



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
DEPARTAMENTO DE ECONOMIA – DECON – SEDE
BACHAREL EM CIÊNCIAS ECONÔMICAS

RHUAAN PEDRO RODRIGUES DO NASCIMENTO

ESTRUTURA FAMILIAR E RESULTADOS EDUCACIONAIS NO BRASIL:
Evidências empíricas sobre arranjos monoparentais e desempenho escolar

RECIFE

2026

RHUAAN PEDRO RODRIGUES DO NASCIMENTO

ESTRUTURA FAMILIAR E RESULTADOS EDUCACIONAIS NO BRASIL
Evidências empíricas sobre arranjos monoparentais e desempenho escolar

Trabalho de Conclusão de Curso apresentado ao
curso de economia como requisito para obtenção
do título de Bacharel em ciências econômicas.

Orientadora: Sônia Maria Fonseca Pereira Oliveira
Gomes

RECIFE

2026

RHUAAN PEDRO RODRIGUES DO NASCIMENTO
ESTRUTURA FAMILIAR E RESULTADOS EDUCACIONAIS NO BRASIL: evidências
empíricas sobre arranjos monoparentais e desempenho escolar

Trabalho de Conclusão de Curso apresentado ao
curso de economia como requisito para obtenção
do título de Bacharel em ciências econômicas.

Aprovado em: _____

BANCA EXAMINADORA

Prof^ª. Dr^ª. Sônia Maria Fonseca Pereira Oliveira Gomes (Orientadora)
Universidade Federal Rural de Pernambuco

Prof. Dr. Ewerton Felipe de Melo Araújo (Examinador)
Universidade Federal Rural de Pernambuco

Prof. Dr. Luís Eduardo Barboza Carazza (Examinador)
Universidade Federal Rural de Pernambuco

DEDICATÓRIA

À minha mãe, Silvânia Pedro da Silva, que abriu mão da própria casa, foi para longe e ainda me ajudou financeiramente para que eu não desistisse. Ela acreditou que eu ia conseguir antes de eu acreditar em mim mesmo.

À minha esposa, Ana Letícia Marques do Carmo, pelo amor que sustentou cada etapa dessa caminhada.

AGRADECIMENTOS

À professora Sônia Maria Fonseca Pereira Oliveira Gomes, orientadora deste trabalho, que acreditou em mim num momento em que poucos acreditariam. Quando eu era estagiário, estava longe da família e vivia com menos de um salário, ela abriu uma porta que mudou o rumo da minha vida. Foi por meio do projeto Ecometrix, sob sua orientação, que descobri meu caminho na análise de dados, recebi uma indicação para minha primeira oportunidade profissional real e comecei a construir a carreira que tenho hoje. Este trabalho é também um reflexo do que aprendi com ela.

A Flávio Adriano Jesus de Araújo, colega do projeto Ecometrix e parceiro de trabalho, que me encorajou a dar o passo mais importante da minha trajetória profissional. Hoje dividimos não só o mesmo campo de atuação, mas a satisfação de termos chegado juntos até aqui.

EPIÍGRAFE

“Porque sou eu que conheço os planos que tenho para vocês, diz o Senhor, planos de fazê-los prosperar e não de causar dano, planos de dar a vocês esperança e um futuro.”

Jeremias 29:11

RESUMO

Este trabalho analisa a associação entre viver em famílias chefiadas exclusivamente pela mãe e o desempenho escolar de estudantes brasileiros do 5º e 9º anos do ensino fundamental, com foco nas proficiências em Língua Portuguesa e Matemática. Utilizando os microdados do Sistema de Avaliação da Educação Básica — SAEB 2023, estimam-se modelos de regressão linear múltipla em quatro blocos sequenciais de controles, incorporando características individuais, familiares, escolares e de envolvimento parental. As variáveis de acompanhamento familiar são utilizadas como proxies do capital social intrafamiliar definido por Coleman (1988). Os resultados indicam que, no 5º ano, estudantes de famílias chefiadas só pela mãe apresentam desvantagem bruta de desempenho que desaparece após o controle por características socioeconômicas, replicando o achado central de Lima, Carvalho e Silva (2021) de que o arranjo familiar funciona mais como marcador de condições socioeconômicas do que como determinante autônomo do desempenho. No 9º ano, o coeficiente bruto também é negativo, mas inverte de sinal após controles, sugerindo a presença de viés de seleção amostral associado à evasão escolar diferencial entre grupos. Os resultados contribuem para o debate sobre os determinantes do desempenho educacional no Brasil e reforçam que políticas voltadas ao apoio de famílias monoparentais devem considerar tanto a dimensão econômica quanto o fortalecimento das condições de acompanhamento escolar.

Palavras-chave: desempenho escolar; arranjos familiares monoparentais; capital social intrafamiliar; SAEB 2023; desigualdade educacional.

ABSTRACT

This study analyzes the association between living in mother-only headed households and the academic performance of Brazilian students in the 5th and 9th grades of elementary education, focusing on proficiency in Portuguese Language and Mathematics. Using microdata from the 2023 Basic Education Assessment System (SAEB), multiple linear regression models were estimated in four sequential blocks of controls, incorporating individual, family, school, and parental involvement characteristics. Family monitoring variables were used as proxies for intrafamily social capital as defined by Coleman (1988). The results indicate that, in the 5th grade, students from mother-only households show a raw performance disadvantage that disappears after controlling for socioeconomic characteristics, replicating the central finding of Lima, Carvalho e Silva (2021) that family structure functions more as a marker of socioeconomic conditions than as an autonomous determinant of academic performance. In the 9th grade, the raw coefficient is also negative but reverses sign after controls, suggesting the presence of sample selection bias associated with differential school dropout between groups. The findings contribute to the debate on the determinants of educational performance in Brazil and reinforce that policies aimed at supporting single-parent families should address both the economic dimension and the strengthening of school monitoring conditions.

Keywords: academic performance; single-parent family arrangements; intrafamily social capital; SAEB 2023; educational inequality.

LISTA DE ILUSTRAÇÕES

Gráfico 1 — Distribuição da proficiência por arranjo familiar — SAEB 2023.....	30
Gráfico 2 — Score médio de capital social intrafamiliar por arranjo familiar — SAEB 2023	31
Gráfico 3 — Evolução do coeficiente da dummy 'família só mãe' por bloco de controles (SAEB 2023).....	35

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas por arranjo familiar – 5º ano do ensino fundamental (SAEB 2023).....	28
Tabela 2 – Estatísticas descritivas por arranjo familiar – 9º ano do ensino fundamental (SAEB 2023).....	29
Tabela 3 – Coeficiente da dummy família só mãe por bloco de controles – Língua Portuguesa, 5º ano (SAEB 2023)	32
Tabela 4 – Coeficiente da dummy família só mãe por bloco de controles – Matemática, 5º ano (SAEB 2023)	32
Tabela 5 – Coeficiente da dummy família só mãe por bloco de controles – Língua Portuguesa, 9º ano (SAEB 2023)	33
Tabela 6 – Coeficiente da dummy família só mãe por bloco de controles – Matemática, 9º ano (SAEB 2023)	33
Tabela 7 – Coeficientes completos: Língua Portuguesa – 5º ano – Bloco 4 (SAEB 2023))	36
Tabela 8 – Coeficientes completos: Matemática – 5º ano – Bloco 4 (SAEB 2023)	37
Tabela 9 – Coeficientes completos: Língua Portuguesa – 9º ano – Bloco 4 (SAEB 2023)	39
Tabela 10 – Coeficientes completos: Matemática – 9º ano – Bloco 4 (SAEB 2023)	40
Tabela 11 – Descritivas por arranjo familiar detalhado – 5º e 9º anos (SAEB 2023)	42
Tabela 12 – Modelo com três categorias de arranjo familiar – 9º ano, Bloco 4 (SAEB 2023)	43

LISTA DE ABREVIATURAS

DECON	Departamento de Economia
IBGE	Instituto Brasileiro de Geografia e Estatística
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
LP	Língua Portuguesa
MT	Matemática
NCS	Nível de Capital Social
NSE	Nível Socioeconômico
SAEB	Sistema de Avaliação da Educação Básica
UFRPE	Universidade Federal Rural de Pernambuco

LISTA DE SÍMBOLOS

Y_i	Proficiência do estudante i em Língua Portuguesa ou Matemática
β_0	Intercepto do modelo de regressão
β_1	Coefficiente associado à variável indicadora de família chefiada apenas pela mãe
$MaeSolo_i$	Variável dummy indicadora de estudante que vive em família chefiada exclusivamente pela mãe
β_2	Vetor de coeficientes associados às características individuais do aluno
X_i	Vetor de características individuais do estudante (sexo, idade, cor/raça)
β_3	Vetor de coeficientes associados às características familiares
F_i	Vetor de características familiares e socioeconômicas
β_4	Vetor de coeficientes associados às características escolares
E_i	Vetor de variáveis da escola e da localização
β_5	Vetor de coeficientes associados às proxies de capital social intrafamiliar
CS_i	Vetor de proxies de capital social intrafamiliar (envolvimento parental)
ε_i	Termo de erro aleatório
i	Índice que identifica o estudante

SUMÁRIO

1	INTRODUÇÃO.....	16
2	OBJETIVOS.....	18
2.1	Objetivo Geral	18
2.2	Objetivos Específicos	18
3	REVISÃO DA LITERATURA.....	18
3.1	Arranjos familiares e desempenho escolar no Brasil: evidências empíricas	18
3.2	Capital social intrafamiliar e desempenho escolar.....	20
3.3	Síntese da literatura e lacunas de pesquisa	22
4	METODOLOGIA.....	23
4.1	Base de dados	24
4.2	Modelo Empírico.....	26
5	RESULTADOS.....	27
5.1	Análise Descritiva.....	27
5.2	Capital Social Intrafamiliar por Arranjo Familiar	30
5.3	Estimações por Blocos de Controles	31
5.4	Coeficientes Completos do Modelo — Bloco 4.....	35
5.5	Análise de Robustez: Heterogeneidade no Grupo Biparental	41
6	CONSIDERAÇÕES FINAIS	44

1 INTRODUÇÃO

Nas quatro últimas décadas, as famílias brasileiras tiveram suas composições mudadas. A taxa de fecundidade, que costumava ser de 4,35 filhos por mulher em 1980, caiu para 2,38 em 2000 e atingiu 1,55 em 2022, o menor nível histórico do Brasil (IBGE, 2025). No mesmo período, os casamentos formais perderam espaço: a taxa de nupcialidade era de 12,2 por grupo de 100 mil habitantes em 1980 e recuou para 5,6 em 2023 (IBGE, 2025). Os divórcios seguiram sentido contrário. Em 2022 aconteceram cerca de 420 mil divórcios, um aumento cerca de 9% acima do ano anterior o que fez com que a taxa subisse de cerca de 2,5 para 2,8 de separações matrimoniais por mil adultos, segundo o IBGE (2025). Esse conjunto de mudanças está diretamente ligado à transformação do papel da mulher na sociedade brasileira. A taxa de participação feminina no mercado de trabalho cresceu de forma expressiva nas últimas décadas, chegando a 53,3% em 2022 (IBGE, 2025). A escolaridade feminina acompanhou esse avanço: em 2022, 20,7% das mulheres com 25 anos ou mais tinham nível superior completo, proporção superior à dos homens, 15,8%.

