificar quais fatores qualitativos no nível escolar estão associados a eficiência com foco no desempenho. Nesse sentido, recomenda-se, principalmente, aqueles relacionados à atuação do professor.

REFERÊNCIAS

- AARONSON, D.; BARROW, L.; SANDER, W. **Teachers and student achievement in the Chicago public high schools**. Journal of labor Economics, [s.l.] v. 25, n. 1, p. 95-135, 2007.
- AGÜERO, J. M.; BELECHE, T. **Test-Mex**: Estimating the effects of school year length on student performance in Mexico. Mexico, Journal of Development Economics, v. 103, p.353-361, 2013.
- AIGNER, D. J.; CHU, S. **On estimating the industry production function**. The American Economic Review, v. 58, n. 4, p. 826-839, 1968.
- ALBERNAZ, A.; FERREIRA, F. H.G.; FRANCO, C. **Qualidade e equidade na educação fundamental brasileira**. Rio de Janeiro: Pesquisa e planejamento econômico, v. 32, n. 3, 2002.
- ALBUQUERQUE, M. C. C. **Microeconomia**. São Paulo: McGraw-Hill, 1986.
- ALTINOK, N. **A Macroeconomic Estimation of the Education Production Function**. 2007.
- AMMERMÜLLER, A.; HEIJKE, H.; WÖßMANN, L. Schooling quality in Eastern Europe: educational production during transition. **Economics of Education Review**, v. 24, n. 5, p. 579-599, 2005.
- ARAÚJO, A. J. N.; CHEIN, F.; PINTO, C. C. X. Ensino profissionalizante, desempenho escolar e inserção produtiva: uma análise com dados do ENEM. **Pesquisa e Planejamento Econômico**, Brasília, v. 48, n. 1, 2018.
- ASSUNCAO, M. V. D.; ARAUJO, A. G.; ALMEIDA, M. R. O background familiar e sua influência no acesso ao Ensino Técnico Profissional. **Revista Administração Pública**, Rio de Janeiro, v. 53, n. 3, p.542-559, 2019.
- BACOLOD, M. P.; TOBIAS, J. L. Schools, school quality and achievement growth: Evidence from the Philippines. **Economics of education review**, v. 25, n. 6, p. 619-632, 2006.
- BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. **Management science**, v. 30, n. 9, p. 1078-1092, 1984.
- BARBOSA, M. E. F.; FERNANDES, C. A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em Matemática dos alunos da 4ª série. Promoção, ciclos e avaliação educacional. Porto Alegre: **ArtMed**, p.155-172, 2001.
- BARROS, R. P. et al. **Determinantes do desempenho educacional no Brasil**. IPEA, 2001.
- BARTHOLO, T. L.; COSTA, M. Evidence of a school composition effect in Rio de

Janeiro public schools. **Ensaio: Avaliação e políticas públicas de Educação**, Rio de Janeiro, v. 24, n. 92, p. 498-521, 2016.

BASSETTO, C. F. Background familiar e desempenho escolar: uma abordagem com variáveis binárias a partir dos resultados do Saesp. **Revista Brasileira de Estudos de População**, v. 36, 2019.

BECKER, B. Social disparities in children's vocabulary in early childhood. Does pre-school education help to close the gap? **British Journal of Sociology**, [s. l.], v. 62, n. 1, p. 69-88, 2011.

BENEVIDES, A. A.; SOARES, A. B. Diferencial de desempenho de alunos das escolas militares: o caso das escolas públicas do Ceará. **Nova economia**, Belo Horizonte, v. 30, n. 1, p. 317-343, 2020.

BERNAL, P; MITTAG, N.; QURESHI, J. A. Estimating effects of school quality using multiple proxies. **Labour Economics**, v. 39, p. 1-10, 2016.

BESANKO, D.; BRAEUTIGAM, R. R. / **Microeconomics**. 4th ed. New York. John Wiley & Sons, 2010.

BEUCHERT, L. et al. The short-term effects of school consolidation on student achievement: Evidence of disruption? **Economics of Education Review**, v. 65, p. 31-47, 2018.

BLACK, D. A.; SMITH, J. A. Estimating the returns to college quality with multiple proxies for quality. **Journal of labor Economics**, v. 24, n. 3, p. 701-728, 2006.

BLANCO, E. Efectos escolares sobre los aprendizajes en México: una perspectiva centrada en la interacción escuela-entorno. **Papeles de población**, Toluca, v. 17, n. 69, p. 219-256, 2011.

BOWLES, S. Towards an educational production function. **National Bureau of Economic Research**, 1970.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **A redação do Enem 2022: cartilha do participante**. Brasília, 2022.

BRASIL. **Decreto n.º 7.566, DE 23 de setembro de 1909**. Dispõe sobre a criação nas capitais dos Estados das Escolas de Aprendizes Artífices para o ensino profissional primário e gratuito.

BRASIL. **Decreto nº 5.154 de 23 de julho de 2004**. Regulamenta o § 2º do art. 36 e os arts. 39 a 41 da Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, e dá outras providências.

BRASIL. **Decreto nº 7.234, de 19 de julho de 2010**. Dispõe sobre o Programa Nacional de Assistência Estudantil-PNAES.

