

Relação entre características individuais, familiares e escolares com o atraso escolar: uma análise do contexto brasileiro atual

Fransuellen Paulino Santos¹

Resumo

Este trabalho analisa como o contexto familiar e escolar podem afetar a distorção idade-série das crianças e adolescentes, dada a conjuntura atual do país, isto é, considerando o cenário político e que o país enfrentou, assim como o restante do mundo, diversos problemas devido a Covid-19. Utiliza-se os microdados da PNADC de 2018 e 2020 para as 27 Unidades da Federação. O cálculo das probabilidades de distorção idade-série foi realizado pelo modelo logit. Concluiu-se que, o contexto familiar e escolar são fatores importantes a serem considerados na redução da chance de defasagem, assim como algumas características individuais.

Palavras-Chave: distorção idade-série; Covid-19; ambiente escolar.

Area temática: 1 - Economia

1. Introdução

¹ Doutoranda em Economia no programa de pós-graduação do Centro de Desenvolvimento e Planejamento Regional da Universidade Federal de Minas Gerais.

O objetivo deste trabalho é entender o modo como o contexto familiar e escolar podem afetar a distorção idade-série das crianças e adolescentes, dada a conjuntura atual do país, isto é, considerando que o país enfrentou, assim como o restante do mundo, diversos problemas devido a pandemia Covid-19. Além disso, o período de análise (2018 e 2020) selecionado também tem como finalidade pensar em como se comportou o contexto educacional no governo atual, onde observa-se uma redução de investimento em educação. Assim, para além dos problemas vinculados a pandemia houve uma necessidade de adaptação do ensino básico.

O período pandêmico foi e ainda é um cenário de grandes incertezas em diversos aspectos. Sendo um deles relacionados a educação, principalmente, das crianças que estão vivenciando um momento de diversas mudanças. A educação básica é entendida como um direito de todos e assegurada pela Constituição Federal de 1988. O ensino remoto foi uma novidade e improvisado dentro desse novo contexto. A questão é que em tempos “normais” o ensino básico enfrenta algumas dificuldades persistentes, tais como: professores mal remunerados, cargas excessivas de trabalho, além de as escolas serem mal equipadas (PINTO, 2014). Contudo, com a Covid-19, a qualidade da educação, passa a depender fortemente, para além dos insumos e da valorização de seus profissionais, de fatores socioeconômicos (ZAJAC, 2022).

Segundo a pesquisa TIC Domicílios 2020, apenas 45% das casas brasileiras tem computador; 90% dos usuários de internet das classes D e E acessam a rede exclusivamente pelo celular e somente 10% se conectam tanto pelo aparelho móvel quanto pelo computador. Analisando por localização, na área rural apenas 16% conseguem ter acesso aos dois dispositivos, ou seja, 84% dos usuários de internet na zona rural possuem acesso apenas pelo telefone celular. Essa é uma realidade que prejudica a garantia do Direito à Educação em condições normais. Sendo o ano letivo cumprido à distância, a situação só piora.

À vista disso faz-se relevante pensar em como o aluno tende a ser afetado e como características familiares e escolares podem influenciar. Destaca-se que essa é uma análise preliminar e que o intuito é ampliar esse período de análise, considerando também o ano de 2022.

Além desta introdução, este artigo está dividido em outras quatro seções. A segunda seção trata da revisão de literatura, em que são descritos os principais resultados encontrados. A terceira seção retrata as questões metodológicas, tais como as bases de dados, variáveis e método econométrico empregado. Na quarta seção são apresentadas

algumas estatísticas descritivas, seguida da apresentação e análise das estimações. Por fim, a última seção conclui este trabalho.

2. Revisão de literatura

Desde a década de 1960 que as análises sobre os fatores relacionados à qualidade da educação dos indivíduos tem sido foco de diversos estudos na área da economia da educação. Isso ocorreu principalmente com a divulgação do Coleman Report, que tinha como objetivo verificar as distinções no nível de aprendizagem de diferentes grupos da sociedade norte-americana. Os resultados apontavam a importância de diversos fatores, mas os mais relevantes estariam relacionados com o background familiar, enquanto fatores como infraestrutura escolar possuíam pouca influência sobre o desempenho (COLEMAN et al., 1966; MURNANE; WILETT, 2011).

No caso brasileiro, as investigações sobre a desigualdade de oportunidades educacionais tem sido foco de diversos estudos, de modo que existe um bom entendimento sobre os principais padrões destas desigualdades (RIBEIRO, 2011; MONTALVÃO, 2011). Consolidou-se a ideia de existência de desigualdades entre indivíduos com origem social nos meios urbano e rural e de desigualdade de raça, principalmente nos níveis educacionais mais altos. De acordo com Ribeiro (2014), a desigualdade relacionada com a origem de classe dos indivíduos é persistente ao longo do tempo.

O Brasil passou por significativas mudanças em seu sistema educacional nas últimas décadas. Sendo que a partir da atual Lei de Diretrizes e Bases da Educação Nacional, Lei nº 9.394/96, a educação infantil e o ensino fundamental e médio passaram a compor a Educação Básica. O ensino fundamental era a única etapa considerada obrigatória na educação nacional até o ano de 2009. Ao passo que, essa condição foi alterada pela Emenda Constitucional (EC) nº 59/2009 que amplia a obrigatoriedade para os indivíduos com idade entre 4 e 17 anos.