A consequência mais visível em termos de estrutura familiar foi o avanço de famílias chefiadas por mulheres. Em 2000, elas representavam 22,2% das famílias brasileiras; em 2010 esse percentual chegou a 37,3% e em 2022 alcançou 48,8%, quase o dobro do registrado duas décadas anteriores (IBGE, 2025). Entre essas famílias, vale destacar aquelas em que a mãe cuida dos filhos sem a presença do cônjuge. O Censo 2022 identificou cerca de 7,8 milhões de mulheres nessa situação, representando 13,5% do total de famílias, uma alta frente aos 11,6% de 2000 (IBGE, 2025). Segundo Rizzotto, França e Frio (2018), essas famílias apresentam um nível socioeconômico menor socioeconômico em comparação aos arranjos biparentais, além de menor renda *per capita*. Dados do Censo 2000 já mostravam que 56,9% das crianças de 0 a 6 anos que viviam em domicílios chefiados por mulheres viviam com renda familiar de até dois salários-mínimos (IBGE, 2025). Esse dado é relevante, pois a estrutura familiar é um dos principais ambientes onde o desempenho escolar das crianças se forma. A literatura registra, de modo amplo, que as condições do domicílio estão diretamente associadas ao desempenho das crianças nas avaliações padronizadas (ALBERNAZ; FERREIRA; FRANCO, 2002; ALVES, 2010).

Rizzotto, França e Frio (2018) identificam diferenças entre arranjos familiares não só em renda, mas também nas escolhas de matrículas em escolas privadas, o que indica que os efeitos vão além do desempenho medido em testes. Lima, Carvalho e Silva (2021), por sua vez, analisando dados da Prova Brasil de 2015, mostram que o arranjo familiar funciona mais como marcador das condições socioeconômicas do que como determinante causal autônomo do

desempenho, após o controle por renda e características escolares. A teoria oferece duas lentes complementares para entender esses resultados. Coleman (1988) destaca que o capital social intrafamiliar, o conjunto de relações entre responsáveis e filhos que favorecem a construção de capital humano nas gerações seguintes, é tão decisivo quanto os recursos materiais. Mesmo quando o adulto tem escolaridade elevada, o rendimento escolar da criança pode ser comprometido se o tempo dedicado a ela é escasso. Nas famílias chefiadas exclusivamente pela mãe, a dupla jornada de trabalho e cuidado tende a comprimir justamente esse espaço relacional (COLEMAN, 1988; BENATTI et al., 2021). A renda familiar e a escolaridade dos responsáveis também figuram entre os principais determinantes do desempenho das crianças, de modo que famílias com menor capacidade financeira tendem a oferecer menos recursos educacionais aos filhos, perpetuando desigualdades entre gerações (ALBERNAZ; FERREIRA; FRANCO, 2002; LIMA; CARVALHO; SILVA, 2021).

Neste âmbito, o presente trabalho analisa a associação entre viver em famílias chefiadas por mães e o desempenho escolar de estudantes brasileiros do 5º e 9º anos do ensino fundamental, com foco nas proficiências em Língua Portuguesa e Matemática. Para tanto, utilizam-se os microdados do Sistema de Avaliação da Educação Básica SAEB 2023, edição mais recente com microdados públicos disponíveis e cobertura nacional (INEP, 2024). Esses dados reúnem informações sobre desempenho escolar, composição do domicílio, características socioeconômicas e variáveis de envolvimento parental estas últimas utilizadas como *proxies* do capital social intrafamiliar definido por Coleman (1988).

A metodologia adotada combina levantamento bibliográfico, análise de dados secundários e estimações econométricas com o intuito de encontrar relações entre o desempenho escolar de estudantes do ensino fundamental e a estrutura familiar brasileira. Na parte bibliográfica, o referencial teórico apoia-se na teoria do capital social (COLEMAN, 1988) e nos estudos empíricos sobre estrutura familiar e desempenho educacional no Brasil, articulados com evidências sobre os efeitos das condições socioeconômicas do domicílio sobre a trajetória escolar das crianças.

Desta forma, além deste capítulo introdutório, esta monografia organiza-se em mais cinco capítulos. O segundo apresenta os objetivos geral e específicos da pesquisa. O terceiro é dedicado à revisão da literatura, com o referencial sobre capital humano e capital social e a sistematização dos estudos empíricos que relacionam arranjos familiares e desempenho escolar no Brasil. O quarto capítulo descreve a metodologia a ser utilizada. O quinto expõe e discute os resultados estimados e, por fim, o sexto traz as considerações finais.

2 OBJETIVOS

2.1 Objetivo Geral

Analisar a associação entre viver em famílias chefiadas por mães e o desempenho escolar de estudantes do 5º e 9º anos do ensino fundamental no Brasil, com foco nos testes de proficiência em Língua Portuguesa e Matemática, a partir dos microdados do Sistema de Avaliação da Educação Básica — SAEB 2023.

2.2 Objetivos Específicos

- Descrever de que forma se dá a associação entre arranjo familiar e o desempenho acadêmico de alunos;
- Descrever de que forma se dá a associação entre capital social intrafamiliar e o desempenho escolar;
- Descrever o perfil socioeconômico e o envolvimento parental dos estudantes em famílias chefiadas por mães em comparação aos que vivem com ambos os responsáveis, com base nos microdados do SAEB 2023;
- Estimar a associação entre viver em família chefiada por mães e as proficiências em Língua Portuguesa e Matemática dos alunos do 5º e 9º anos, com e sem controles socioeconômicos e escolares;
- Verificar se a associação estimada varia segundo a série escolar, o sexo do aluno, o tipo de escola e o nível socioeconômico da família;
- Examinar em que medida os indicadores de envolvimento parental — frequência de conversas sobre a escola, incentivo aos estudos e participação em reuniões — se associam ao desempenho escolar, operando como *proxies* do capital social intrafamiliar definido por Coleman (1988).

3 REVISÃO DA LITERATURA

3.1 Arranjos familiares e desempenho escolar no Brasil: evidências empíricas

Nos últimos tempos houve um crescimento da literatura brasileira sobre a análise da associação entre o desempenho escolar e a composição familiar, sendo que os resultados têm apontado para a mesma direção, ou seja, os arranjos familiares ou a forma como as famílias são constituídas explicam pouco quando as condições socioeconômicas são controladas no ambiente escolar. O que tende a influenciar não é com quem a criança mora, e sim se os adultos

se envolvem na educação dessa criança (PALERMO; SILVA; NOVELLINO, 2014; MARTINS; TEIXEIRA, 2021; LIMA; CARVALHO; SILVA, 2021). Essa descoberta não encerra o debate, mas abre uma discussão mais difícil: se o meio onde essa criança vive é só um marcador de pobreza ou riqueza, o que se associa ao desempenho dela?

Palermo, Silva e Novellino (2014), usando modelo com três níveis aplicado a dados da Prova Brasil 2007 para alunos do 5º ano da rede municipal do Rio de Janeiro, descobriram que o arranjo biparental teve efeito positivo sobre o desempenho em matemática, tendo 1,2 pontos a mais que os alunos com estrutura familiar monoparental sendo uma diferença considerada pequena. No entanto, o que chama atenção nesse artigo não é esse resultado e sim a construção que as autoras desenvolveram. Elas mostram o nível de capital social, que é adquirido quando os pais frequentam reuniões escolares, incentivam a leitura e os estudos, ajudam no dever de casa e a frequência com que eles conversam com os filhos sobre o dia a dia na escola, teve associação positiva com o desempenho escolar, de modo que maiores níveis de envolvimento parental estavam associados as notas mais altas em matemática (PALERMO; SILVA; NOVELLINO, 2014). Esse resultado mostra que o envolvimento parental no dia a dia escolar também importa, exatamente como Coleman (1988) previa ao definir o capital social intrafamiliar.

Martins e Teixeira (2021), em seguida, encontraram resultados estatisticamente mais significativas. Usando modelo hierárquico com dados do SAEB 2013 para alunos do 9º ano de todo o Brasil, mostraram que os estudantes com famílias monoparentais tinham desempenho menor do que os alunos com arranjos familiares biparentais. Em Língua Portuguesa, alunos que moravam só com a mãe tiveram, em média, 0,42 ponto a menos que os de famílias biparentais; os que moravam só com o pai ficaram 1,46 pontos abaixo. Em Matemática, as diferenças foram de 0,82 e 1,45 ponto. As justificativas apresentadas pelos autores iam na direção de menor disponibilidade financeira e de tempo desse responsável na criação da criança e do acompanhamento dos estudos. Uma outra descoberta importante foi a de que os alunos cujas famílias são chefiadas apenas pela mãe têm desempenho maior que aquelas chefiadas apenas pelo pai, a diferença de 1,03 ponto em português e 0,63 em matemática, foram pontuados como associação maior da escolaridade materna sobre a trajetória escolar dos filhos (MARTINS; TEIXEIRA, 2021).

Lima, Carvalho e Silva (2021) chegaram à conclusão diferente ao analisar dados da Prova Brasil de 2015, com alunos do 5º e 9º anos de escolas públicas de todo o Brasil. Controlando por características socioeconômicas e do ambiente escolar, os autores encontraram efeito negativo e de pequena magnitude: no 5º ano, alunos que moravam só com a mãe

apresentaram nota em média 1,63 ponto inferior à dos que viviam com ambos os pais; no 9º ano essa diferença recuou para 0,19 ponto. Ambos foram estatisticamente significativos, mas a magnitude levou os autores a concluir que o arranjo familiar funciona mais como marcador das condições socioeconômicas do que como determinante autônomo do desempenho (LIMA; CARVALHO; SILVA, 2021). A diferença de resultados entre Martins e Teixeira (2021) e Lima, Carvalho e Silva (2021) pode ser explicada pelas distintas estratégias de controle adotadas: o primeiro trabalho usa um índice socioeconômico construído a partir de bens de consumo; o segundo inclui controles mais amplos sobre o ambiente escolar e a trajetória do aluno.

Os três estudos revelam uma lacuna que o presente trabalho busca preencher. Palermo, Silva e Novellino (2014) construíram um indicador de capital social familiar e mostraram sua relevância, mas a análise teve restrição ao 5º ano de uma única rede municipal do Rio de Janeiro. Martins e Teixeira (2021) cobriram todo o Brasil, mas usaram dados de 2013 e analisaram apenas o 9º ano. Lima, Carvalho e Silva (2021) avançaram em abrangência, mas cobriram apenas matemática e dados de 2015, oito anos antes do SAEB 2023, e não testaram se o envolvimento parental se associa à relação entre arranjo e desempenho. O presente trabalho usa o SAEB 2023, que abrange escolas públicas e privadas, analisa tanto Língua Portuguesa quanto matemática para o 5º e 9º anos, incorpora de forma sistemática as variáveis de envolvimento parental como proxies do capital social intrafamiliar definido por Coleman (1988) dimensão que os trabalhos anteriores reconheceram como relevante, mas não exploraram de forma direta e estruturada em nível nacional.

3.2 Capital social intrafamiliar e desempenho escolar

O capital social parte de uma ideia concreta: relações entre pessoas produzem recursos. Coleman (1988) define o conceito de capital social pela função que este cumpre e que são aspectos da estrutura social que viabilizam certas ações que, sem eles, não ocorreriam. O que o distingue de outras formas de capital está no lugar onde existe, ou seja, o capital físico está em objetos; capital humano, dentro das pessoas, em habilidades e escolaridade acumuladas e o capital social está nas relações, e só quem participa delas pode acessá-lo.

Putnam (1993) levou essa noção para o plano coletivo, associando capital social a redes de participação cívica, normas de reciprocidade e confiança entre membros de uma comunidade. Bourdieu (1980) seguiu caminho diferente. Para ele, trata-se dos recursos que um indivíduo consegue mobilizar por pertencer a redes de relações relativamente estáveis, e sua distribuição não é neutra, ela tende a reproduzir posições sociais. As ênfases não são as mesmas,

mas as três abordagens partem do mesmo reconhecimento de que vínculos sociais têm valor e que esse valor varia conforme a estrutura e a intensidade das relações.