BRASIL. **Lei nº 11.892, de 29 de dezembro de 2008**. Institui a Rede Federal de Educação Profissional, Científica e Tecnológica, cria os Institutos Federais de Educação, Ciência e Tecnologia, e dá outras providências.

- BRAULT, M.; JANOSZ, M.; ARCHAMBAULT, I. Effects of school composition and school climate on teacher expectations of students: A multilevel analysis. **Teaching and Teacher Education**, v. 44, p. 148-159, 2014.
- CANGUSSU, R. C.; SALVATO, M. A.; NAKABASHI, L. Uma análise do capital humano sobre o nível de renda dos estados brasileiros: MRW versus Mincer. **Estudos Econômicos**, São Paulo, v. 40, n. 1, p. 153-183, 2010.
- CARRASCO, A.; MARTÍN, E. S. Voucher system and school effectiveness: Reassessing school performance difference and parental choice decision-making. **Estudios de Economía**, [s. l.], v. 39, n. 2, p. 123–141, 2012.
- CASADO, F. L. Análise envoltória de dados: conceitos, metodologia e estudo da arte na educação superior. **Revista Sociais e Humanas**, v. 20, n. 1, p. 59-71, 2007.
- CASTELLÓ-CLIMENT, A.; HIDALGO-CABRILLANA, A. The role of educational quality and quantity in the process of economic development. **Economics of Education Review**, vol. 31, n. 4, 391–409, 2012.
- CECHELLA, G. A interiorização do ensino público federal e o desenvolvimento local. **Brazilian Journal of Development**, v. 5, n. 1, p. 618-630, 2019.
- CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European journal of operational research**, v. 2, n. 6, p. 429-444, 1978.
- CHOI, E. J.; MOON, H. R.; RIDDER, G. Estimation of an education production function under random assignment with selection. **American Economic Review**, v. 104, n. 5, p. 206-11, 2014.
- CHUDGAR, A.; QUIN, E. Relationship between Private Schooling and Achievement: Results from Rural and Urban India. **Economics of Education Review**, [s. l.], v. 31, n. 4, p. 376–390, 2012.
- COLEMAN, J. S. et al. **Equality of Educational Opportunity**. Washington, US Government Printing Office, 1966.
- COOPER, S. T.; COHN, E. Estimation of a frontier production function for the South Carolina educational process. **Economics of Education Review**, v. 16, n. 3, p. 313-327, 1997.
- COSTA, L. O.; ARRAES, R. A. E.; GUIMARÃES, D. B. Estabilidade dos professores e qualidade do ensino de escolas públicas. **Economia Aplicada**, v. 19, n. 2, p. 261-298, 2015.
- CURI, A. Z.; MENEZES FILHO, N. A. Mensalidade escolar, background familiar e os resultados do Exame Nacional do Ensino Médio (ENEM). **Pesquisa e planejamento econômico**, v. 43. n. 2, 2013.
- CURI, A. Z.; SOUZA, A. P. Medindo a qualidade das escolas: evidências para o Brasil. **Economia Aplicada**, Ribeirão Preto, v. 19, n. 3, p. 541-574, 2015.

DELPRATO, M.; CHUDGAR, A. Factors associated with private-public school performance: Analysis of TALIS-PISA link data. **International Journal of Educational Development**, v. 61, p. 155-172, 2018.

DEWEY, J.; HUSTED, T. A.; KENNY, L. W. The ineffectiveness of school inputs: a product of misspecification? **Economics of Education Review**, v. 19, n. 1, p.27- 45, 2000.

DRIESSEN, G.; AGIRDAG, O.; MERRY, M. S. The gross and net effects of primary school denomination on pupil performance. **Educational Review**, [s. l.], v. 68, n. 4, p. 466–480, 2016.

DUTRA, R. S. et al. Determinantes do desempenho educacional dos Institutos Federais do Brasil no Exame Nacional do Ensino Médio. **Educação e Pesquisa**. São Paulo, v. 45, 2019.

DUTRA, R. S. O que mudou no desempenho educacional dos Institutos Federais do Brasil? **Ensaio: avaliação e políticas públicas em educação**, Rio de Janeiro, v. 27, n. 104, p.631-653, 2019.

DUTRA, R. S.; COELHO, A. C. D.; DUTRA, G. B. M. Indicadores educacionais e proficiência no Enem: um estudo nos institutos federais do Brasil. **Revista Meta: Avaliação**, v. 11, n. 31, p. 124-153, 2019.

ESPINOSA, A. M. G. Estimating the education production function for cognitive and non-cognitive development of children in vietnam through structural equation modeling using young lives data base. **Institute of Education, University College London**, 2017.

FARREL, M. **The measurement of production efficiency**. **Journal of the Royal Statistical Society**, v. 120, n. 3, p. 253-290, 1957.

FELÍCIO, F.; FERNANDES, R. O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de São Paulo. **Anais do XXXIII Encontro Nacional de Economia**, 2005.

FERRÃO, M. E.; FERNANDES, C. O efeito-escola e a mudança-dá para mudar? Evidências da investigação Brasileira. **REICE: Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, v. 1, n. 1, p. 5, 2003.