Alguns estudos apontam para a existência de uma relação entre repetência, abandono e evasão. Os trabalhos realizados por Ribeiro (1991), Machado (2005) e Fritsch et. al (2014) afirmam que a repetência é o principal problema que requer atenção, pois há um aumento das chances de o aluno evadir posteriormente. Barros e Mendonça (1998) afirmam que as reprovações, apresentam um efeito negativo tanto para a autoestima dos estudantes, aumentando a probabilidade de reprovações subsequentes, quanto para o estado, pois há maiores gastos com as reprovações. Além disso, alguns trabalhos mostraram que a defasagem também afeta o desempenho do estudante quando

comparados àqueles que estão na idade adequada para a série (FERRÃO et al., 2001; FRANCO, 2008; SOARES; SATYRO, 2008; POZZOBON et al., 2017).

As oportunidades escolares possuem uma distribuição desigual, devido às diferentes categorias sociais e isso continua sendo verificado (MONT'ALVÃO, 2014; RIBEIRO et al., 2015). De modo que, essas desvantagens repercutem durante a trajetória social dos indivíduos, não sendo uma simples desigualdade de origem. Isso ocorre, principalmente, entre aqueles que provêm das camadas populares e dos grupos raciais negros, que possuem desvantagens mais acentuadas em relação aos demais grupos sociais cujas posições socioprofissionais dependem de elevados graus de escolarização (LOURO, 2017).

A família também exerce influência sobre a trajetória do estudante, inclusive na entrada da vida adulta. Assim, é usual atribuir à família o papel preponderante na qualidade de vida de seus membros e também em suas decisões relativas à saúde, moradia, educação e participação no mercado de trabalho. Assim, tende a ter um papel de influência na forma como a criança socializa, que pode estar diretamente relacionada com comportamento escolar, atitudes, valores, aspirações futuras e aspectos emocionais (LEONE et al, 2010; SEBEN, 2010).

Segundo Ribeiro (2004), uma das formas de a família interferir nas transições para a vida adulta é por meio das transmissões intergeracionais. As características familiares em que os adolescentes e jovens são criados influenciam de diversas maneiras suas chances de progressão e permanência no sistema educacional. De modo que, aqueles que são de classes mais baixas tendem a permanecer menos tempo na escola e começar a trabalhar mais cedo. Ao passo que, os de classes mais altas pendem a estudar por mais tempo. Mas, o autor destaca que a maioria dessas diferenças se reduziram significativamente ao longo das décadas.

Ao longo dos anos ocorreram algumas mudanças na organização da família, tais como as configurações familiares, o tamanho da família e as relações entre os membros que a compõe. Essas mudanças, causaram importantes impactos sobre os recursos disponíveis para as crianças, tanto dentro da família, quanto dentro da sociedade, pois também modificou o tamanho das coortes de nascimento e a estrutura etária das populações (LAM; MARTELETO, 2008).

Rios-Neto et al. (2002) realizaram uma análise considerando as características individuais e das escolas, e observaram que com a inclusão de características da escola, há uma redução da importância das características da família. Neste estudo, com a inclusão de variáveis sobre a escolaridade média local dos professores do ensino

fundamental ocorreu a redução do coeficiente da variável educação da mãe, que afeta positivamente a probabilidade de não defasagem. Isso pode ser o indicativo de um efeito substituição entre estes dois aspectos. Barros et al. (2001) também aponta para a importância do nível educacional da mãe para determinar o aumento/redução na distorção idade-série.

Nesse conjunto de características familiares, os estudos de Machado e Gonzaga (2007) e Pontili e Kassouf (2008) encontraram que quanto maior a renda, maior idade e a escolaridade dos pais menor a probabilidade de distorção idade-série dos filhos. Esses autores também analisaram as características individuais, assim ser do sexo masculino e não branco apresentam um nível maior de probabilidade à defasagem idade-escolaridade.

Existem muitos estudos sobre essa temática, mas esta seção buscou realizar uma breve revisão de literatura sobre o tema. Assim, percebe-se a importância da inserção das características da família e da escola nesta análise, bem como das características dos indivíduos. Assim, pode-se perceber que existem diversas consequências na ocorrência da defasagem idade-série (abandono, evasão e baixo desempenho) e que os grupos da população são afetados de maneira distinta. A partir dessa contextualização, esse trabalho tem como objetivo estudar o modo como o contexto familiar e escolar podem afetar a distorção idade-série das crianças e adolescentes nos anos de 2018 e 2020.

3. Metodologia

Nesta seção, são expostas inicialmente o método econométrico utilizado e, num segundo momento, as definições adotadas para a obtenção das variáveis de interesse.

3.1 Modelo *Logit*

Tendo em vista que o objetivo deste trabalho é entender o efeito das características familiares e escolares sobre a distorção idade-série, utiliza-se de um método econométrico que possa analisar de que modo tais características tendem a afetar a probabilidade de o aluno repetir alguma etapa do ensino ou não, já que a variável dependente em questão é do tipo binária. Sendo assim, recorre-se a modelos de escolha qualitativa, neste caso optou-se pelo modelo *logit* que, permite estimar a probabilidade de defasagem com base em uma função de distribuição logística. Cabe destacar que os coeficientes do modelo *logit* são estimados a partir do método da máxima verossimilhança.