Na formulação de Coleman (1988), o capital social passa menos pelos atributos isolados dos adultos e mais pelos recursos que surgem das relações que cercam a criança. Quando a análise é trazida para dentro da família, o ponto central deixa de ser apenas a escolaridade dos responsáveis e passa a incluir o tipo de vínculo que eles constroem com os filhos no cotidiano. A diferença para o capital humano é importante aqui. O capital humano está dentro da pessoa, sob a forma de escolaridade, habilidades e repertórios cognitivos; o capital social, por sua vez, está entre as pessoas, nas relações que facilitam orientação, acompanhamento e apoio. Por isso, Coleman (1988) sustenta que a educação dos pais, sozinha, não resolve o problema: ela pode ter pouco efeito sobre a trajetória escolar da criança quando não vem acompanhada de presença, atenção e interação familiar.

No plano intrafamiliar, esse capital social pode ser observado em três frentes. A primeira envolve obrigações e expectativas, isto é, a ideia de que o adulto se percebe responsável pelo desenvolvimento da criança e de que a criança espera esse suporte. A segunda aparece nos canais de informação, quando os responsáveis conversam sobre a escola, acompanham o progresso do aluno e mantêm contato com o ambiente escolar. A terceira está nas normas e sanções, expressas em regras sobre estudar, fazer dever e frequentar as aulas. O ganho teórico dessa leitura está em mostrar que alta escolaridade dos responsáveis não significa, automaticamente, alto capital social. Pode existir muito capital humano e pouco capital social quando o adulto está distante da rotina da criança, física ou relacionalmente. Também pode ocorrer o contrário: um responsável com baixa escolaridade, mas presente e envolvido, pode gerar um ambiente mais favorável ao desempenho escolar (COLEMAN, 1988).

Esse debate ganha força quando se observam as famílias chefiadas por mães. O problema não está, por si só, no arranjo monoparental, como se a ausência de um segundo responsável determinasse automaticamente piores resultados. O ponto mais delicado parece estar na compressão do tempo disponível para produzir essas relações. Quando trabalho remunerado, cuidado, supervisão doméstica e acompanhamento escolar se concentram em uma só pessoa, a chance de enfraquecimento do capital social intrafamiliar cresce.

A revisão bibliográfica de Benatti et al. (2021) ajuda a sustentar esse entendimento ao mostrar que, nas trajetórias monoparentais, as práticas de cuidado, as redes de apoio e a forma de exercício da parentalidade seguem centrais para os desfechos das crianças. Assim, mais do que perguntar com quem o aluno mora, interessa observar se há tempo, disponibilidade e envolvimento real na sua vida escolar. E esta chave analítica se aproxima do modo como

Palermo, Silva e Novellino (2014) operacionalizam o Nível de Capital Social (NCS) entendido como um indicador do interesse dos pais pela educação dos filhos e da supervisão exercida sobre sua vida escolar, com base em incentivos, presença na escola e conversas sobre o cotidiano escolar.

3.3 Síntese da literatura e lacunas de pesquisa

De uma forma geral, os estudos discutidos ao longo deste capítulo apontam mais convergências do que rupturas. A principal delas é que a estrutura familiar, isoladamente, tem poder explicativo limitado sobre o desempenho escolar quando comparada ao peso das condições socioeconômicas, da trajetória do aluno e do ambiente escolar. Isso aparece tanto nos trabalhos que analisam diretamente a proficiência em avaliações padronizadas quanto naqueles que observam decisões familiares ligadas ao investimento educacional. Em termos mais amplos, a literatura sugere que o arranjo familiar funciona menos como causa autônoma e mais como parte de um conjunto de condições que cercam o processo de escolarização.

As divergências entre os resultados não parecem decorrer de uma contradição substantiva entre os autores, mas das diferentes estratégias empíricas adotadas. Albernaz, Ferreira e Franco (2002) examinam determinantes mais gerais do desempenho escolar e mostram a força da composição socioeconômica dos alunos. Alves (2010) destaca o papel das escolhas familiares e dos estabelecimentos escolares sobre a aprendizagem em um contexto urbano específico. Rizzotto, França e Frio (2018) analisam como os arranjos familiares influenciam a decisão de matrícula na rede pública ou privada, mostrando que os efeitos do arranjo familiar sobre as oportunidades educacionais podem surgir antes mesmo de aparecerem nas notas, por meio das escolhas de matrícula. Já Lima, Carvalho e Silva (2021) aproximam-se mais diretamente do problema aqui estudado e concluem que os arranjos familiares têm pouco efeito sobre a proficiência média quando fatores socioeconômicos, escolares e da trajetória do aluno são controlados. Benatti et al. (2021), por sua vez, ajudam a relativizar interpretações deterministas ao mostrar que a monoparentalidade, por si só, não deve ser entendida como fator automático de risco ao desenvolvimento infantil.

A lacuna que permanece está menos na identificação da existência de diferenças entre arranjos familiares e mais na compreensão dos mecanismos que conectam estrutura familiar e desempenho escolar. Ainda são escassos, na literatura brasileira, estudos recentes de abrangência nacional que articulem composição do domicílio, proficiência em diferentes disciplinas e indicadores mais diretos de envolvimento parental. Também são limitados os trabalhos que tratam esse envolvimento como dimensão empiricamente observável do capital

social intrafamiliar. É justamente nesse ponto que se insere o presente estudo. Ao utilizar os microdados do SAEB 2023, cobrindo escolas públicas e privadas, o 5º e o 9º anos e variáveis de acompanhamento familiar, a pesquisa busca avançar sobre uma lacuna ainda presente na literatura brasileira: examinar, de forma mais atualizada e estruturada, a associação entre viver em famílias chefiadas por mães, o envolvimento parental e o desempenho escolar em Língua Portuguesa e Matemática.

4 METODOLOGIA

A metodologia adotada combina levantamento bibliográfico, análise de dados secundários e estimções econométricas visando encontrar relações entre o desempenho escolar de estudantes do ensino fundamental, a estrutura familiar brasileira e o capital social das famílias brasileiras. No âmbito bibliográfico, o estudo apoia-se na teoria do capital social (COLEMAN, 1988) e nos estudos empíricos resultantes.

A estratégia econométrica adotada nesta pesquisa será a estimação de modelos de regressão linear múltipla, ajustados separadamente para as proficiências em Língua Portuguesa e Matemática. A escolha desse procedimento se justifica pelo fato de as variáveis dependentes do estudo serem contínuas e expressarem diretamente o desempenho dos estudantes nas escalas do SAEB. Além disso, a regressão linear permite avaliar a associação entre a variável principal de interesse e a proficiência dos alunos, controlando simultaneamente um conjunto de características individuais, familiares, entre elas o capital social e escolares.

A etapa empírica da pesquisa foi conduzida em *Python*¹, linguagem escolhida por sua capacidade de lidar com bases extensas, reprodutibilidade analítica e flexibilidade no tratamento de dados educacionais. Para a organização, limpeza e manipulação dos microdados do SAEB 2023, foi utilizada a biblioteca *pandas*², com apoio de *numpy*³ para operações numéricas. As estimções dos modelos econométricos foram realizadas com a biblioteca *statsmodels*⁴, enquanto a produção de estatísticas descritivas e visualizações gráficas contou

¹ *Python*: linguagem de programação de código aberto, amplamente utilizada em análise de dados e pesquisa científica pela sua versatilidade e vasta biblioteca de pacotes estatísticos.

² *pandas*: biblioteca *Python* para manipulação e análise de dados tabulares, permitindo leitura, limpeza e transformação de bases de dados de grande volume.

³ *numpy*: biblioteca *Python* voltada para operações matemáticas e matriciais de alto desempenho, utilizada como base por diversas outras bibliotecas científicas.

⁴ *statsmodels*: biblioteca *Python* especializada em estimação de modelos estatísticos e econométricos, incluindo regressão linear com suporte a pesos amostrais e erros-padrão robustos.

com o uso de *matplotlib*⁵. Quando necessário, arquivos em formato Excel foram lidos com apoio de *openpyxl*⁶.

4.1 Base de dados

A base de dados para a análise empírica é composta pelos microdados do Sistema de Avaliação da Educação Básica – SAEB, 2023, produzidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). O SAEB é uma avaliação externa em larga escala voltada à produção de indicadores sobre qualidade, equidade e eficiência da educação básica brasileira, sendo aplicado periodicamente em etapas específicas da escolarização. Na edição de 2023, o sistema contemplou estudantes do 2º, 5º e 9º anos do ensino fundamental e da 3ª série do ensino médio regular e da 4ª série do ensino médio técnico integrado, em escolas públicas e privadas localizadas em áreas urbanas e rurais.

Para os fins deste estudo, o recorte foi delimitado aos estudantes do 5º e 9º anos do ensino fundamental, pois essas etapas reúnem, simultaneamente, informações de proficiência em Língua Portuguesa e Matemática, além das respostas ao questionário contextual do aluno, fundamentais para a construção das variáveis de arranjo familiar, escolaridade materna, nível socioeconômico e envolvimento parental.

No SAEB 2023, a aplicação de testes para a avaliação da Língua Portuguesa e da Matemática para o 5º e o 9º anos se dão de forma censitária nas escolas públicas e amostral nas escolas privadas, preservando a comparabilidade da série histórica dessas disciplinas. A população de referência da avaliação exclui escolas com menos de dez estudantes matriculados na etapa considerada, turmas multisseriadas, turmas de correção de fluxo, Educação de Jovens e Adultos, ensino normal/magistério, educação especial exclusiva e escolas indígenas que não ministram a Língua Portuguesa como primeira língua. Dessa forma, os resultados do SAEB 2023 são obtidos com base na população efetivamente elegível segundo os critérios metodológicos definidos pelo INEP. E para as estimações serão considerados os pesos amostrais correspondentes a cada área, de modo a respeitar o desenho da avaliação e produzir estimativas estatisticamente consistentes.

Quanto ao desenho amostral, o SAEB organiza-se em estratos, conglomerados e pesos. Nesta pesquisa, as estimativas incorporam os pesos amostrais do aluno (PESO_ALUNO_LP e PESO_ALUNO_MT), que corrigem as probabilidades desiguais de seleção entre as redes —

⁵ *matplotlib*: biblioteca *Python* para criação de gráficos e visualizações estatísticas, utilizada na produção das figuras apresentadas neste trabalho.

⁶ *openpyxl*: biblioteca *Python* para leitura e escrita de arquivos no formato Excel (.xlsx), utilizada no acesso a arquivos auxiliares da pesquisa.

uma vez que a aplicação é censitária nas escolas públicas e amostral nas privadas, sem ponderação as estimativas seriam dominadas pela rede pública e deixariam de refletir a população de interesse. O conglomerado, definido pela escola, é tratado por meio dos erros padrão clusterizados (HC1) no nível da escola, que reconhecem a não independência entre alunos da mesma unidade. Os estratos da avaliação correspondem a recortes de unidade da federação, dependência administrativa e localização; suas dimensões são incorporadas ao modelo na forma de variáveis de controle (região, dependência administrativa e localização urbana ou rural), de modo que a comparação entre arranjos familiares ocorre dentro de condições estruturais semelhantes.

No tratamento dos dados ausentes, optou-se pela exclusão das observações que não apresentavam resposta nas variáveis utilizadas na construção do arranjo familiar, dos controles socioeconômicos e das proxies de capital social (exclusão listwise). Reconhece-se que a não resposta pode não ser completamente aleatória e que sua remoção, ainda que necessária à estimação, pode afetar a representatividade da amostra analisada. Por essa razão, o número de observações varia entre os blocos de controle, conforme indicado nas tabelas de resultados, refletindo a perda de casos com informação incompleta nas variáveis adicionadas em cada etapa.

As variáveis dependentes da pesquisa serão os indicadores de desempenho em português e matemática, identificados no dicionário SAEB como PROFICIENCIA_LP_SAEB e PROFICIENCIA_MT_SAEB, respectivamente. A variável principal de interesse será uma *dummy* indicadora de estudante que vive em família chefiada apenas pela mãe, construída a partir das respostas do questionário do aluno sobre presença de mãe ou madrasta no domicílio e ausência de pai ou padrasto.