FERRÃO, M. E.; LEITE, I. C.; BELTRÃO, K. I. **Introdução à modelagem multinível em Avaliação Educacional**. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística – IBGE, 2001.

FIGLIO, D. N. Functional form and the estimated effects of school resources. **Economics of education review**, v. 18, n. 2, p. 241-252, 1999.

FLORES, L. B. **Determinantes da Eficiência das Escolas Públicas no Exame Nacional do Ensino Médio (ENEM): Uma Análise de Dois Estágios**. Orientador: Roberto Tatiwa Ferreira. 2019. 53 f. Dissertação (Mestrado em Economia). Universidade Federal do Ceará, Fortaleza, 2019.

FRANCO, A. M. P.; MENEZES FILHO, N. A. Os determinantes do aprendizado com dados de um painel de escolas do SAEB. **Economia Aplicada**, v. 21, n. 3, p. 525-548, 2017.

FRANCO, C. et al. Qualidade e equidade em educação: reconsiderando o significado de "fatores intraescolares". **Ensaio: avaliação e políticas públicas em educação**, v. 15, n. 55, p. 277-298, 2007.

GALVAO, F. V. Gastos com salários e desempenho em matemática: uma análise baseada nas escolas municipais de SBC. **Educação e Pesquisa**, São Paulo, v. 44, 2018.

GAVIRIA, J.; MARTÍNEZ-ARIAS, R.; CASTRO, M. Un estudio multinivel sobre los factores de eficacia escolar en países en desarrollo: El caso de los recursos en Brasil. **Education Policy Analysis Archives**, v. 12, n. 20, p. 1-31, 2004.

GELMAN, A.; HILL, J. **Data analysis using regression and multilevel/hierarchical models**. Cambridge university press, 2006.

GLEWWE, P. W. et al. School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010. **National Bureau of Economic Research**, 2011.

GLEWWE, P.; KREMER, M. Schools, teachers, and education outcomes in developing countries. **Handbook of the Economics of Education**, v. 2, p. 945-1017, 2006.

GLICK, P.; SAHN, D. E. Cognitive skills among children in Senegal: Disentangling the roles of schooling and family background. **Economics of Education Review**, v. 28, n. 2, p. 178-188, 2009.

GREMAUD, A. P.; FELÍCIO, F.; BIONDI, R. L. **Indicador de efeito escola: uma metodologia para a identificação dos sucessos escolares a partir dos dados da prova Brasil**. Brasília: INEP/MEC, 2007 (Texto para Discussão, n. 27).

GRIEW, P. et al. The school effect on children's school time physical activity: the PEACH Project. **Preventive Medicine**, [s. l.], v. 51, n. 3/4, p. 282-286, 2010.

GYIMAH-BREMpong, K.; GYAPONG, A. O. Characteristics of education production functions: An application of canonical regression analysis. **Economics of Education Review**, v. 10, n. 1, p. 7-17, 1991.

HÆGELAND, T.; RAAUM, O.; SALVANES, K. G. Pennies from heaven? Using exogenous tax variation to identify effects of school resources on pupil achievement. **Economics of Education Review**, v. 31, n. 5, p. 601-614, 2012.

HAHN, Y.; WANG, L. C.; YANG, H. Does greater school autonomy make a difference? Evidence from a randomized natural experiment in South Korea. **Journal of Public Economics**, v. 161, p. 15-30, 2018.

HANUSHEK, E. A. Assessing the effects of school resources on student performance: An update. **Educational evaluation and policy analysis**, v. 19, n. 2,

p. 141-164, 1997.

HANUSHEK, E. A. Assessing the effects of school resources on student performance: An update. **Educational evaluation and policy analysis**, v. 19, n. 2, p. 141-164, 1997.

HANUSHEK, E. A. Interpreting recent research on schooling in developing countries. **The world bank research observer**, v. 10, n. 2, p. 227-246, 1995.

HANUSHEK, E. A. School resources. **Handbook of the Economics of Education**, v. 2, p. 865-908, 2006.

HANUSHEK, E. A. The economic value of higher teacher quality. **Economics of Education review**, v. 30, n. 3, p. 466-479, 2011.

HANUSHEK, E. A. The economics of schooling: Production and efficiency in public schools. **Journal of economic literature**, v. 24, n. 3, p. 1141-1177, 1986.

HANUSHEK, E. A. The failure of input-based schooling policies. **The economic journal**, v. 113, n. 485, p. 64-98, 2003.

HANUSHEK, E. A. The impact of differential expenditures on school performance. **Educational researcher**, v. 18, n. 4, p. 45-62, 1989.

HANUSHEK, E. A. Throwing money at schools. **Journal of policy analysis and management**, v. 1, n. 1, p. 19-41, 1981.

HANUSHEK, E. A. What matters for achievement: updating Coleman on the influence of families and schools. **Education Next**, v. 16, n. 2, p. 22-30, 2016.

HANUSHEK, E. A.; WÖßMANN, L. The role of education quality for economic growth. **World Bank policy research working paper**, n. 4122, 2007.

HANUSHEK, Eric A. Conceptual and empirical issues in the estimation of educational production functions. **Journal of human Resources**, p. 351-388, 1979.