Uma definição de defasagem idade-série é sugerida por Machado e Gonzaga (2007, p. 456): “a criança é considerada atrasada em termos educacionais se não tem o total de anos de estudo completos compatível com a sua idade no início de cada ano

letivo”. Assim, a variável dependente da equação de regressão é relacionada com a probabilidade de ocorrer um de dois possíveis eventos, 1 para distorção idade-série e 0 caso contrário. Logo, os estudantes considerados com distorção idade-série são aqueles que estão acima da idade ideal em uma determinada série. No Brasil, o ingresso no ensino fundamental acontece aos seis anos completos e tem duração de nove anos. Portanto, espera-se que a criança termine essa etapa escolar aos quatorze anos e, assim, sucessivamente.

Esse método tem como principal interesse a probabilidade de resposta, ou seja:

$$P(y = 1|x) = P(y = 1|x_1, x_2, \dots, x_k) \quad (1)$$

onde x representa o conjunto de variáveis independente selecionadas. Deste modo, a variável dependente pode assumir dois valores (0 ou 1) e para que isso ocorra, assume-se a probabilidade de resposta linear em um conjunto de parâmetros β_j , conforme descrito na equação 2.

$$P(y = 1|x) = F(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k) = G(\beta_0 + x\beta) \quad (2)$$

sendo que F é uma função que assume valores estritamente entre zero e um para todos os números z reais; e $x\beta = \beta_1 x_1 + \dots + \beta_k x_k$.

Desta forma, a função de distribuição cumulativa de uma variável aleatória logística padrão F será representada conforme demonstrado na equação 3.

$$F(z) = \frac{\exp(z)}{[1+\exp(z)]} = \Delta(z) \quad (3)$$

A interpretação dos coeficientes β ocorre por meio da razão de chances, isto é, demonstram a relação entre as probabilidades de sucesso ou de fracasso em relação ao estado de referência. Como exposto por Cirino (2008), os coeficientes para modelos de escolha binária não possuem significado econômico relevante. Portanto, as informações neste trabalho são analisadas as probabilidades e os efeitos marginais das variáveis de interesse.

3.2 Fonte de dados e variáveis analisadas

Os dados utilizados são oriundos da Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) que consiste de uma amostra representativa da população brasileira, para os anos de 2018 e 2020. A população de interesse é formada pelos estudantes em idade escolar matriculados no ensino básico regular público e privado no Brasil. Dado o novo sistema de ensino básico, com 12 anos de estudo e ingresso aos 6 anos completos no primeiro ano, são considerados defasados aqueles estudantes que, com 7 anos de idade não tinham um ano completo de escolaridade, e assim sucessivamente. A amostra

utilizada é composta de todas as crianças e adolescentes de 7 a 17 anos de idade, matriculadas em escolas da rede pública ou privada de todos os estados brasileiros. Assim, no caso de crianças e adolescentes que estão fora do sistema escolar, a discussão sobre distorção idade-escolaridade não faz sentido, pois nesse caso o problema é que eles deixaram de estudar permanentemente e não simplesmente que eles estão atrasados. Esses dois problemas possuem uma ligação, conforme a literatura aponta, mas não são iguais, isto é, cada um possui sua especificidade e deve ser analisado individualmente.

Cabe destacar que a variável dependente está mais relacionada às repetências do que ao abandono, uma vez que são consideradas apenas as crianças e jovens que estão matriculadas. Além disso, o recorte de idade escolhido se justifica pelo formato da variável dependente, pois é apenas neste intervalo que é possível a existência de defasagem idade-série positiva ou zero, sendo que para estudantes com menos de 7 anos ou mais de 17, a defasagem idade-série será sempre nula ou positiva, respectivamente.

As variáveis independentes utilizadas nas estimações foram escolhidas de duas fontes, a primeira relativa as características do aluno e do núcleo familiar foram extraídas da PNADC, enquanto que as variáveis relacionadas ao ambiente escolar foram retiradas dos Indicadores Educacionais do INEP para os dois anos. Essas variáveis são disponibilizadas pelo INEP, por Unidade da Federação e a rede de ensino. Desta forma, como a totalização ocorre em dois períodos para o ensino fundamental (anos iniciais e finais) e também para o ensino médio, cada indivíduo recebeu os valores relativos à sua unidade da federação, esfera de ensino e período (inicial do ensino fundamental, final do ensino fundamental ou ensino médio).

As variáveis selecionadas nos microdados da PNADC são: idade; cor de pele; branco (cor de pele - brancos = 1 e não-brancos = 0, sendo amarelos enquadrados como brancos); mulher (sexo feminino = 1 e masculino = 0); urbano - se reside na zona urbana ou rural (urbano = 1 e rural = 0); controles por região (com cinco *dummies* referentes às regiões Norte, Nordeste, Sudeste, Sul, tendo como resultado base o da região Centro-Oeste); se a mãe mora ou não no mesmo domicílio (se mora = 1, se não mora = 0, o que inclui mães que já morreram ou que não se tem conhecimento); a escolaridade da mãe; o número de membros da família; logaritmo da renda familiar per capita; se esse domicílio recebe auxílio Bolsa Família; o número de pessoas ocupadas da família; número de pessoas com educação de nível superior na família; número de crianças na família. As variáveis relativas às características das escolas, são: média de alunos por turma, número de horas-aula diárias e o percentual de professores com ensino superior.