Para captar as dimensões sobre o capital social intrafamiliar, serão utilizadas variáveis do bloco Q10 do questionário do aluno, que captam frequência de conversa sobre a escola, incentivo aos estudos, estímulo à realização das tarefas, incentivo à frequência às aulas e comparecimento às reuniões de pais. Também serão incluídos controles individuais, familiares e escolares, como nível socioeconômico do aluno, escolaridade da mãe, sexo, idade, cor ou raça, dependência administrativa da escola, localização e região do país.

O score de capital social intrafamiliar foi construído como média simples das cinco variáveis do bloco Q10, atribuindo-se peso igual a cada indicador. Essa escolha segue a operacionalização adotada por Palermo, Silva e Novellino (2014), que combinam incentivos, presença na escola e conversas sobre o cotidiano escolar em um único indicador de envolvimento parental. Na ausência, na literatura, de evidência que estabeleça pesos

diferenciados entre essas dimensões, a atribuição de pesos iguais constitui a alternativa mais conservadora, por não pressupor, a priori, que um tipo de envolvimento seja mais relevante que outro. A baixa correlação observada entre os cinco indicadores (detalhada no Apêndice) reforça que cada um capta uma faceta distinta do acompanhamento familiar, o que é coerente com seu tratamento como componentes de igual peso.

4.2 Modelo Empírico

O modelo a ser estimado nesta pesquisa tem como objetivo analisar a associação entre o arranjo familiar e o desempenho escolar dos estudantes, controlando por características individuais, familiares, escolares e relacionais. Para isso, serão estimados modelos de regressão linear múltipla, separadamente para as proficiências em Língua Portuguesa e Matemática.

De forma geral, o modelo pode ser representado pela Equação 1 seguinte:

$$Y_i = \beta_0 + \beta_1 MaeSolo_i + \beta_2 X_i + \beta_3 F_i + \beta_4 E_i + \beta_5 CS_i + \varepsilon_i \quad (1)$$

em que Y_i representa a proficiência do estudante i em Língua Portuguesa ou Matemática; $MaeSolo_i$ é a variável indicadora de estudante que vive em família chefiada exclusivamente pela mãe; X_i , corresponde às características individuais do aluno; F_i reúne características familiares e socioeconômicas; E_i representa variáveis da escola e da localização CS_i corresponde às proxies de capital social intrafamiliar; e ε_i é o termo de erro aleatório.

E, a interpretação do coeficiente β_1 é central para o estudo, pois indica a diferença média de proficiência entre estudantes em famílias chefiadas apenas pela mãe e os demais alunos do grupo de comparação. Caso esse coeficiente diminua de magnitude após a inclusão das variáveis de capital social, isso poderá ser interpretado como evidência de que parte dessa associação está relacionada ao nível de envolvimento parental e à qualidade das interações familiares no contexto educacional.

As estimações serão realizadas em etapas sucessivas. Em um primeiro modelo, será incluída apenas a variável de interesse principal, a fim de observar sua associação bruta com o desempenho escolar. Em seguida, serão adicionadas variáveis de controle individuais e familiares, como sexo, idade, cor ou raça, nível socioeconômico do aluno e escolaridade da mãe. Em um terceiro momento, serão incorporadas variáveis escolares e territoriais, como dependência administrativa da escola, localização urbana ou rural e região do país. Por fim, serão acrescentadas as variáveis de envolvimento parental, tratadas nesta pesquisa como proxies do capital social intrafamiliar, permitindo verificar se parte da associação entre arranjo

familiar e desempenho escolar é absorvida por diferenças nas relações de acompanhamento e monitoramento educacional dentro do domicílio. Essa opção por uma estratégia de estimação em blocos tem importância analítica para os objetivos do trabalho. Ela permite verificar se a associação entre viver em família chefiada apenas pela mãe e o desempenho escolar permanece estável ou se se reduz à medida que entram no modelo variáveis socioeconômicas, escolares e relacionais.

O uso de mínimos quadrados ponderados (MQP), em vez dos ordinários (MQO), decorre do desenho amostral do SAEB: os alunos são selecionados com probabilidades desiguais entre estratos, de modo que certas regiões e tipos de escola ficam sobre ou sub-representados na amostra. Sem os pesos, as estimativas de β seriam viesadas (GUJARATI, 2006). A ponderação corrige essa distorção e faz com que os resultados se refiram à população de interesse — estudantes do 5º e 9º anos das redes públicas brasileiras — e não apenas à amostra observada.

Os erros padrão robustos HC1, clusterizados por escola, foram adotados por duas razões práticas. A variância dos resíduos tende a diferir entre grupos de alunos com perfis socioeconômicos distintos — o que caracteriza heterocedasticidade. Além disso, alunos da mesma escola não são independentes entre si: compartilham professores, infraestrutura e cultura institucional. Tratar observações intracluster como independentes subestima os erros padrão e infla a significância dos coeficientes, um problema bem documentado na literatura econométrica (GUJARATI, 2006; WHITE, 1980).

5 RESULTADOS

5.1 Análise Descritiva

Esta seção apresenta os resultados da pesquisa em duas etapas. Primeiro, examina-se o perfil descritivo dos estudantes segundo o arranjo familiar, comparando as características observáveis dos grupos biparental e só mãe. Em seguida, apresentam-se as estimações por blocos de controles.

As Tabelas 1 e 2 apresentam as estatísticas descritivas para os alunos do 5º e do 9º anos, respectivamente, diferenciando-os por arranjo familiar. As diferenças entre grupos foram testadas por meio de teste t com variâncias heterogêneas.

Tabela 1 – Estatísticas descritivas por arranjo familiar – 5º ano do ensino fundamental (SAEB 2023)

Variável	Total	Biparental	Só mãe	Diferença
Proficiência LP (SAEB)	219,16	220,54	213,15	−7,38***
Proficiência MT (SAEB)	230,68	232,27	223,73	−8,54***
INSE (nível socioecon.)	5,18	5,26	4,81	−0,45***
Feminino (%)	48,0%	47,0%	49,0%	+2,0 p.p.***
Não branco (%)	68,0%	67,0%	71,0%	+4,0 p.p.***
Distorção idade-série (%)	8,0%	7,0%	11,0%	+4,0 p.p.***
Mãe c/ EM ou mais (%)	33,0%	34,0%	31,0%	−3,0 p.p.***
Escola pública (%)	80,0%	79,0%	87,0%	+8,0 p.p.***
Área urbana (%)	91,0%	90,0%	92,0%	+2,0 p.p.***
Score capital social (0–2)	1,62	1,63	1,58	−0,06***

Nota: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. Diferença = (Só mãe) – (Biparental). N = 1.626.261 (só mãe: 325.405; biparental: 1.300.856). Pesos amostrais aplicados.

Fonte: SAEB 2023/INEP. Elaboração própria.

No 5º ano, a comparação bruta das proficiências revela uma desvantagem clara para os estudantes de famílias chefiadas só pela mãe. Em Língua Portuguesa, a diferença média não ponderada é de −7,38 pontos (biparental: 220,54; só mãe: 213,15), e em Matemática é de −8,54 pontos (biparental: 232,27; só mãe: 223,73), ambas estatisticamente significativas ($p < 0,001$). Esse padrão é consistente com o esperado pela literatura: famílias monoparentais apresentam, em média, nível socioeconômico mais baixo (INSE de 4,81 contra 5,26 nas biparentais), maior concentração em escolas públicas (87% contra 79%) e maior proporção de alunos com distorção idade-série (11% contra 7%). A diferença bruta de desempenho é, portanto, provavelmente confundida por essas características observáveis, o que motiva a análise de regressão apresentada a seguir.

O score de capital social intrafamiliar — construído como média simples dos cinco indicadores de envolvimento parental do bloco Q10 do questionário do aluno, com escala de 0 (nunca) a 2 (sempre ou quase sempre) — é ligeiramente inferior no grupo só mãe (1,58 contra 1,63 no grupo biparental), diferença estatisticamente significativa. Essa diferença, embora pequena em magnitude, sugere que a compressão de tempo característica das mães que chefiam o domicílio sozinhas se reflete marginalmente em menor frequência de práticas de acompanhamento escolar, em linha com o argumento de Coleman (1988) sobre o capital social intrafamiliar.

Tabela 2 – Estatísticas descritivas por arranjo familiar – 9º ano do ensino fundamental (SAEB 2023)

Variável	Total	Biparental	Só mãe	Diferença
Proficiência LP (SAEB)	263,02	263,75	260,35	-3,41***
Proficiência MT (SAEB)	262,03	263,29	257,36	-5,94***
INSE (nível socioecon.)	5,21	5,31	4,83	-0,48***
Feminino (%)	49,0%	48,0%	50,0%	+2,0 p.p.***
Não branco (%)	64,0%	63,0%	67,0%	+4,0 p.p.***
Distorção idade-série (%)	11,0%	10,0%	13,0%	+3,0 p.p.***
Mãe c/ EM ou mais (%)	50,0%	51,0%	48,0%	-3,0 p.p.***
Escola pública (%)	82,0%	81,0%	87,0%	+6,0 p.p.***
Área urbana (%)	91,0%	90,0%	93,0%	+3,0 p.p.***
Score capital social (0-2)	1,58	1,60	1,54	-0,06***

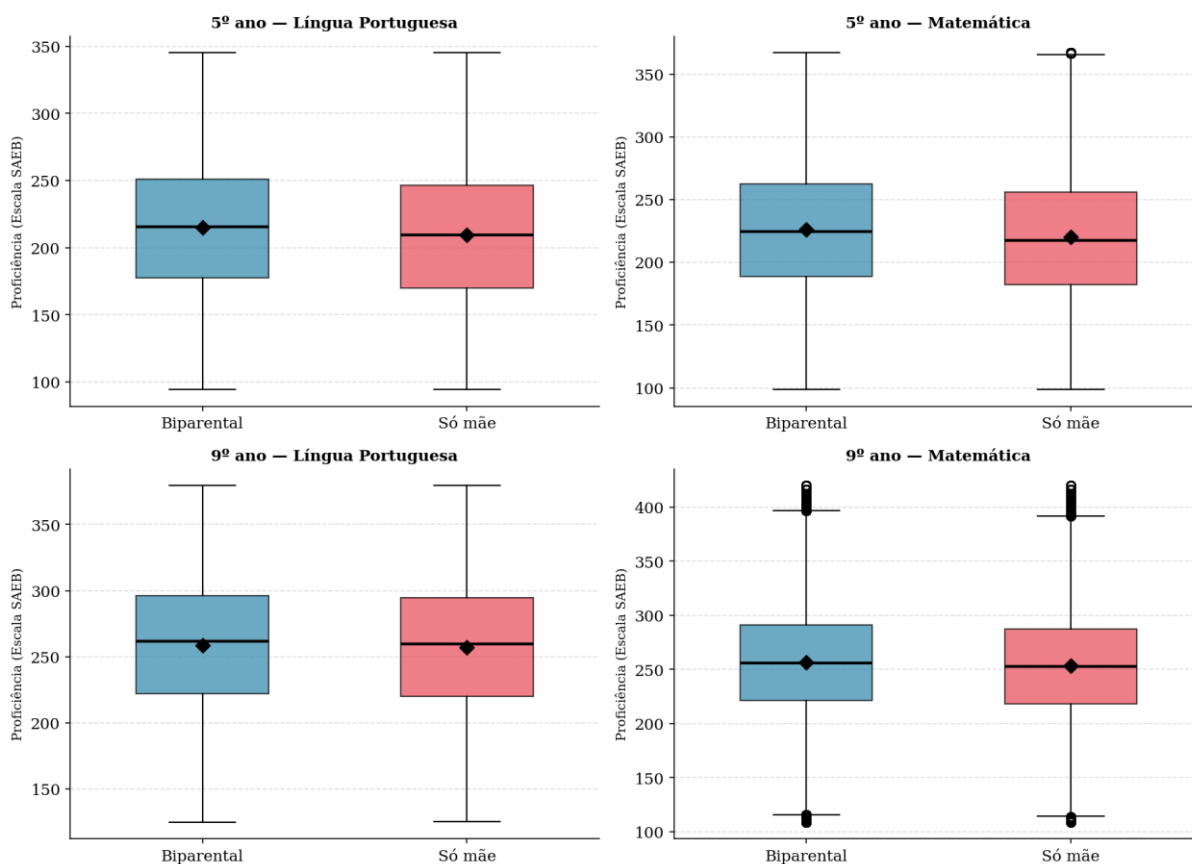
Nota: *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. Diferença = (Só mãe) – (Biparental). N = 1.633.208 (só mãe: 364.435; biparental: 1.268.773). Pesos amostrais aplicados.