HARBISON, R. W.; HANUSHEK, E. **A Educational performance of the poor: lessons from rural Northeast Brazil**. Oxford University Press, 1992.

HARBISON, R.; HANUSHEK, E. **Educational performance of the poor: lessons from rural Northeast Brazil**. The World Bank Group. 384 p. 1992.

HILL, H. C. The Coleman Report, 50 years on: What do we know about the role of schools in academic inequality? **The ANNALS of the American Academy of Political and Social Science**, v. 674, n. 1, p. 9-26, 2017.

HOBIN, E. P. et al. A multilevel examination of school and student characteristics associated with physical education class enrollment among high school students. **Journal of School Health**, [s. l.], v. 80, n. 9, p. 445-452, 2010.

HUMLUM, M. K.; SMITH, N. Long-term effects of school size on students' outcomes. **Economics of Education Review**, v. 45, p. 28-43, 2015.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Anuário estatístico do Brasil**. Departamento de Divulgação Estatística, IBGE, 2020.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Brasil no Pisa 2018** [recurso eletrônico]. – Brasília, 2020.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Brasil no PISA 2015: análises e reflexões sobre o desempenho dos estudantes brasileiros/OCDE-Organização para a Cooperação e Desenvolvimento Econômico**. São Paulo: Fundação Santillana, 2016.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP).2020. **Cartilha SAEB 2019**.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Exame Nacional do Ensino Médio (Enem)**.

JESUS, G. R.; LAROS, J. A. Eficácia escolar: regressão multinível com dados de avaliação em larga escala. **Avaliação psicológica**, Porto Alegre, v. 3, n. 2, p. 93-106, 2004.

JURADO, J. C. Z. Análisis multinivel del rendimiento escolar en matemáticas para cuarto grado de Educación Básica Primaria en Colombia. **Sociedade e Economia**, Cali, n. 25, p. 205-235, 2013.

JURADO, J. C. Z. Un estudio multinivel del rendimiento escolar en matemáticas para tercer grado de educación básica primaria en América Latina. **Sociedade e Economia**, Cali, n. 30, p. 91-120, 2016.

KIM, H.; RHEE, D. Toilets for education: Evidence from Kenya's primary school-level data. **International Journal of Educational Development**, v. 70, 2019.

KIM, Y. The effects of school choice on achievement gaps between private and public high schools: Evidence from the Seoul high school choice program. **International Journal of Educational Development**, v. 60, p. 25-32, 2018.

KITCHENHAM, B. Procedures for performing systematic reviews. Keele, UK, **Keele University**, v.33, n. 2004, p.1-26, 2004.

KUNZE, N. C. O surgimento da rede federal de educação profissional nos primórdios do regime republicano brasileiro. **Revista brasileira da educação profissional e tecnológica**, v. 2, n. 2, p. 8-24, 2009.

LAROS, J. A.; MARCIANO, J. L.; ANDRADE, J. M. Fatores associados ao desempenho escolar em português: um estudo multinível por regiões. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 20, n. 77, p. 623-646, 2012.

LEE, V. E. A necessidade dos dados longitudinais na identificação do efeito-escola. **Revista Brasileira de Estudos Pedagógicos**, v. 91, n. 229, 2010.

- LEON, G.; VALDIVIA, M. Inequality in school resources and academic achievement: Evidence from Peru. **International Journal of Educational Development**, v. 40, p. 71-84, 2015.
- LIMA, R. **Mercado de trabalho: o capital humano e a teoria da segmentação**. 1980.
- LINS, M. P. E.; ALMEIDA, B. F.; BARTHOLO JUNIOR, R. Avaliação de desempenho na pós-graduação utilizando a Análise Envoltória de Dados: o caso da Engenharia de Produção. **Revista Brasileira de Pós-graduação**, Brasília, n. 1, p. 41-56, 2004.
- LOCKHEED, M. E.; BRUNS, B. **School effects on achievement in secondary Mathematics and Portuguese in Brazil**. Washington, D. C.: The World Bank, 1990. (Working Paper, 525).
- MASCI, C. et al. Bivariate multilevel models for the analysis of mathematics and reading pupils' achievements. **Journal of Applied Statistics**, [s. l.], v. 44, n. 7, p. 1296–1317, 2017.
- MAS-COLELL, A.; WHINSTON, M. D.; GREEN, J. R. **Microeconomic theory**. New York: Oxford University Press, 1995.
- MCVICAR, D.; MOSCHION, J.; RYAN, C. Achievement effects from new peers: Who matters to whom? **Economics of Education Review**, v. 66, p. 154-166, 2018.
- MELLO, J.; C.; B.; S. *et al.* Curso de análise de envoltória de dados. **XXXVII Simpósio Brasileiro de Pesquisa Operacional**, p. 2520-2547, 2005.
- MELO, S. G.; MORAIS, A. Clima escolar como fator protetivo ao desempenho em condições socioeconômicas desfavoráveis. **Cadernos de Pesquisa**, São Paulo, v. 49, n. 172, p. 10-34, 2019.
- MENEZES-FILHO, N. A. **Os determinantes do desempenho escolar do Brasil**. 2007.
- MINAYA, V.; AGASISTI, T. Evaluating the Stability of School Performance Estimates over Time. **Fiscal Studies**, [s. l.], v. 40, n. 3, p. 401–425, 2019.
- MINCER, J. Changes in wage inequality, 1970-1990. **NBER Working Paper**, 1996.
- MINCER, J. The distribution of labor incomes: a survey with special reference to the human capital approach. **Journal of economic literature**, v. 8, n. 1, p. 1-26, 1970.
- MINISTÉRIO DA EDUCAÇÃO E DESPORTO (MEC). **Histórico da Rede Federal de Educação Profissional e Tecnológica**. Brasília, 2016.
- MINISTÉRIO DA EDUCAÇÃO E DESPORTO (MEC). **Manual para cálculo dos indicadores de gestão das Instituições da Rede Federal de Educação Profissional, Científica e Tecnológica: indicadores, definições, fórmulas de cálculo e critérios de agregação**. Brasília, 2016.
- MORAES, J.; DIAS, B. F. B.; MENEZES, D. T. de. Efeito escola na educação básica:

observações a partir da perspectiva Bourdieusiana. Universidade Federal da Paraíba. **Revista Temas em Educação**, v. 29, n. 1, 2020.

MOREIRA, N. P.; BENEDICTO, G. C. DE; CARVALHO, F. DE M. Discussão de alguns condicionantes da eficiência em universidades federais brasileiras a partir do Reuni. **Revista do Serviço Público**, v. 70, n. 3, 2019.

NASCIMENTO, M. M.; CAVALCANTI, C.; OSTERMANN, F. Dez anos de instituição da Rede Federal de Educação Profissional, Científica e Tecnológica: o papel social dos institutos federais. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 101, n. 257, p. 120-145, 2020.

OSWALD, Y.; BACKES-GELLNER, U. Learning for a bonus: How financial incentives interact with preferences. **Journal of Public Economics**, v. 118, p. 52-61, 2014.

OTRANTO, C. R. Criação e implantação dos institutos federais de educação, ciência e tecnologia–IFETS. **Revista Retta**, n. 1, p. 89-110, 2010.

PEÑA, Carlos Rosano. Um modelo de avaliação da eficiência da administração pública através do método análise envoltória de dados (DEA). **Revista de Administração Contemporânea**, Curitiba, v. 12, n.1, p.83-106, 2008

PINDYCK, R. S.; RUBNFIELD, D. L., **Microeconomia**. 8ª ed., São Paulo: Pearson, 2013.

PRITCHETT, L.; FILMER, D. What education production functions really show: a positive theory of education expenditures. **Economics of Education review**, v. 18, n. 2, p. 223-239, 1999.

RAUDENBUSH, S. W.; WILLMS, J. D. The estimation of school effects. **Journal of educational and behavioral statistics**, v. 20, n. 4, p. 307-335, 1995.

RIANI, J. L. R.; RIOS-NETO, E. L. G. Background familiar versus perfil escolar do município: qual possui maior impacto no resultado educacional dos alunos brasileiros? **Revista Brasileira de Estudos de População**, v. 25, n. 2, p. 251-269, 2008.

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

ROCHA, A. B.; FUNCHAL, B. Mais recursos, melhores resultados? As relações entre custos escolares diretos e desempenho no Ensino Médio. **Revista Administração Pública**, Rio de Janeiro, v. 53, n. 2, p. 291-309, 2019.

ROCHA, F. B. N. et al. Análise dos microdados de matemática do ENEM de 2017–2019 do Nordeste. **Research, Society and Development**, v. 11, n. 10, 2022.

ROCKOFF, J. E. The impact of individual teachers on student achievement: Evidence from panel data. **American economic review**, v. 94, n. 2, p. 247-252, 2004.

RODRIGUES, L. O. **Ensaio sobre o diferencial de desempenho escolar entre**

alunos de escolas rurais e urbanas no Brasil. Dissertação (Mestrado em Economia Rural) – Centro de Ciências Agrárias, Universidade Federal do Ceará, Fortaleza, Ceará, 2017.

RODRIGUEZ, M.; CORREA, B. Impacto del contexto municipal sobre el desempeño académico individual. **Lecturas de Economía**, Medellín, n. 90, p. 159-193, 2019.

RODRÍGUEZ-JIMÉNEZ, O. R.; MURILLO-TORRECILLA, F. J. Estimación del efecto escuela para Colombia. Magis. **Revista Internacional de Investigación en Educación**, v. 3, n. 6, p. 299-316, 2011.

SAKELLARIOU, C. Private or public school advantage? Evidence from 40 countries using PISA 2012-Mathematics. **Applied Economics**, [s. l.], v. 49, n. 29, p. 2875–2892, 2017.

SALVATO, M. A.; FERREIRA, P. C. G.; DUARTE, A. J. M. O impacto da escolaridade sobre a distribuição de renda. **Estudos Econômicos**. São Paulo, v. 40, n. 4, p. 753-791, 2010.

SAMPAIO, B. et al. Desempenho no vestibular, background familiar e evasão: evidências da UFPE. **Economia Aplicada**, Ribeirão Preto, v. 15, n. 2, p. 287-309, 2011.

SANSANI, S. The effects of school quality on long-term health. **Economics of Education Review**, v. 30, n. 6, p. 1320-1333, 2011.