A análise será realizada separando entre rede de ensino público e privada, dado que tende a ter características distintas. Além disso, os resultados são apresentados em três modelos para cada ano (2018 e 2020). O intuito disso é verificar como essas variáveis podem influenciar no impacto uma das outras, verificar se há efeito substituição, conforme apontado por Rios-Neto et al. (2002). Assim, são realizadas as seguintes estimações: modelo (1) conta com características individuais e de localização do aluno; modelo (2) adiciona as características familiares; e modelo (3) acrescenta as informações do ambiente escolar.

4. Resultados e discussões

Os resultados serão analisados para os anos de 2018 e 2020, foram considerados, respectivamente, um total de 63.348 e 53.255 observações em cada ano. Mesmo com a melhoria nos níveis educacionais no Brasil ao longo dos últimos anos, a distorção idade-série ainda pode ser considerada elevada. Para o ano de 2018, 41,9% dos indivíduos com idade entre 7 e 17 anos apresentam defasagem, conforme pode ser observado na Tabela 1. Essa porcentagem apresenta uma pequena redução, passando para 40,9% em 2020.

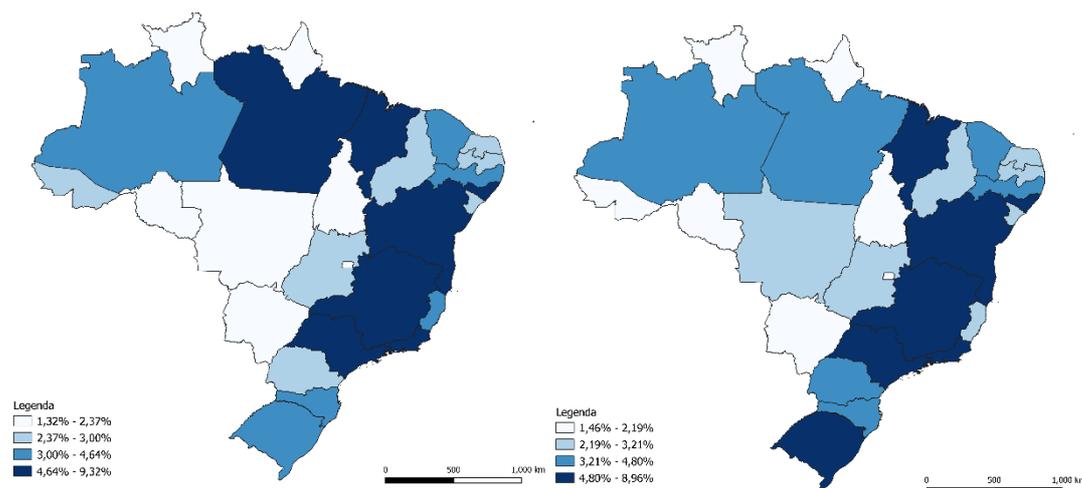
Tabela 1 – Percentual de pessoas com e sem distorção idade-série

Categoria	2018	2020
Defasado	41,9%	40,9%
Não defasado	58,1%	59,1%
Tamanho da amostra	68.348	53.255

Fonte: Elaborado a partir dos microdados da PNADC.

A Figura 1 apresenta a distribuição do percentual de indivíduos defasados nas amostras por unidade da federação para cada ano analisado. Analisando os mapas é possível observar que não correram muitas mudanças na distribuição das distorções idade-série. Nota-se que houve uma redução na porcentagem de alunos defasados no Acre, Pará e Espírito Santo. Em contrapartida, verifica-se um aumento de estudantes fora da série recomendada nos estados do Mato Grosso, Paraná e Rio Grande do Sul.

Figura 1 – Mapas da distribuição da distorção idade-série por Unidade da Federação



Fonte: Elaborado a partir dos microdados da PNADC.

Em relação às estatísticas descritivas das variáveis independentes, pela Tabela 2, observa-se que em 2018 cerca de 42,68% da amostra era do sexo feminino, com uma média de 13 anos de idade, com 29,68% de cor branca ou amarela. A grande maioria residia com a mãe (71,43%) em área urbana (64,20%). A média do logaritmo da renda per capita familiar foi 5,81 e apenas 0,59% recebia bolsa família. Sobre as pessoas do domicílio, em média 1 está ocupada, sendo que 12,63% possui ensino superior e em média os domicílios possuem 2 crianças. Em relação às variáveis do ambiente escolar, pode-se constatar que a média de horas-aula diária foi de 4,52, enquanto foram 23 alunos em média por turma. O percentual de docentes com ensino superior foi de 80,36%.

A Tabela 2 também apresenta as médias e desvios padrão das variáveis independentes para o ano de 2020. Como pode-se visualizar, não ocorreram muitas mudanças. Novamente, observa-se que 57,08% da amostra era do sexo feminino, com uma média de 13 anos de idade, sendo que 30,78% da cor branca ou amarela. Novamente, a amostra majoritariamente residia com a mãe, cerca de 89,52% e em área urbana (67,53%). Quanto as variáveis que caracterizam o ambiente escolar, a média de horas-aula diária foi de 4,58, enquanto foram 23 alunos em média por turma. O percentual de docentes com ensino superior foi de 87,08%.