Fonte: SAEB 2023/INEP. Elaboração própria.

No 9º ano, o padrão descritivo é semelhante ao do 5º ano, embora com diferenças de menor magnitude. Alunos de famílias chefiadas só pela mãe apresentam proficiência média inferior tanto em Língua Portuguesa (260,35 contra 263,75; diferença = -3,41 pontos; $p < 0,001$) quanto em Matemática (257,36 contra 263,29; diferença = -5,94 pontos; $p < 0,001$). O perfil socioeconômico do grupo só mãe no 9º ano também é desfavorável em relação ao biparental: INSE médio de 4,83 contra 5,31, maior concentração em escolas públicas (87% contra 81%) e maior proporção de alunos com distorção idade-série (13% contra 10%). Essas diferenças de composição reforçam, mais uma vez, que a comparação bruta não isola o efeito do arranjo familiar sobre o desempenho.

A Figura 1 apresenta os boxplots das distribuições de proficiência por arranjo familiar para as quatro combinações de série e disciplina. Em todas elas, a mediana e a média (indicadas pelo diamante) do grupo biparental são ligeiramente superiores às do grupo só mãe, confirmando a desvantagem bruta observada nas tabelas descritivas. A amplitude das distribuições é similar entre os grupos, indicando que a variabilidade intragrupo é elevada em ambos os casos.

Gráfico 1 — Distribuição da proficiência por arranjo familiar — SAEB 2023

Distribuição da Proficiência por Arranjo Familiar — SAEB 2023

Nota: Diamante indica a média amostral. Fonte: SAEB 2023/INEP. Elaboração própria.

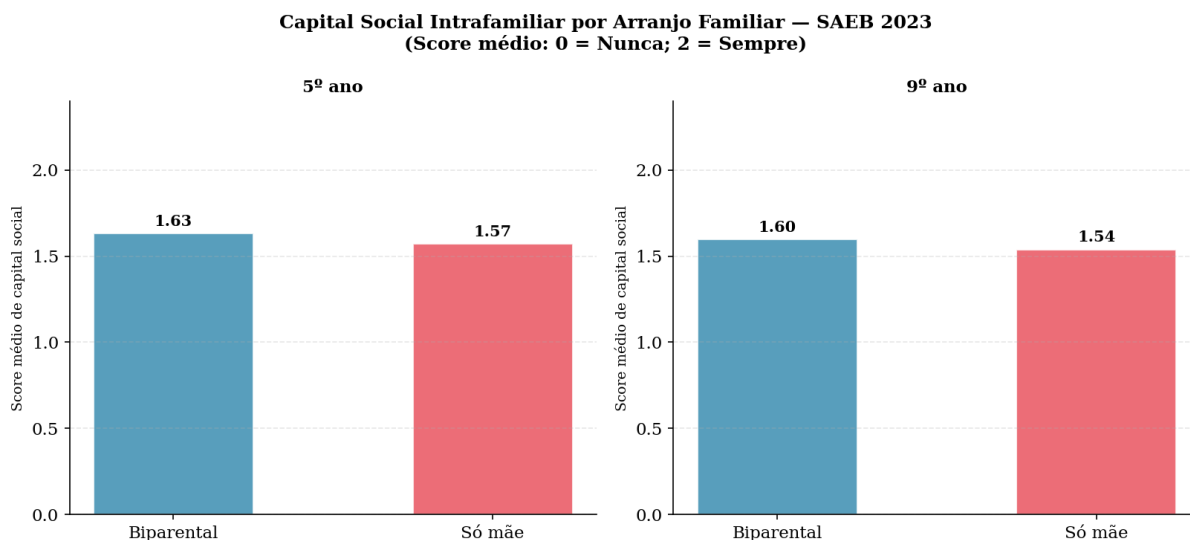
5.2 Capital Social Intrafamiliar por Arranjo Familiar

A Figura 2 apresenta o score médio de capital social intrafamiliar por arranjo familiar, separadamente para o 5º e o 9º anos. O score foi construído como média simples das cinco variáveis do bloco Q10 do questionário do aluno — frequência de conversa sobre a escola, incentivo aos estudos, incentivo à realização das tarefas, incentivo à frequência às aulas e comparecimento às reuniões de pais —, codificadas em escala de 0 (nunca ou quase nunca) a 2 (sempre ou quase sempre), conforme o dicionário do SAEB 2023.

Em ambas as séries, o score médio é ligeiramente inferior no grupo só mãe: 1,58 contra 1,63 no 5º ano, e 1,54 contra 1,60 no 9º ano. Embora a diferença seja pequena em magnitude, ela é estatisticamente significativa em ambos os casos ($p < 0,001$). Esse resultado é consistente com a hipótese de Coleman (1988) de que a compressão de tempo nas famílias monoparentais tende a reduzir a frequência das interações escolares entre responsável e filho. A semelhança entre os scores dos dois grupos indica, porém, que o arranjo familiar per se não determina

grandes diferenças no envolvimento parental — o que desloca a atenção para os mecanismos que explicam como a estrutura familiar se associa ao desempenho escolar.

Gráfico 2 — Score médio de capital social intrafamiliar por arranjo familiar — SAEB 2023



Nota: Score médio das variáveis Q10b–Q10f (0 = nunca ou quase nunca; 2 = sempre ou quase sempre). Fonte: SAEB 2023/INEP. Elaboração própria.

A matriz de correlação entre as cinco variáveis de capital social e as proficiências — apresentada no Apêndice — confirma que não há multicolinearidade grave entre as proxies: a correlação mais elevada entre pares de variáveis Q10 é de 0,508 (entre 'incentiva a estudar' e 'incentiva a fazer tarefa' no 9º ano), valor abaixo do limiar convencional de 0,70. Isso valida a inclusão simultânea das cinco variáveis no modelo. Chama atenção, no 9º ano, o fato de 'incentiva a estudar' e 'incentiva a fazer tarefa' apresentarem correlação negativa com as proficiências (respectivamente $-0,006$ e $-0,020$ com LP; $-0,017$ e $-0,036$ com MT). A explicação mais plausível é causalidade reversa: pais tendem a intensificar o incentivo quando o filho já apresenta dificuldades, de modo que a variável capta a resposta ao baixo desempenho, e não sua causa (COLEMAN, 1988; PALERMO; SILVA; NOVELLINO, 2014).

5.3 Estimações por Blocos de Controles

As Tabelas 3 a 6 apresentam os coeficientes da dummy família só mãe nos quatro modelos estimados em sequência para o 5º e o 9º anos, em Língua Portuguesa e Matemática. O Bloco 1 inclui apenas a variável de interesse; o Bloco 2 acrescenta controles individuais e familiares; o Bloco 3 incorpora variáveis escolares e territoriais; e o Bloco 4 adiciona as proxies de capital social intrafamiliar.

Tabela 3 – Coeficiente da dummy família só mãe por bloco de controles – Língua Portuguesa, 5º ano (SAEB 2023)

	(1)	(2)	(3)	(4)
Família só mãe	−7,381*** (0,269)	−0,217 (0,240)	−1,089*** (0,232)	−0,685*** (0,245)
R ²	0,003	0,148	0,168	0,168
N	1.623.476	1.572.133	1.572.133	1.450.157

Nota: coeficiente da dummy “família só mãe” em cada especificação, com erro padrão entre parênteses. Colunas: (1) sem controles; (2) + individuais e familiares; (3) + escolares e regionais; (4) + capital social. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

Tabela 4 – Coeficiente da dummy família só mãe por bloco de controles – Matemática, 5º ano (SAEB 2023)

	(1)	(2)	(3)	(4)
Família só mãe	−8,540*** (0,278)	−0,125 (0,244)	−1,010*** (0,238)	−0,744*** (0,251)
R ²	0,004	0,166	0,185	0,183
N	1.623.476	1.572.133	1.572.133	1.450.157

Nota: coeficiente da dummy “família só mãe” em cada especificação, com erro padrão entre parênteses. Colunas: (1) sem controles; (2) + individuais e familiares; (3) + escolares e regionais; (4) + capital social. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

Para o 5º ano, os resultados são coerentes com a literatura. No Bloco 1, sem nenhum controle, o coeficiente da dummy só mãe é negativo e altamente significativo: −7,38 pontos em Língua Portuguesa ($p < 0,001$) e −8,54 pontos em Matemática ($p < 0,001$). Essa desvantagem bruta é expressiva e reflete as diferenças de composição socioeconômica entre os grupos, documentadas na seção anterior.

Com a entrada dos controles individuais e familiares no Bloco 2 — notadamente o INSE e a escolaridade materna —, o coeficiente cai abruptamente e perde significância estatística (LP: −0,22, $p = 0,37$; MT: −0,13, $p = 0,61$). Esse resultado é a principal evidência de que a desvantagem observada no Bloco 1 se deve, em grande parte, às piores condições socioeconômicas das famílias só mãe, e não ao arranjo familiar em si. Esse padrão replica diretamente o achado central de Lima, Carvalho e Silva (2021), que com dados do SAEB 2015

também encontraram que o arranjo familiar perde relevância após o controle por características socioeconômicas.

No Bloco 3, com a adição das variáveis escolares e territoriais, o coeficiente recua levemente para terreno negativo e recupera significância: $-1,09$ pontos em LP e $-1,01$ em MT (ambos $p < 0,001$). Esse movimento indica que parte dos estudantes de famílias só mãe frequenta escolas com menor qualidade observável confirmado pela maior concentração em escolas públicas (87% contra 79% no 5º ano), e que esse canal está adicionalmente associado à desvantagem de desempenho. No modelo completo (Bloco 4), após a inclusão das proxies de capital social, os coeficientes se reduzem marginalmente para $-0,69$ pontos em LP e $-0,74$ em MT, ambos significativos ao nível de 1%. A magnitude residual é pequena, mas a persistência da associação negativa após controles completos sugere a existência de canais não captados pelas covariáveis observadas.

Tabela 5 – Coeficiente da dummy família só mãe por bloco de controles – Língua Portuguesa, 9º ano (SAEB 2023)

	(1)	(2)	(3)	(4)
Família só mãe	$-3,405^{***}$ (0,278)	$+2,390^{***}$ (0,231)	$+1,352^{***}$ (0,219)	$+1,304^{***}$ (0,225)
R ²	0,001	0,162	0,179	0,179
N	1.629.804	1.576.009	1.576.009	1.515.117

Nota: coeficiente da dummy “família só mãe” em cada especificação, com erro padrão entre parênteses. Colunas: (1) sem controles; (2) + individuais e familiares; (3) + escolares e regionais; (4) + capital social. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

Tabela 6 – Coeficiente da dummy família só mãe por bloco de controles – Matemática, 9º ano (SAEB 2023)

	(1)	(2)	(3)	(4)
Família só mãe	$-5,937^{***}$ (0,321)	$+1,426^{***}$ (0,265)	$+0,425^*$ (0,249)	$+0,374$ (0,256)
R ²	0,002	0,160	0,188	0,191
N	1.629.804	1.576.009	1.576.009	1.515.117

Nota: coeficiente da dummy “família só mãe” em cada especificação, com erro padrão entre parênteses. Colunas: (1) sem controles; (2) + individuais e familiares; (3) + escolares e regionais; (4) + capital social. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

No 9º ano, o padrão é distinto do 5º ano e merece análise mais cuidadosa. No Bloco 1, o coeficiente bruto é negativo e significativo: $-3,41$ pontos em LP ($p < 0,001$) e $-5,94$ pontos em MT ($p < 0,001$). A desvantagem bruta dos estudantes de famílias só mãe existe, portanto, também no 9º ano. Esse resultado, por si só, já é alinhado à literatura.