SCHMIDT, M. A. **Os Institutos de Educação, Ciência e Tecnologia: um estudo da expansão da Rede Federal de Educação Profissional e Tecnológica.** Dissertação (Mestrado em Educação) – Programa de Pós-graduação em Educação, Universidade de Passo Fundo, Passo Fundo, 2010.

SCHMITT, J.; WADSWORTH, J. Is there an impact of household computer ownership on children's educational attainment in Britain? **Economics of Education review**, v. 25, n. 6, p. 659-673, 2006.

SEABRA, T.; CARVALHO, H.; ÁVILA, P. The effect of the ethnic composition of schools on primary school maths results of pupils of immigrant origin. **Portuguese Journal of Social Science**, [s. l.], v. 18, n. 1, p. 9–26, 2019.

SHEPHARD, R. W. **Theory of cost and production functions.** Princeton University Press, 1970.

SOARES, J. F. Avaliação da qualidade da educação escolar brasileira. O sociólogo e as políticas públicas. In: Schwartzman, L. F. (Org.). **Ensaio em homenagem a Simon Schwartzman.** Rio de Janeiro: Editora FGV, p. 215-242, 2009.

SOARES, J. F. Melhoria do desempenho cognitivo dos alunos do ensino fundamental. **Cadernos de pesquisa**, v. 37, p. 135-160, 2007.

SOARES, J. F. O efeito da escola no desempenho cognitivo de seus alunos. REICE: **Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación**, v. 2, n. 2, p. 6, 2004.

SOARES, J. F.; CANDIAN, J. F. O efeito da escola básica brasileira: as evidências do PISA e do SAEB. **Revista Contemporânea de Educação**, v. 2, n. 4, p. 163-181, 2007.

SOARES, J. F.; COLLARES, A. C. M. Recursos familiares e o desempenho cognitivo dos alunos do ensino básico brasileiro. **Dados**, Rio de Janeiro, v. 49, n. 3, p.615-650, 2006.

SOARES, S.; SÁTYRO, N. **O impacto de infraestrutura escolar na taxa de distorção idade-série das escolas brasileiras de ensino fundamental: 1998 a 2005**. 2008.

SOUZA, A. M. Determinantes da aprendizagem em escolas municipais. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 13, n. 49, p. 413-434, 2005.

SOUZA, W. P. S. F.; OLIVEIRA, V. R.; ANNEGUES, A. C. 2018. Background familiar e desempenho escolar: uma abordagem não paramétrica. **Pesquisa e Planejamento Econômico**, v. 48, n.2, 2018

SUGUIY, Takao. **Eficiência versus satisfação no transporte público. Um estudo das práticas nas cidades brasileiras**. Tese (Doutorado em Engenharia Elétrica). Programa de Pós-graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 2017.

TAVARES, P.; CAMELO, R.; PACIÊNCIA, L. Uma análise do papel das escolas e das redes de ensino sobre as desigualdades de oportunidades educacionais. **Economia Aplicada**, v. 22, n. 2, p. 47-80, 2018.

THAPA, A. Public and Private School Performance in Nepal: An Analysis Using the SLC Examination. **Education Economics**, [s. l.], v. 23, n. 1, p. 47–62, 2015.

TODD, P. E.; WOLPIN, K. I. On the specification and estimation of the production function for cognitive achievement. **The Economic Journal**, v. 113, n. 485, p. F3-F33, 2003.

TODOS PELA EDUCAÇÃO. **Relatório Anual de Acompanhamento do Educação Já! Balanço 2019 e perspectivas 2020**. 1ª Edição, março de 2020.

TRAVITZKI, R. **ENEM: limites e possibilidades do Exame Nacional do Ensino Médio enquanto indicador de qualidade escolar**. 320 p. Tese (Doutorado em educação) – Programa de Pós-Graduação em Educação, Área de Concentração: Filosofia e Educação. Universidade de São Paulo. São Paulo, s.n., 2013.

VARDARDOTTIR, A. The impact of classroom peers in a streaming system. **Economics of Education Review**, v. 49, p. 110-128, 2015.

VARIAN, H. R. **Microeconomic Analysis**. 3ª ed., New York: Norton & Company, 1992.

VARIAN, H. R. **Microeconomia: uma abordagem moderna**. 9ª ed., Rio de Janeiro: Campus, 2012.

VASCONCELOS, J. C. et al. Infraestrutura escolar e investimentos públicos em Educação no Brasil: a importância para o desempenho educacional. **Ensaio: Avaliação e Políticas Públicas em Educação**, 2020.

VIANA, G.; LIMA, J. F. de. Capital humano e crescimento econômico. **Interações (Campo Grande)**, Campo Grande, v. 11, n. 2, p. 137-148, 2010.

VIGGIANO, E.; MATTOS, C. O desempenho de estudantes no Enem 2010 em diferentes regiões brasileiras. **Revista Brasileira de Estudos Pedagógicos**, v. 94, n. 237, p. 417-438, 2013.

VINHA, L. G. A; LAROS, J. A. El Uso de Datos de Evaluaciones Educativas en Brasil, Chile y Argentina. **Revista Iberoamericana de Evaluación Educativa**, [s. l.] v. 9, n. 2, p. 127–148, 2016.