Tabela 2 – Estatísticas descritivas da amostra

Variáveis	2018		2020	
	Média	Desvio Padrão	Média	Desvio Padrão
Idade	13,3209	0,0170	13,3107	0,0197
Branco	0,2968	0,0027	0,3078	0,0031
Mulher	0,4268	0,0029	0,5708	0,0034
Urbano	0,6420	0,0028	0,6753	0,0032

Norte	0,1897	0,0023	0,1792	0,0026
Nordeste	0,4078	0,0029	0,3894	0,0033
Sudeste	0,2039	0,0024	0,2076	0,0027
Sul	0,1167	0,0019	0,1319	0,0023
Centro-Oeste	0,0819	0,0016	0,0919	0,0019
Mora com a mãe	0,7143	0,0027	0,8952	0,0021
Tamanho da família	4,7509	0,0107	4,5868	0,0115
Logaritmo da renda <i>per capita</i> familiar	5,8149	0,0060	6,0639	0,0061
Bolsa Família	0,0059	0,0005	0,0025	0,0003
Pessoas ocupadas no domicílio	1,4061	0,0061	1,2799	0,0068
Pessoas com ensino superior no domicílio	0,1263	0,0025	0,1702	0,0033
Número de crianças no domicílio	2,3339	0,0083	2,1876	0,0087
Média de alunos por turma	23,2534	0,0143	23,1625	0,0174
Horas-aula	4,5153	0,0019	4,5760	0,0026
Docentes com ensino superior	80,3553	0,0660	87,0815	0,0623

Fonte: Elaborado a partir dos microdados da PNADC.

A Tabela 3 apresenta os resultados da estimação do modelo logit para os alunos que frequentam o ensino básico na rede pública para os anos de 2018 e 2020. Conforme exposto o modelo (1) é o mais básico, composto apenas por características individuais e de localização. Assim, nota-se que para o ano de 2018, todas as variáveis foram estatisticamente significativas, enquanto no ano de 2020, somente a dummy que representa a região Sul não foi estatisticamente significativa.

Na segunda estimação, onde são adicionadas as características do núcleo familiar, para ambos os anos é possível observar que algumas variáveis que eram significativas deixaram de ser. Tanto em 2018 quanto em 2020, as variáveis que indicam que o aluno mora com a mãe, educação da mãe log da renda per capita familiar, de número de pessoas com ensino superior e de número de crianças no domicílio foram estatisticamente significativas.

Por fim, a terceira estimação, o modelo completo com as características do ambiente escolar é apresentado para os dois anos. Para o ano de 2018, constata-se que quanto mais velha a criança ou adolescente maior será a chance de ele estar defasado. Enquanto isso, ser mulher e morar com a mãe reduzem a probabilidade de estar defasado, esse resultado corrobora com os estudos expostos na seção de revisão bibliográfica (RIOS-NETO et al., 2002; MACHADO; GONZAGA, 2007; PONTILI; KASSOUF, 2008). O nível de educação da mãe também auxilia na redução da chance de defasagem do estudante.

Outras características da família que apresentaram impacto na probabilidade de distorção idade-série é o log da renda per capita familiar, pessoas com ensino superior e

número de crianças no domicílio. Sendo que as duas primeiras reduzem as chances de defasagem e a última aumenta. Assim, quanto maior a renda, mais pessoas com ensino superior e menor número de crianças no domicílio menor será a probabilidade de defasagem.

Quanto a localização, ser da área urbana não impacta na probabilidade de estar defasado, mas os indivíduos que estudam na rede de ensino pública e que são das regiões Norte e Sudeste apresentam maior probabilidade de estar fora da série correta em comparação com os do Centro-Oeste. Enquanto isso, os estudantes do Nordeste e Sul do país possuem menor chance de estar defasado quando comparados aos do Centro-Oeste.

Percebe-se através dos efeitos marginais que há uma correlação negativa entre a média de alunos na turma, a porcentagem de docentes com ensino superior e a probabilidade distorção idade-série. Dessa forma, quanto maior média de alunos por turma e maior a quantidade de docentes com ensino superior menor serão as chances daquele aluno estar com atraso escolar. O primeiro resultado é contraintuitivo, onde um aumento no tamanho das turmas reduz a chance de ocorrência de defasagem, esperava-se o contrário. Porém, pode estar ocorrendo o que foi apontado por Portella et al. (2017). Os autores explicam que pode ser que em turmas muito pequenas esse evento de distorção não ocorra, de modo que exista um tamanho ótimo de turma.

Para o ano de 2020 os resultados são bem semelhantes em termos de sinal e significância estatística. Entretanto, algumas questões merecem atenção. Por exemplo, neste modelo (3) de 2020, ser mulher passou a ser uma característica que aumenta a probabilidade de defasagem. Ou seja, em 2020 as meninas passaram a ter uma maior distorção idade-série quando comparadas aos meninos. Isso pode estar relacionado a várias questões, inclusive em alguns casos, podem estar relacionadas ao fato de que as meninas tendem a ajudar mais os pais em casa, desde a afazeres domésticos a auxiliar no cuidado com outros irmãos. Então, com o momento de pandemia, isso pode ter afetado o tempo de dedicação aos estudos desse grupo.