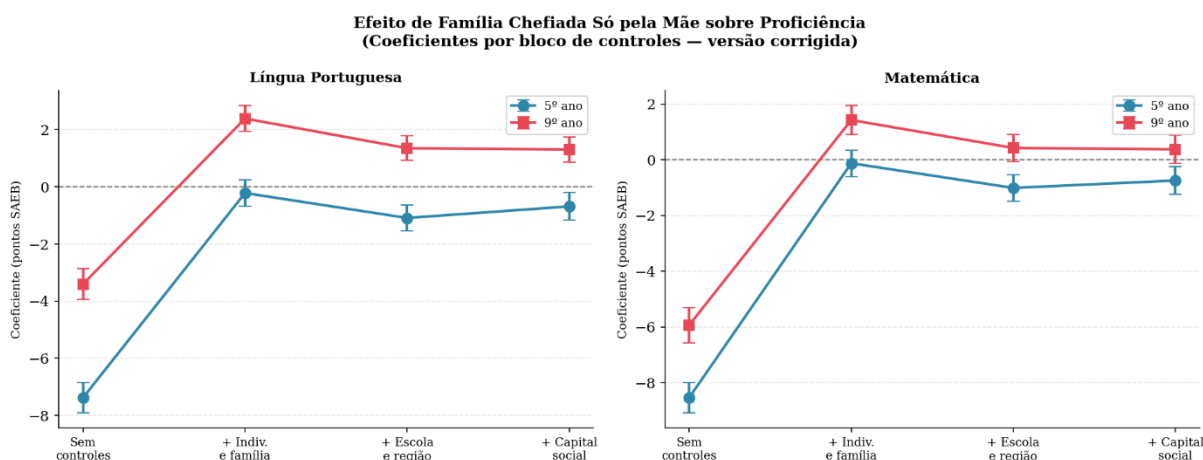
O comportamento singular ocorre no Bloco 2: com a entrada dos controles de INSE e escolaridade materna, o coeficiente inverte de sinal e torna-se positivo — $+2,39$ pontos em LP e $+1,43$ em MT, ambos altamente significativos ($p < 0,001$). Esse fenômeno indica que, ao se comparar estudantes de famílias só mãe e biparentais com as mesmas condições socioeconômicas e o mesmo nível de escolaridade materna, os primeiros apresentam desempenho ligeiramente superior. A inversão de sinal após controles é um achado que, embora não seja o resultado esperado pela literatura, tem interpretação econômica: o coeficiente bruto negativo refletia a diferença de composição socioeconômica entre os grupos; uma vez que essa diferença é controlada, emerge uma relação condicionada positiva.

Uma explicação plausível para esse resultado no 9º ano envolve um mecanismo de seleção amostral. O 9º ano corresponde ao último ano do ensino fundamental, e a chegada até essa etapa requer que o aluno não tenha abandonado o sistema escolar. Estudantes de famílias só mãe em condições socioeconômicas adversas enfrentam maior risco de evasão ao longo da trajetória escolar, de modo que os que chegam ao 9º ano podem representar um grupo mais selecionado — com maior resiliência, motivação ou suporte familiar — do que seus pares biparentais. Esse viés de sobrevivência está documentado na literatura de determinantes da evasão escolar no Brasil (LIMA; CARVALHO; SILVA, 2021) e é uma limitação inerente a análises seccionais com dados de avaliações aplicadas apenas aos presentes.

Nos Blocos 3 e 4, o coeficiente do 9º ano se reduz progressivamente: em LP passa para $+1,35$ (Bloco 3) e $+1,30$ (Bloco 4), mantendo significância ao nível de 1%. Em MT, o coeficiente cai para $+0,43$ no Bloco 3 (significativo apenas a 10%) e $+0,37$ no Bloco 4, sem significância estatística convencional ($p = 0,14$). Para Matemática no 9º ano, portanto, o arranjo familiar não apresenta associação estatisticamente robusta com o desempenho após o controle completo das covariáveis. Para Língua Portuguesa, persiste uma associação positiva pequena mas significativa.

A Figura 3 sintetiza visualmente a trajetória dos coeficientes ao longo dos quatro blocos. Para o 5º ano, os coeficientes partem de valores negativos expressivos, convergem para próximo de zero no Bloco 2 — quando o INSE é controlado — e retornam levemente ao campo negativo nos Blocos 3 e 4. Para o 9º ano, o coeficiente bruto é negativo, inverte de sinal no Bloco 2 e decresce gradualmente até o Bloco 4, sem retornar ao campo negativo em LP. Esse contraste entre as duas séries é a principal contribuição empírica do trabalho.

Gráfico 3 — Evolução do coeficiente da dummy 'família só mãe' por bloco de controles (SAEB 2023)



Nota: Barras verticais representam intervalo de confiança de 95%. Linha tracejada em zero para referência.

Fonte: SAEB 2023/INEP. Elaboração própria.

5.4 Coeficientes Completos do Modelo — Bloco 4

As Tabelas 7 a 10 apresentam os coeficientes completos do modelo no Bloco 4 — o modelo com todos os controles —, separadamente para cada combinação de série e disciplina. Os sinais dos controles são amplamente coerentes com a literatura.

O INSE apresenta coeficientes positivos e altamente significativos em todas as especificações (variando de +4,42 a +8,72 pontos por unidade), confirmando a centralidade do nível socioeconômico como preditor do desempenho educacional (PALERMO; SILVA; NOVELLINO, 2014; LIMA; CARVALHO; SILVA, 2021). A variável de escola pública é negativa e altamente significativa em todos os modelos (−15,65 a −26,07 pontos), refletindo as diferenças estruturais entre redes de ensino, mesmo após o controle pelo INSE individual.

O sexo do aluno apresenta o comportamento esperado: meninas têm desempenho superior em Língua Portuguesa (coeficientes entre +6,99 e +11,94 pontos) e inferior em Matemática (entre −6,25 e −8,41 pontos). A variável de cor ou raça indica desvantagem

consistente para alunos não brancos em todas as especificações (entre -6,70 e -10,89 pontos), resultado que reflete desigualdades estruturais no sistema educacional brasileiro (ALBERNAZ; FERREIRA; FRANCO, 2002). A distorção idade-série associa-se negativamente ao desempenho em todas as especificações (entre -22,42 e -25,99 pontos), o que é esperado dado que alunos com trajetória escolar interrompida por reprovações tendem a apresentar menor proficiência.

Entre as variáveis de capital social, destaca-se o incentivo à frequência às aulas (q10e) como preditor com coeficiente positivo e grande magnitude no 5º ano (+8,12 em LP e +6,89 em MT), declinando para valores menores mas ainda positivos no 9º ano (+2,48 em LP e +1,04 em MT). O comparecimento às reuniões de pais (q10f) também apresenta coeficiente positivo e significativo em todas as especificações (+1,70 a +3,09 pontos). Por outro lado, 'incentiva a estudar' (q10c) e 'incentiva a fazer tarefa' (q10d) apresentam coeficientes negativos especialmente no 9º ano, resultado interpretável como causalidade reversa: pais intensificam o incentivo quando o filho já está com dificuldades escolares (COLEMAN, 1988; PALERMO; SILVA; NOVELLINO, 2014)

Tabela 7 – Coeficientes completos: Língua Portuguesa – 5º ano – Bloco 4 (SAEB 2023))

Variável	Coeficiente
Constante	173,322*** (1,209)
Família só mãe (ref.: biparental)	-0,685*** (0,245)
Feminino (ref.: masculino)	6,995*** (0,205)
Não branco (ref.: branco)	-6,702*** (0,246)
Distorção idade-série	-22,421*** (0,280)
Mãe c/ EM ou mais	15,677*** (0,245)
INSE (nível socioeconômico)	6,731*** (0,145)

Variável	Coefficiente
Escola pública (ref.: privada)	-16,138*** (0,494)
Área urbana (ref.: rural)	7,605*** (0,170)
Região Norte (ref.: Sudeste)	-10,486*** (0,249)
Região Nordeste (ref.: Sudeste)	-5,735*** (0,286)
Região Sul (ref.: Sudeste)	1,640*** (0,261)
Região Centro-Oeste (ref.: Sudeste)	-1,478*** (0,272)
Conversa sobre a escola (Q10b)	0,812*** (0,159)
Incentiva a estudar (Q10c)	0,080 (0,236)
Incentiva a fazer tarefa (Q10d)	-0,225 (0,211)
Incentiva a ir às aulas (Q10e)	8,123*** (0,279)
Vai às reuniões de pais (Q10f)	2,660*** (0,158)

Nota: coeficientes do modelo completo (Bloco 4), com erro padrão entre parênteses. As categorias de referência estão indicadas por “ref.”. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

Tabela 8 – Coeficientes completos: Matemática – 5º ano – Bloco 4 (SAEB 2023)

Variável	Coefficiente
Constante	189,906*** (1,297)

Variável	Coefficiente
Família só mãe (ref.: biparental)	−0,744*** (0,251)
Feminino (ref.: masculino)	−8,413*** (0,209)
Não branco (ref.: branco)	−6,880*** (0,254)
Distorção idade-série	−24,054*** (0,270)
Mãe c/ EM ou mais	15,324*** (0,247)
INSE (nível socioeconômico)	8,718*** (0,148)
Escola pública (ref.: privada)	−15,653*** (0,515)
Área urbana (ref.: rural)	5,775*** (0,170)
Região Norte (ref.: Sudeste)	−12,839*** (0,258)
Região Nordeste (ref.: Sudeste)	−7,782*** (0,299)
Região Sul (ref.: Sudeste)	0,983*** (0,262)
Região Centro-Oeste (ref.: Sudeste)	−3,894*** (0,278)
Conversa sobre a escola (Q10b)	0,072 (0,165)
Incentiva a estudar (Q10c)	−0,485** (0,245)
Incentiva a fazer tarefa (Q10d)	−0,982*** (0,222)
Incentiva a ir às aulas (Q10e)	6,886*** (0,295)

Variável	Coefficiente
Vai às reuniões de pais (Q10f)	3,087*** (0,161)

Nota: coeficientes do modelo completo (Bloco 4), com erro padrão entre parênteses. As categorias de referência estão indicadas por “ref.”. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

Tabela 9 – Coeficientes completos: Língua Portuguesa – 9º ano – Bloco 4 (SAEB 2023)

Variável	Coefficiente
Constante	246,825*** (1,279)
Família só mãe (ref.: biparental)	1,304*** (0,225)
Feminino (ref.: masculino)	11,938*** (0,208)
Não branco (ref.: branco)	-10,894*** (0,238)
Distorção idade-série	-25,987*** (0,223)
Mãe c/ EM ou mais	14,070*** (0,198)
INSE (nível socioeconômico)	4,418*** (0,142)
Escola pública (ref.: privada)	-17,921*** (0,554)
Área urbana (ref.: rural)	6,845*** (0,159)
Região Norte (ref.: Sudeste)	-4,341*** (0,240)
Região Nordeste (ref.: Sudeste)	-2,009*** (0,269)
Região Sul (ref.: Sudeste)	0,500* (0,272)