WALTENBERG, F. D. Teorias econômicas de oferta de educação: evolução histórica, estado atual e perspectivas. **Educação e Pesquisa**, v. 32, n. 1, p. 117-136, 2006.

WILLMS, J. D.; RAUDENBUSH, S. W. A longitudinal hierarchical linear model for estimating school effects and their stability. **Journal of educational measurement**, v. 26, n. 3, p. 209-232, 1989.

APÊNDICE A – PROCESSO DE SELEÇÃO DOS ESTUDOS EMPÍRICOS

1 - OBJETIVO

Este Protocolo de Revisão Sistemática da Literatura (PRSL) apresenta a estrutura metodológica para a execução da etapa de revisão da literatura sobre o efeito da escola em instituições de ensino de acordo com o método proposto por Kitchenham (2004).

1.1 EQUIPE

Quadro 1 - Equipe Responsável para Revisão Sistemática de Literatura (RSL)

Nome	Papel	Afiliação
Dr. Tiago Wickstrom Alves	Orientador	UNISINOS/PPG Economia
Flávio Elizario de Souza	Acadêmico Doutorado	UNISINOS/PPG Economia

Fonte: Elaborado pelo autor (2020).

2. ESTRATÉGIA DE BUSCA

Na estratégia de busca, são definidas a questão principal a ser respondida juntamente com os critérios de pesquisa, as perguntas norteadoras e as bases de dados.

2.1. QUESTÃO PRINCIPAL

Para esta RSL, a questão principal que se busca responder é: Qual o efeito da escola no desempenho acadêmico dos estudantes identificados pela literatura em instituições de ensino e quais as principais metodologias empíricas aplicadas? Seguem-se os critérios da pesquisa:

Quadro 2 - Descrição dos critérios da pesquisa de RSL

Crítérios	Descrição
População	Produções científicas sobre efeito escola em instituições de ensino nos últimos 10 anos.
Intervenção	Estudos realizados em Instituições de ensino (preferencialmente ensino básico, ou seja, fundamental e médio).

Controle	Estudos com uso do efeito escola ou recursos escolares no desempenho dos estudantes.
Resultado	Qualidade aferida por meio da relação entre desempenho do estudante com a intervenção escolar.
Contexto de Aplicação	Impacto da intervenção escolar no desempenho estudantil.

Fonte: Elaborado pelo autor (2020).

Além da questão principal, delinea-se a pesquisa, conforme perguntas norteadoras apresentadas no quadro 3.

Quadro 3 - Perguntas norteadoras da pesquisa

Pergunta	Descrição da Pergunta
P1	O que a literatura empírica dos últimos 10 anos apresenta sobre o impacto da escola no desempenho acadêmico dos estudantes?
P2	Quais as principais metodologias utilizadas para se identificar o efeito da escola no desempenho dos estudantes?
P3	Qual a importância dos recursos escolares no contexto dos países desenvolvidos e em desenvolvimento?

Fonte: Elaborado pelo autor (2020).

2.2 BASES DE DADOS

2.2.1 Métodos de pesquisas de fontes

A pesquisa foi realizada em bases de dados disponíveis via *web*, a partir das palavras-chave identificadas por meio do referencial teórico. A pesquisa deu-se especificamente em artigos revisados pelos pares. No processo de busca, foram escolhidas as *strings* encontradas nos Títulos, Resumos e Palavras-chave. Com a filtragem, foram avaliados o título e o resumo para seleção de quais artigos seriam analisados. Após essa etapa, todos os artigos selecionados foram lidos, com a inclusão e/ou exclusão conforme critérios preestabelecidos.

2.2.2 Bases de Dados pesquisadas

Para esta revisão sistemática de literatura, foram utilizadas as bases de dados EBSCOhost, Science Direct e Scielo. A escolha dessas bases ocorreu pela amplitude de cada uma. A metodologia de busca foi realizada conforme segue.

2.2.3 EBSCOhost

Para essa base de dados, seguiu-se as *strings* de busca, conforme Glewwe *et al.* (2011) que realizaram uma atualização dos estudos sobre os determinantes do desempenho dos alunos em países em desenvolvimento, no período de 1990 a 2011.