Tabela 3 - Efeitos marginais a partir do modelo *Logit* para os anos de 2018 e 2020 para a rede pública

Variáveis	Ano = 2018			Ano = 2020		
	(1)	(2)	(3)	(1)	(2)	(3)
Idade	0,0258*** (0,0010)	0,3170*** (0,0040)	0,3559*** (0,0048)	0,0226*** (0,0012)	0,3211*** (0,0045)	0,3644*** (0,0054)
Branco	-0,0636*** (0,0067)	-0,0224** (0,0113)	-0,0126 (0,0115)	-0,0511*** (0,0080)	-0,0149 (0,0124)	0,0013 (0,0128)
Mulher	-0,0906*** (0,0057)	-0,0699*** (0,0095)	-0,0659*** (0,0096)	0,0904*** (0,0069)	0,0728*** (0,0109)	0,0680*** (0,0109)
	-0,0449***	-0,0009	0,0094	-0,0311***	-0,0120	0,0019

Urbano	(0,0056)	(0,0096)	(0,0097)	(0,0072)	(0,0116)	(0,0116)
Norte	0,0900*** (0,0106)	0,0659*** (0,0187)	0,0372** (0,0197)	0,0406*** (0,0124)	0,0100 (0,0188)	-0,0461** (0,0186)
Nordeste	0,0756*** (0,0095)	0,0144 (0,0162)	-0,0432** (0,0186)	0,0469*** (0,0111)	0,0074 (0,0170)	-0,0807*** (0,0189)
Sudeste	-0,0215** (0,0100)	-0,0216 (0,0163)	0,0474*** (0,0169)	-0,0417*** (0,0117)	-0,0252 (0,0178)	-0,0172 (0,0176)
Sul	0,0227** (0,0110)	0,0320* (0,0181)	-0,0335** (0,0175)	0,0014 (0,0126)	-0,0003 (0,0185)	-0,0930*** (0,0168)
Mora com a mãe		-0,0588*** (0,0214)	-0,0511** (0,0214)		-0,0664*** (0,0237)	-0,0593** (0,0243)
Escolaridade da mãe		-1,1464*** (0,0148)	-1,1216*** (0,0168)		-1,1479*** (0,0168)	-1,1233*** (0,0191)
Tamanho da família		-0,0002 (0,0054)	0,0036 (0,0054)		0,0038 (0,0058)	0,0073 (0,0058)
Logaritmo da renda per capita familiar		-0,0349*** (0,0069)	-0,0307*** (0,0070)		-0,0283*** (0,0083)	-0,0281*** (0,0082)
Bolsa Família		0,1032 (0,0728)	0,1082 (0,0708)		0,0358 (0,0927)	0,0715 (0,1044)
Pessoas ocupadas no domicílio		-0,0013 (0,0056)	0,0004 (0,0057)		-0,0103 (0,0066)	-0,0102 (0,0066)
Pessoas com ensino superior no domicílio		-0,0579*** (0,0129)	-0,0540*** (0,0131)		-0,0551*** (0,0122)	-0,0543*** (0,0122)
Número de crianças no domicílio		0,0221*** (0,0069)	0,0147** (0,0071)		0,0185** (0,0081)	0,0105 (0,0081)
Média de alunos por turma			-0,0532*** (0,0019)			-0,0653*** (0,0024)
Horas-aula			0,0218 (0,0154)			0,0989*** (0,0158)
Docentes com ensino superior			-0,0031*** (0,0007)			-0,0028*** (0,0009)
Teste de significância global do modelo – Teste de Wald						
Valor calculado	1389,08	7137,78	5825,22	729,82	6083,52	5006,25
Prob>Chi²	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

*** Significativo a 1%; ** Significativo a 5%; * Significativo a 10%; Desvio-padrão entre parênteses.

Fonte: Elaborado a partir dos microdados da PNADC.

Quanto a localização, a região Sudeste deixou de ser significativa, conforme apresentado na Tabela 3. Enquanto as demais, em comparação com o Centro-Oeste, reduzem a probabilidade de defasagem. Outra variável que não apresentou significância estatística para o ano de 2020 foi o número de crianças no domicílio. Por fim, a variável que indica a média de horas-aula passou a ser significativa e afetando positivamente a chance de distorção idade-série. O percentual de docentes com ensino superior reduz essa probabilidade, assim, uma hipótese levantada por Portella et al. (2017) é de que a dedicação do docente, dada pelo número de horas-aula e por sua formação, tende a ser um benefício para o aluno, mas também pode elevar o nível de exigência. Isto é, ocorre

uma melhora no ambiente de aprendizado e ao mesmo tempo elevasse o nível de conhecimento exigido dos estudantes.

A Tabela 4 apresenta as estimações dos modelos logit para a rede de ensino privada. No geral, as tendências são bem semelhantes. Mas, cabe destacar que algumas variáveis deixaram de ser significativas, tais como: em 2018, se o indivíduo está localizado nas regiões Norte ou Sudeste, se mora com a mãe, log da renda per capita familiar e número de crianças no domicílio; em 2020, se o indivíduo mora com a mãe, a média de horas-aula e o percentual de docentes com ensino superior.

No ano de 2018 observa-se que os fatores que impactam positivamente na probabilidade de distorção idade-série são idade e tamanho da família. Portanto, isso indica que quanto mais velho o estudante, maior a chance de defasagem. Bem como famílias maiores tendem a elevar a ocorrência de defasagem para os indivíduos que estudam na rede de ensino privada. O que pode estar relacionado a atenção dos responsáveis aos estudos da criança ou adolescente, devido a maior cobrança das escolas da rede privada. Ser mulher e estar localizada na área urbana reduz a probabilidade de defasagem. A escolaridade da mãe e o número de pessoas com ensino superior no domicílio também reduzem as chances, o mesmo é observado para as características do ambiente escolar, média de alunos por turma e docentes com ensino superior.