Variável	Coefficiente
Região Centro-Oeste (ref.: Sudeste)	-1,677*** (0,258)
Conversa sobre a escola (Q10b)	2,035*** (0,179)
Incentiva a estudar (Q10c)	-1,267*** (0,266)
Incentiva a fazer tarefa (Q10d)	-4,553*** (0,189)
Incentiva a ir às aulas (Q10e)	2,479*** (0,315)
Vai às reuniões de pais (Q10f)	1,702*** (0,165)

Nota: coeficientes do modelo completo (Bloco 4), com erro padrão entre parênteses. As categorias de referência estão indicadas por “ref.”. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

Tabela 10 – Coeficientes completos: Matemática – 9º ano – Bloco 4 (SAEB 2023)

Variável	Coefficiente
Constante	259,960*** (1,434)
Família só mãe (ref.: biparental)	0,374 (0,256)
Feminino (ref.: masculino)	-6,249*** (0,237)
Não branco (ref.: branco)	-9,962*** (0,274)
Distorção idade-série	-25,814*** (0,226)
Mãe c/ EM ou mais	12,969*** (0,224)
INSE (nível socioeconômico)	5,931*** (0,161)

Variável	Coefficiente
Escola pública (ref.: privada)	-26,072*** (0,635)
Área urbana (ref.: rural)	2,050*** (0,168)
Região Norte (ref.: Sudeste)	-6,234*** (0,267)
Região Nordeste (ref.: Sudeste)	-2,271*** (0,306)
Região Sul (ref.: Sudeste)	2,006*** (0,304)
Região Centro-Oeste (ref.: Sudeste)	-1,572*** (0,292)
Conversa sobre a escola (Q10b)	1,863*** (0,191)
Incentiva a estudar (Q10c)	-1,082*** (0,289)
Incentiva a fazer tarefa (Q10d)	-4,797*** (0,227)
Incentiva a ir às aulas (Q10e)	1,043*** (0,369)
Vai às reuniões de pais (Q10f)	2,559*** (0,184)

Nota: coeficientes do modelo completo (Bloco 4), com erro padrão entre parênteses. As categorias de referência estão indicadas por “ref.”. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

5.5 Análise de Robustez: Heterogeneidade no Grupo Biparental

Os resultados apresentados nas seções anteriores tratam o grupo biparental como categoria homogênea, reunindo sob um mesmo rótulo arranjos familiares com composições distintas. A variável Q07b do questionário do SAEB 2023 identifica a presença do pai ou padrasto no domicílio sem distinguir entre pai biológico e padrasto. Para investigar se essa heterogeneidade interna ao grupo de referência influencia os coeficientes estimados, utilizou-se a variável Q07c, disponível nos microdados do SAEB 2023, que identifica especificamente

a presença do pai biológico no domicílio. Isso permitiu decompor o grupo biparental em duas subcategorias: biparental biológico (mãe e pai biológico presentes) e biparental reconstituído (mãe e padrasto presentes, sem pai biológico).

A Tabela 11 apresenta as estatísticas descritivas das três categorias de arranjo familiar para o 5º e o 9º anos.

Tabela 11– Descritivas por arranjo familiar detalhado – 5º e 9º anos (SAEB 2023)

Arranjo familiar	N	Prof. LP	Prof. MT	INSE médio
5º ANO				
Biparental biológico	317.938	202,83	214,49	5,09
Biparental reconstituído	1.089.650	217,83	229,02	5,05
Só mãe	325.405	209,46	220,02	4,70
9º ANO				
Biparental biológico	203.527	251,30	249,09	5,15
Biparental reconstituído	1.150.022	259,67	257,55	5,10
Só mãe	364.435	256,79	253,15	4,71

Nota: Biparental biológico = Q07a=B e Q07c=B. Biparental reconstituído = Q07a=B, Q07b=B e Q07c=A. Só mãe = Q07a=B e Q07b=A. Fonte: SAEB 2023/INEP. Elaboração própria.

Os dados da Tabela 11 revelam um padrão relevante: em ambas as séries, o grupo biparental biológico apresenta proficiência média inferior à das famílias só mãe e, especialmente, inferior à das famílias biparentais reconstituídas. No 9º ano, a diferença bruta entre biparental biológico e reconstituído é de 8,37 pontos em LP e 8,46 em MT, com INSE praticamente idêntico entre os dois subgrupos (5,15 vs. 5,10). Esse resultado é consistente com a literatura internacional, que indica que os desfechos educacionais de crianças em famílias reconstituídas são frequentemente similares aos de crianças em famílias monoparentais maternas (GINTHER; POLLAK, 2004), e com Lima, Carvalho e Silva (2021), que já identificaram diferenças de desempenho entre famílias reconstituídas e demais arranjos no SAEB 2015, atribuindo-as ao menor investimento educacional de padrastos em relação a filhos de relacionamentos anteriores.

A Tabela 12 apresenta os coeficientes do modelo com três categorias, estimado com o biparental biológico como categoria de referência, para o 9º ano.

Tabela 12 – Modelo com três categorias de arranjo familiar – 9º ano, Bloco 4 (SAEB 2023)

Variável	Coefficiente
LP - Só mãe vs biparental biológico	+9,552*** (0,433)
LP - Reconstituído vs biparental biológico	+9,060*** (0,409)
MT - Só mãe vs biparental biológico	+10,616*** (0,447)
MT - Reconstituído vs biparental biológico	+11,156*** (0,415)

Nota: modelo com três categorias de arranjo familiar (categoria de referência: biparental biológico), 9º ano, Bloco 4, com erro padrão entre parênteses. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$. MQP com pesos amostrais. Erros padrão HC1 clusterizados por escola. Fonte: SAEB 2023/INEP. Elaboração própria.

Os resultados da Tabela 12 são esclarecedores. Quando o biparental biológico é utilizado como referência, tanto a família só mãe quanto a reconstituída apresentam coeficientes positivos e de magnitude similar: em LP, +9,55 e +9,06 pontos; em MT, +10,62 e +11,16 pontos. A diferença entre só mãe e reconstituída é estatisticamente indistinguível, indicando que os dois arranjos têm desempenho equivalente após controles socioeconômicos.

Esse resultado permite reinterpretar o coeficiente positivo da dummy só mãe observado nos modelos binários da Seção 5.3. A inversão de sinal após controles não reflete vantagem substantiva do arranjo só mãe sobre o biparental em geral, mas sim a heterogeneidade interna ao grupo biparental: ao misturar famílias biológicas com desempenho condicionado mais baixo e famílias reconstituídas com desempenho condicionado mais alto, o grupo de referência assume uma média intermediária que, após os controles, posiciona-se abaixo do grupo só mãe. Do ponto de vista metodológico, esse resultado reforça a importância de tratar a categoria biparental com maior granularidade em estudos futuros, agenda que está além do escopo deste trabalho mas que os dados do SAEB 2023 permitem explorar de forma promissora. Coleman (1988) já havia sinalizado que as obrigações parentais tendem a ser mais fracas entre padrastos do que entre pais biológicos, em função da importância social do parentesco biológico e da institucionalização incompleta das famílias reconstituídas — o que contribuiria para explicar o menor capital social intrafamiliar desse subgrupo e, consequentemente, seus padrões distintos de desempenho escolar.

6 CONSIDERAÇÕES FINAIS

Este trabalho investigou a associação entre a monoparentalidade materna e o desempenho escolar de estudantes do 5º e do 9º anos do ensino fundamental brasileiro, utilizando os microdados do Sistema de Avaliação da Educação Básica de 2023 (SAEB 2023). A partir de uma estratégia empírica baseada na inclusão sequencial de blocos de controles socioeconômicos, familiares, escolares e de capital social intrafamiliar, o estudo buscou desentrelaçar o efeito do arranjo familiar das condições materiais e relacionais do ambiente domiciliar. O conjunto dos resultados converge para uma conclusão central: a estrutura familiar, tomada isoladamente, possui capacidade explicativa limitada sobre o desempenho escolar quando comparada às condições socioeconômicas subjacentes. A monoparentalidade materna funciona, em grande medida, como marcador de vulnerabilidade socioeconômica, e não como determinante autônomo da proficiência escolar.

Os achados relativos ao 5º ano ilustram com clareza essa dinâmica. A desvantagem bruta observada entre estudantes de famílias monoparentais maternas — da ordem de 7,38 pontos em Língua Portuguesa e 8,54 pontos em Matemática em relação às famílias biparentais — desaparece em termos de significância estatística quando se introduzem os controles socioeconômicos individuais e familiares, reduzindo-se a valores próximos de zero. Essa trajetória dos coeficientes ao longo dos blocos de controles é teoricamente informativa: ela indica que a desvantagem inicial capturada pelo modelo irrestrito não reflete uma penalidade do arranjo familiar em si, mas antes a correlação entre monoparentalidade materna e condições de privação socioeconômica. Esse resultado é consistente com a hipótese do capital humano de Becker (1993), segundo a qual os investimentos realizados no ambiente familiar — em termos de tempo, recursos financeiros e estímulos educacionais — são os determinantes primários da formação de habilidades cognitivas nas crianças. Famílias chefiadas exclusivamente por mulheres encontram-se, na média, em posição de maior restrição orçamentária e de tempo disponível para esses investimentos, o que explica a correlação observada no nível bruto. A evidência encontrada dialoga diretamente com Lima, Carvalho e Silva (2021), que, com dados da Prova Brasil de 2015, identificaram padrão análogo de esvaziamento do coeficiente do arranjo familiar após controles socioeconômicos, e reforça a interpretação de que a monoparentalidade atua primordialmente como indicador de vulnerabilidade, não como variável causal independente.

O padrão observado no 9º ano revela uma complexidade adicional que merece interpretação cautelosa. Nessa etapa, após a inclusão dos controles socioeconômicos, o coeficiente da variável de monoparentalidade materna inverte de sinal, tornando-se positivo e estatisticamente significativo em Língua Portuguesa no modelo completo. Esse resultado não deve ser interpretado como evidência de que a monoparentalidade materna produz vantagem educacional, mas sim como indicativo de que fatores não observados pelo modelo atuam de forma assimétrica entre os grupos ao longo da trajetória escolar. A hipótese mais plausível, ainda que não testável diretamente a partir dos dados seccionais disponíveis, envolve viés de seleção amostral decorrente de evasão diferencial: estudantes de famílias monoparentais que permanecem matriculados e avaliados no 9º ano podem constituir um grupo sistematicamente diferente daquele observado no 5º ano, na medida em que a evasão tende a ser mais pronunciada entre os indivíduos mais vulneráveis nas fases iniciais da escolarização. Nesse sentido, o grupo residual poderia apresentar características não observadas de resiliência, motivação ou suporte extrafamiliar que explicariam parte do diferencial positivo condicional. Essa leitura é compatível com a perspectiva de Heckman (2006), que demonstra como desigualdades

acumuladas no início do ciclo de vida inclusive aquelas decorrentes de privações no ambiente familiar tendem a se manifestar de forma heterogênea nas trajetórias de longo prazo, com implicações sobre a composição dos grupos que chegam às etapas finais da escolarização básica. A ausência de dados longitudinais impede qualquer afirmação causal a esse respeito, e análises futuras com painéis individuais seriam necessárias para testar essa hipótese de forma rigorosa.