Quadro 4 - Estratégia de busca por *strings*, período, descrição e resultado

<p>Strings e palavras-chave: "school effect " AND ("class size" OR "school size" OR "Student teacher ratio" OR "Pupil teacher ratio" OR "School expenditure*" OR "expenditure per pupil" OR "textbook*" OR "instructional material*" OR "Workbook*" OR "exercise book*" OR "computer*" OR "laptop*" OR "internet" OR "school infrastructure" OR "Facilities" OR "Building condition*" OR "Laborator*" OR "lab" OR "labs" OR "Librar*" OR "Desk*" OR "Teaching tools" OR "teaching guide*" OR "blackboard*" OR "chalk*" OR "electricity" OR "table*" OR "bench*" OR "chair*" OR "roof*" OR "wall*" OR "floor*" OR "window*" OR "bathroom*" OR "plumbing" OR "teacher quality" OR "teacher efficacy" OR "teacher knowledge" OR "teacher salar*" OR "teacher training" OR "teacher experience" OR "teacher education" OR "teacher absenteeism" OR "teacher gender" OR "class preparation" OR "lesson planning" OR "homework" OR "evaluation" OR "follow-up" OR "monitoring of pupil performance" OR "testing" OR "remedial program*" OR "teaching practices" OR "instructional time" OR "length of instructional program" OR "hours" OR "school day" OR "curriculum" OR "principal quality" OR "principal training" OR "principal education" OR "principal experience" OR "staff assessment*" OR "teacher assessment" OR "school inspection*" OR "parent* involvement" OR "production function" OR "school resources" OR "school inputs" OR "School quality" OR "Pedagogical inputs" OR "pedagogical resources")</p>
<p>Período: 2010 a 2019</p>
<p>Descrição: alguns filtros foram utilizados dentro da ferramenta de busca, como: texto completo; revistas acadêmicas (analisadas por especialistas); tipos de documentos periódicos científicos; expansores, aplicar assuntos equivalentes.</p>
<p>Resultado: após a aplicação dos filtros, 54 artigos foram encontrados, dos quais 28 duplicados, restando ao final 26 artigos. Após o processo de leitura foram excluídos 13 artigos, restando ao final 13 trabalhos.</p>

Fonte: Elaborado pelo autor (2020).

2.2.4 Science Direct

Para essa base de dados, as *strings* de busca foram adaptadas devido à quantidade de caracteres suportadas pelo buscador da plataforma. Para priorizar uma busca qualitativa, a pesquisa foi direcionada para periódicos na área de economia e educação. O período considerado foi de 2010 a 2021.

Quadro 5 - Estratégia de busca por *strings*, período, descrição e resultado

Strings e palavras-chave: ("school effect" OR "efeito escola" OR "efeito-escola" OR "effects of school" AND "school resource" OR "recursos escolares" OR "education production function" OR "student achievement" OR "economics of education")
Período: 2010 a 2021
Descrição: o filtro foi para: Research articles; Publication title; após as seguintes revistas foram selecionadas com as respectivas quantidades em parênteses: Economics of Education Review (17) International Journal of Educational Development (9) Journal of Public Economics (5) Studies in Educational Evaluation (5) Teaching and Teacher Education (3) Labour Economics (2) Regional Science and Urban Economics (2) Journal of Development Economics (2) World Development (2)
Resultado: após a aplicação dos filtros, foram encontrados 47 artigos, desses 16 foram refutados, restando 31. Após o processo de leitura de todo o documento, 15 foram excluídos, restando ao final 16 artigos.

Fonte: Elaborado pelo autor (2020).

2.2.5 Scielo

Para essa base de dados, as palavras-chave foram usadas em português e espanhol para se ampliar os estudos para o contexto da América Latina e do Brasil. O período considerado foi de 2010 a 2020.

Quadro 6 - Estratégia de busca por strings, período, descrição e resultado

Strings e palavras-chave: "recursos escolares" OR "efeito escola" OR "efecto escuela" OR "insumos escolares" OR "entradas escolares" OR "suministros escolares"
Período: 2010 a 2020
Descrição: foram usadas nas buscas os seguintes filtros: citável, artigo, todos os índices
Resultado: após a aplicação dos filtros, foram encontrados 18 artigos, destes 2 foram refutados, restando ao final 16 pesquisas. Após o processo de leitura de todo o documento, 6 foram refutados, restando ao final 10 artigos.

Fonte: Elaborado pelo autor (2020).

2.2.6 Resumo das Bases de Dados

Após o processo de busca e filtro dos trabalhos pesquisados por base de dados, restou a seguinte configuração, conforme o quadro 7:

Quadro 7 - Resumo da busca por trabalhos aceitos e excluídos

Base de Dados	Aceitos	Excluídos	Total
EBSCOhost	13	13	26
Science Direct	16	31	47
Scielo	10	8	18
Total	39	52	91

Fonte: Elaborado pelo autor (2020).

2.2.7 Idioma

Para esta Revisão Sistemática de Literatura (RSL), foram adotados o idioma inglês, como principal língua de pesquisa, o português e o espanhol. A escolha foi feita de acordo com a base de dados, conforme apresentado.

2.2.8 Critérios de Inclusão e Exclusão

Os critérios de inclusão e exclusão usados foram os seguintes:

Quadro 8 - Critérios de inclusão de artigos para RSL

Critério	Descrição do critério de inclusão
CI1	Artigos que tratam do efeito da escola no desempenho acadêmico
CI2	Artigos que relacionam recursos escolares com o desempenho do estudante
CI3	Artigos que utilizam análise quantitativa
CI4	Artigos que demonstram o impacto da escola podem ser incluídos
Critério	Descrição do critério de exclusão
CE1	Serão excluídos os artigos que não tratam do efeito escola
CE2	Serão excluídos os artigos que não se relacionam com o desempenho dos estudantes
CE3	Serão excluídos os artigos que não utilizam análise quantitativa

Fonte: Elaborado pelo autor (2020).

Ressalta-se que os artigos que tratam do impacto da escola por meio de outras variáveis diferentes do desempenho poderão ser incluídos, desde que demonstrem relação com a escola.