O modelo completo para o ano de 2020 da rede de ensino privada apresentou pouca diferença para o que foi encontrado na rede de ensino pública. Novamente destaca-se que no ano de 2020 ser mulher passa a ser um fator que afeta positivamente as chances de distorção idade-série.

Tabela 4 - Efeitos marginais a partir do modelo *Logit* para os anos de 2018 e 2020 para a rede privada

Variáveis	Ano = 2018			Ano = 2020		
	(1)	(2)	(3)	(1)	(2)	(3)
Idade	-0,0049** (0,0019)	0,1075*** (0,0053)	0,1134*** (0,0057)	-0,0050** (0,0022)	0,1246*** (0,0058)	0,1320*** (0,0068)
Branco	-0,0194 (0,0136)	0,0071 (0,0124)	0,0116 (0,0120)	-0,0259* (0,0152)	0,0013 (0,0132)	0,0070 (0,0127)
Mulher	-0,0356*** (0,0124)	-0,0199** (0,0100)	-0,0188** (0,0097)	0,0558*** (0,0143)	0,0278** (0,0121)	0,0247** (0,0114)
Urbano	-0,0229 (0,0237)	-0,0326* (0,0210)	-0,0366** (0,0223)	-0,0026 (0,0319)	0,0094 (0,0275)	0,0113 (0,0279)

Norte	0,0274 (0,0259)	-0,0011 (0,0199)	-0,0190 (0,0165)	-0,0469* (0,0242)	-0,0424** (0,0177)	-0,0651*** (0,0124)
Nordeste	0,0386** (0,0208)	-0,0080 (0,0161)	-0,0385** (0,0150)	0,0198 (0,0223)	-0,0143 (0,0186)	-0,0514*** (0,0175)
Sudeste	0,0341* (0,0200)	0,0031 (0,0164)	0,0162 (0,0162)	-0,0014 (0,0214)	-0,0241 (0,0179)	-0,0120 (0,0174)
Sul	-0,0189 (0,0213)	-0,0212 (0,0146)	-0,0388*** (0,0123)	-0,0329 (0,0231)	-0,0255 (0,0178)	-0,0709*** (0,0118)
Mora com a mãe		-0,0228 (0,0234)	-0,0275 (0,0243)		-0,0269 (0,0376)	-0,0179 (0,0339)
Escolaridade da mãe		-0,4005*** (0,0198)	-0,3573*** (0,0213)		-0,4755*** (0,0212)	-0,4040*** (0,0234)
Tamanho da família		0,0156*** (0,0058)	0,0172*** (0,0057)		0,0118 (0,0073)	0,0120* (0,0071)
Logaritmo da renda <i>per capita</i> familiar		-0,0024 (0,0082)	-0,0005 (0,0077)		-0,0210** (0,0084)	-0,0142* (0,0083)
Bolsa Família		-0,0391 (0,0640)	-0,0371 (0,0618)		0,3239 (0,2050)	0,3398 (0,2321)
Pessoas ocupadas no domicílio		-0,0027 (0,0062)	-0,0013 (0,0061)		-0,0009 (0,0083)	-0,0001 (0,0079)
Pessoas com ensino superior no domicílio		-0,0188*** (0,0065)	-0,0163*** (0,0062)		-0,0192** (0,0081)	-0,0184** (0,0075)
Número de crianças no domicílio		-0,0048 (0,0077)	-0,0076 (0,0075)		0,0037 (0,0098)	0,0024 (0,0097)
Média de alunos por turma			-0,0176*** (0,0024)			-0,0257*** (0,0042)
Horas-aula			0,0195 (0,0167)			-0,0121 (0,0270)
Docentes com ensino superior			-0,0018*** (0,0007)			-0,0008 (0,0009)
Teste de significância global do modelo – Teste de Wald						
Valor calculado	43,39	910,87	847,62	40,23	828,71	828,21
Prob>Chi ²	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

*** Significativo a 1%; ** Significativo a 5%; * Significativo a 10%; Desvio-padrão entre parênteses.

Fonte: Elaborado a partir dos microdados da PNADC.

Cabe destacar que todas as estimações obedecem a propriedades estatísticas consistentes. No final das Tabelas 3 e 4 são apresentados os resultados do teste de Wald para identificar a significância global do modelo. Para todos os anos verifica-se que ao nível de significância de 1% rejeita-se a hipótese nula, e assim a relação estudada é significativa, isto é, os modelos estariam bem especificados.

4. Considerações finais

O objetivo deste artigo é compreender o modo como o contexto familiar e escolar tendem a afetar a distorção idade-série das crianças e adolescentes a partir de um modelo de resposta binário, o modelo Logit, para os anos de 2018 e 2020. Nos modelos estimados, destaca-se que a grande maioria das variáveis utilizadas apresentou o coeficiente

significativo e com o sinal esperado. Sendo assim, uma menor probabilidade de distorção idade-série está associada às pessoas do sexo feminino, brancas ou amarelas, cuja mãe mora no mesmo domicílio, com maior nível de educação da mãe e cuja renda familiar per capita é maior.