A dimensão do capital social intrafamiliar, operacionalizada por meio de variáveis de envolvimento parental, revelou-se relevante para a compreensão do desempenho escolar em todas as especificações estimadas. O arcabouço teórico de Coleman (1988) fornece a base interpretativa para esses achados: o capital social intrafamiliar compreendido como o conjunto de obrigações, expectativas, canais de informação e normas de conduta que estruturam as relações entre pais e filhos constitui um recurso que favorece a produção de capital humano nas gerações seguintes. A análise descritiva revelou que famílias monoparentais maternas apresentam, em média, scores ligeiramente inferiores nas variáveis de envolvimento parental, diferença coerente com a restrição de tempo que caracteriza esse arranjo e que, conforme Coleman (1988), tende a comprimir a frequência e a intensidade das interações escolares. Nas estimações, variáveis como o incentivo à frequência às aulas e o comparecimento às reuniões de pais mostraram associação positiva e consistente com a proficiência, confirmando a relevância do acompanhamento ativo, em consonância com Palermo, Silva e Novellino (2014). A presença de coeficientes negativos em algumas variáveis de incentivo especialmente no 9º ano — é interpretada como reflexo de causalidade reversa, e não como evidência de efeito adverso do envolvimento parental: em contextos seccionais, pais tendem a intensificar o acompanhamento precisamente quando o filho já enfrenta dificuldades escolares. Essa limitação não invalida o uso dessas variáveis como proxies de capital social intrafamiliar, mas recomenda cautela na atribuição de sentido causal aos coeficientes estimados. Importa destacar que a inclusão do bloco de capital social não eliminou a associação entre arranjo familiar e desempenho, sugerindo que o envolvimento parental e a estrutura familiar operam, ao menos em parte, por canais distintos sobre a proficiência.

Os achados do presente estudo têm implicações relevantes para o desenho de políticas públicas voltadas à redução das desigualdades educacionais. Se a desvantagem escolar dos estudantes de famílias monoparentais se explica, em sua maior parte, pelas condições socioeconômicas a que esse grupo está sujeito e não pelo arranjo familiar em si, decorre que intervenções centradas exclusivamente na estrutura domiciliar tendem a ser ineficazes ou insuficientes. Políticas de transferência de renda, acesso a serviços de educação infantil de qualidade e ampliação da cobertura das redes de proteção social para famílias chefiadas por mulheres mostram-se mais promissoras do que abordagens que tratam o arranjo monoparental como um fator de risco autônomo a ser corrigido. Essa perspectiva é reforçada pela contribuição de Heckman (2006), que demonstra empiricamente que investimentos na primeira infância especialmente em famílias em situação de vulnerabilidade produzem retornos elevados em termos de habilidades cognitivas e não cognitivas ao longo do ciclo de vida, e que a ausência desses investimentos nos anos iniciais gera déficits dificilmente recuperáveis nas etapas posteriores da escolarização. Nesse sentido, políticas focalizadas em famílias monoparentais maternas com crianças em idade pré-escolar e nos anos iniciais do ensino fundamental seriam prioritárias, dado o maior retorno social dos investimentos realizados nessa janela temporal. Adicionalmente, os resultados sobre envolvimento parental apontam para a utilidade de programas que estimulem e facilitem a participação das mães na vida escolar dos filhos não como substitutos das condições materiais, mas como complementares a políticas de base socioeconômica.

Do ponto de vista metodológico, este trabalho oferece contribuições incrementais em relação à literatura brasileira precedente. A utilização do SAEB 2023 edição mais recente com microdados públicos disponíveis representa uma atualização em relação aos estudos de Lima, Carvalho e Silva (2021) e Martins e Teixeira (2021), baseados em dados de 2015 e 2013, respectivamente. A análise simultânea de duas disciplinas e duas etapas de ensino amplia a cobertura empírica, e a incorporação sistemática de variáveis de envolvimento parental como proxies do capital social intrafamiliar preenche uma lacuna identificada na literatura, que reconhece a relevância teórica dessa dimensão sem explorá-la de forma estruturada em nível nacional. Não obstante essas contribuições, o estudo apresenta limitações que condicionam a interpretação dos resultados. A natureza transversal dos dados do SAEB impede a realização de inferências causais: as associações estimadas expressam correlações condicionais, não efeitos causais, razão pela qual o trabalho adota linguagem compatível com esse limite ao longo de toda a análise. Estratégias de identificação mais robustas como variáveis instrumentais, desenhos de diferenças em diferenças ou estimadores de pareamento seriam necessárias para avançar na atribuição de causalidade. Adicionalmente, a variável de arranjo familiar disponível no SAEB não distingue a origem da monoparentalidade separação, viuvez ou maternidade solo, nem o tempo de exposição ao arranjo, fatores que a literatura internacional identifica como relevantes para os desfechos educacionais (AMATO, 2000; MCLANAHAN; SANDEFUR, 1994). Pesquisas futuras que disponham de dados longitudinais individuais estariam em melhores condições de testar a hipótese de seleção amostral no 9º ano e de investigar os mecanismos causais pelos quais o capital social intrafamiliar atua sobre a proficiência em diferentes etapas da trajetória escolar.

Em síntese, este trabalho contribui para a literatura brasileira sobre estrutura familiar e desempenho escolar ao demonstrar, com dados nacionais recentes, que a monoparentalidade materna não constitui uma desvantagem educacional autônoma, mas opera como marcador de condições de vulnerabilidade socioeconômica que, quando controladas, reduzem ou eliminam o diferencial de desempenho entre os grupos. Essa evidência reposiciona o debate: mais do que o tipo de arranjo domiciliar, são as condições materiais, o acesso a recursos educacionais e o nível de capital social intrafamiliar os fatores que moldam as trajetórias escolares das crianças brasileiras. A compreensão desses mecanismos é condição necessária para a formulação de políticas educacionais que atuem sobre as causas estruturais das desigualdades de aprendizado, em vez de tratar seus sintomas.

REFERÊNCIAS

- ALBERNAZ, Ângela; FERREIRA, Francisco H. G.; FRANCO, Creso. Qualidade e equidade no ensino fundamental brasileiro. *Pesquisa e Planejamento Econômico*, v. 32, n. 3, p. 453-476, 2002.
- ALVES, Fátima. Escolhas familiares, estratificação educacional e desempenho escolar: quais as relações? *Dados*, v. 53, n. 2, p. 447-468, 2010.
- ALVES, Maria Teresa Gonzaga; SOARES, José Francisco; XAVIER, Flavia Pereira. Índice socioeconômico das escolas de educação básica brasileiras. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 22, n. 84, p. 671-704, jul./set. 2014.
- AMATO, Paul R. The consequences of divorce for adults and children. *Journal of Marriage and Family*, v. 62, n. 4, p. 1269-1287, nov. 2000.
- BENATTI, Ana Paula; CAMPEOL, Ângela Rafaela; MACHADO, Maria Salete; PEREIRA, Cíntia Rejane Rodrigues. Famílias monoparentais: uma revisão sistemática da literatura. *Psicologia: Ciência e Profissão*, v. 41, n. spe 3, e209634, 2021.
- BOURDIEU, Pierre. Le capital social: notes provisoires. *Actes de la Recherche en Sciences Sociales*, Paris, v. 31, p. 2-3, jan. 1980.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Microdados do Saeb 2023*: Leia-me. Brasília, DF: Inep, 2024.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Portaria nº 267, de 21 de junho de 2023. Estabelece as diretrizes de realização do Sistema de Avaliação da Educação Básica – Saeb no ano de 2023.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Relatório de amostragem do Saeb 2023*. Brasília, DF: Inep, 2024.
- COLEMAN, James S. Social capital in the creation of human capital. *American Journal of Sociology*, Chicago, v. 94, p. S95-S120, 1988. Supplement.
- GINTHER, Donna K.; POLLAK, Robert A. Family structure and children's educational outcomes: blended families, stylized facts, and descriptive regressions. *Demography*, v. 41, n. 4, p. 671-696, 2004.
- GUJARATI, Damodar N.; PORTER, Dawn C. *Econometria básica*. 5. ed. Porto Alegre: AMGH, 2011.
- HECKMAN, James J. Skill formation and the economics of investing in disadvantaged children. *Science*, v. 312, n. 5782, p. 1900-1902, 2006.
- IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Censo Demográfico 2022*: nupcialidade e família: resultados preliminares da amostra. Rio de Janeiro: IBGE, 2025.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Perfil das mulheres responsáveis pelos domicílios no Brasil: 2000*. Rio de Janeiro: IBGE, 2002.

LIMA, Larissa de Eleterio; CARVALHO, Angelita Alves de; SILVA, Denise Britz do Nascimento. Arranjos familiares e desempenho escolar de alunos do 5º e 9º anos no Brasil em 2015. *Revista Brasileira de Estudos de População*, Rio de Janeiro, v. 38, e0142, p. 1-23, 2021.

MARTINS, Jayne Cecília; TEIXEIRA, Evandro Camargos. As estruturas familiares afetam o desempenho escolar no Brasil? *Revista de Economia do Nordeste*, Fortaleza, v. 52, n. 1, p. 65-76, jan./mar. 2021.

PALERMO, Gabrielle A.; SILVA, Denise Britz do Nascimento; NOVELLINO, Maria Salet Ferreira. Fatores associados ao desempenho escolar: uma análise da proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro. *Revista Brasileira de Estudos de População*, Rio de Janeiro, v. 31, n. 2, p. 367-394, jul./dez. 2014.

PUTNAM, Robert D. *Making democracy work: civic traditions in modern Italy*. Princeton: Princeton University Press, 1993.

RIZZOTTO, Julia Sbroglio; FRANÇA, Marco Tulio Aniceto; FRIO, Gustavo Saraiva. Os arranjos familiares importam no momento de decidir em qual rede de ensino matricular os filhos? *Revista Brasileira de Estudos de População*, v. 35, n. 1, e0066, 2018.

WHITE, Halbert. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, v. 48, n. 4, p. 817-838, 1980.

GLOSÁRIO

Arranjo familiar biparental: composição domiciliar em que a criança reside com ambos os responsáveis, mãe ou madrasta e pai ou padrasto.

Arranjo familiar monoparental: composição domiciliar em que a criança reside com apenas um dos responsáveis. Neste trabalho, refere-se especificamente ao arranjo chefiado exclusivamente pela mãe.

Capital social intrafamiliar: conjunto de relações entre responsáveis e filhos que favorecem a construção de capital humano nas gerações seguintes, incluindo obrigações, expectativas, canais de informação e normas (COLEMAN, 1988).

Dummy: variável binária que assume apenas os valores 0 ou 1, utilizada para representar categorias qualitativas em modelos de regressão.

Erros padrão robustos (HC1): estimadores de variância que permanecem válidos na presença de heterocedasticidade nos resíduos do modelo.

INSE: Indicador de Nível Socioeconômico, calculado pelo INEP a partir de informações sobre bens, serviços e escolaridade dos responsáveis pelos estudantes.

Mínimos Quadrados Ponderados (MQP): método de estimação que atribui pesos distintos às observações, utilizado neste trabalho para incorporar os pesos amostrais do SAEB.

Proficiência: medida de desempenho escolar expressa na escala do SAEB, calculada por meio da Teoria de Resposta ao Item (TRI).

Proxy: variável utilizada para representar indiretamente um conceito não diretamente observável nos dados.

R^2 : coeficiente de determinação; indica a proporção da variância da variável dependente explicada pelo modelo de regressão.

SAEB: Sistema de Avaliação da Educação Básica; avaliação externa em larga escala conduzida pelo INEP para monitorar a qualidade da educação básica brasileira.

TRI (Teoria de Resposta ao Item): metodologia psicométrica utilizada para calcular proficiências a partir das respostas dos estudantes, levando em conta a dificuldade, discriminação e acerto ao acaso de cada item.

APÊNDICES

Apêndice A – Script de análise em Python

O script completo utilizado para leitura dos microdados do SAEB 2023, construção das variáveis, estimação dos modelos de regressão e produção das figuras e tabelas apresentadas neste trabalho foi desenvolvido em Python 3 e está disponível mediante solicitação ao autor. O código utiliza as bibliotecas pandas, numpy, statsmodels e matplotlib, e pode ser reproduzido com os microdados públicos do SAEB 2023 disponibilizados pelo INEP.

ANEXOS

Não se aplica.