Um resultado interessante diz respeito a variável de sexo, pois no ano de 2020 essa relação muda, de modo que ela passou a afetar positivamente a probabilidade de defasagem. Esse fator merece ser mais explorado, principalmente, porque é neste período que o ensino remoto estava em funcionamento, o que pode levantar discussões relevantes a partir de um maior aprofundamento. Outro resultado importante são variáveis do contexto educacional que, além de serem significativas e associadas a um menor número de distorções, foram capazes de tornar mais fraco os coeficientes de educação da mãe.

Este estudo precisa de um aprofundamento, tanto em termos teóricos, como metodológicos. A ideia principal é de entender inicialmente como é esse contexto, dada as bases disponíveis atualmente. Destaca-se a que é de interesse realizar futuramente uma análise utilizando todos os anos (de 2018 a 2022) que compõe o período do governo atual e analisar as questões de redução de financiamento do ensino básico. Também é interessante realizar um estudo mais aprofundado das variáveis relativas ao contexto escolar, pois as que foram utilizadas neste artigo foram obtidas a partir de agregações que podem estar escondendo importantes desigualdades na distribuição dos valores relativos às unidades escolares individuais.

5. Referências bibliográficas

BARROS, R. P. de; MENDONÇA, R. **Consequências da repetência sobre o desempenho educacional**. Brasília: Ministério da Educação. Projeto de Educação Básica para o Nordeste. 1998.

BARROS, Ricardo P.; MENDONÇA, Rosane; SANTOS, Daniel D.; QUINTAES, G. Determinantes do desempenho educacional no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.31, n.1. p. 1-42, abr. 2001.

COLEMAN, J. S. et al. **Equality of Educational Opportunity**. US Department of Health, Education and Welfare. Washington: U.S. Government Printing Office, 1966.

FERRÃO, M. E. et al. O SAEB – Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na escola eficaz. **Revista Brasileira de Estudos de População**, v. 18, p. 111-130, 2001.

FRANCO, A. M. De P. **Os determinantes na qualidade da educação no Brasil**. 2008. Tese (Doutorado em Economia) – Departamento da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (USP), São Paulo, 2008.

FRITSCH, R.; VITELLI, R.; ROCHA, C.S. Defasagem Idade-Série em Escolas Estaduais de Ensino Médio do Rio Grande do Sul. **Revista Brasileira de Estudos Pedagógicos**, v. 95, n. 239, 2014.

LEONE, Eugenia T.; MAIA, Alexandre G.; BALTAR, Paulo E. Mudanças na composição das famílias e impactos sobre a redução da pobreza no Brasil. **Economia e Sociedade**, Campinas, v.19, n.1 (38), p. 59-77, abr. 2010.

MACHADO, D. C. **Escolaridade das crianças no Brasil**: três ensaios sobre a defasagem idade série. Tese de Doutorado. Programa de Pós-Graduação em Economia, Departamento de Economia. Pontifícia Universidade Católica do Rio de Janeiro, PUC-RIO. Rio de Janeiro, RJ, 2005.

MACHADO, D.C.; GONZAGA, G. O impacto dos fatores familiares sobre a defasagem idade-série de crianças no Brasil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 61, n. 4, 2007.

MURNANE, R. J.; WILLETT, J. B. **Methods Matter**: Improving Causal Inference in Educational and Social Science Research. New York: Oxford University Press, 2011.

PINTO, J. M. R. Dinheiro traz felicidade? A relação entre insumos e qualidade na educação. In: PINTO, J. M. R.; SOUZA, S. A. (Orgs.). **Para onde vai o dinheiro? Caminhos e descaminhos do financiamento da educação**. São Paulo: Xamã, 2014.

PONTILI, R. M.; KASSOUF, A. L. Is Age-Grade Distortion in Brazil's primary education system more closely associated to school infrastructure or to family characteristics? **Well-Being and Social Polic.**, v. 4, n.1, p. 29-54, 2008

PORTELLA, Alysson Lorenzon; BUSSMANN, Tanise Brandão; OLIVEIRA, Ana Maria Hermeto de. A relação de fatores individuais, familiares e escolares com a distorção idade-série no ensino público brasileiro. **Nova economia**, v. 27, n. 3, p. 477-509, 2017.

POZZOBON, Magda; MAHENDRA, Fénita; MARIN, Angela Helena. Renomeando o fracasso escolar. **Psicologia Escolar e Educacional**, v. 21, p. 387-396, 2017.

RIBEIRO, S. C. A pedagogia da repetência. **Estudos Avançados**, São Paulo, n. 5, v. 12, p. 7-21, mai./jun. 1991

RIOS-NETO, E. L. G.; CÉSAR, C. C; RIANI, J. de L. R. Estratificação educacional e progressão escolar por série no Brasil. **Política e Planejamento Econômico**, Brasília, v. 32, n. 3, 2002.

SEBEN, Gabriela. **A família em transformação**: aspectos psicossociais da criança em duas distintas configurações familiares. 2010. 65f. Dissertação (Mestrado em Psicologia Clínica) – Faculdade de Psicologia, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2010.

SOARES, S; SÁTYRO, N. O impacto da infraestrutura escolar na taxa de distorção idade-série das escolas brasileiras de Ensino Fundamental – 1998- 2005. **Texto para Discussão**, No. 1338, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília, 2008.

ZAJAC, Danilo. **Ensino remoto na educação básica**. Nota técnica, Epufabc. São Paulo, 2022. Disponível em:< <https://epufabc.proec.ufabc.edu.br/ensino-remoto-na-educacao-basica/>>. Acesso em: 26 de maio de 2022.