

DETERMINANTES DA EVASÃO NO ENSINO SUPERIOR: UMA ABORDAGEM DE RISCOS COMPETITIVOS

Matheus Gomes
Guilherme Hirata

PROGRAMAS DE BONIFICAÇÃO A DOCENTES E RENDIMENTO ESCOLAR: UMA AVALIAÇÃO DO PROGRAMA ESCOLA DE VALOR NO ESTADO DA PARAÍBA

Gabriella da Silva Cavalcanti
Joelson Oliveira Santos
Ana Cláudia Annegues da Silva

SOCIOECONOMIC AND DEMOGRAPHIC DETERMINANTS OF MIGRATION IN BRAZILIAN MUNICIPALITIES: 2000-2010

Luís Abel da Silva Filho
Alexandre Gori Maia

"OS AMIGOS DOS MEUS AMIGOS SÃO MEUS AMIGOS?": AS CONSEQUÊNCIAS DO CAPITAL SOCIAL PARA AS DIFERENÇAS SALARIAIS ENTRE HOMENS E MULHERES NO BRASIL

Fábio Lúcio Rodrigues
Mércia Santos da Cruz
Wallace Patrick Santos de Farias de Souza

RETORNO SALARIAL E DOENÇAS CARDIOVASCULARES: EVIDÊNCIAS PARA O CASO BRASILEIRO

Victor Rodrigues de Oliveira
Giácomo Balbinotto Neto
Paulo de Andrade Jacinto
Eduardo da Silveira

DESIGUALDADE DE OPORTUNIDADES EDUCACIONAIS: INFLUÊNCIA DAS CIRCUNSTÂNCIAS DA INFÂNCIA NO DESEMPENHO ESCOLAR NO BRASIL

Sabrina Martins de Araújo
Cláudia Malbouisson
Wallace Patrick Santos de Farias Souza

pesquisa e planejamento econômico - ppe

ipea

Governo Federal

Ministério do Planejamento e Orçamento

Ministra Simone Nassar Tebet



Fundação pública vinculada ao Ministério do Planejamento e Orçamento, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiros – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidenta

Luciana Mendes Santos Servo

Diretor de Desenvolvimento Institucional

Fernando Gaiger Silveira

Diretora de Estudos e Políticas do Estado, das Instituições e da Democracia

Luseni Maria Cordeiro de Aquino

Diretor de Estudos e Políticas Macroeconômicas

Cláudio Roberto Amitrano

Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais

Aristides Monteiro Neto

Diretora de Estudos e Políticas Setoriais, de Inovação, Regulação e Infraestrutura

Fernanda De Negri

Diretor de Estudos e Políticas Sociais

Carlos Henrique Leite Corseuil

Diretor de Estudos Internacionais

Fábio Vêras Soares

Chefe de Gabinete

Alexandre dos Santos Cunha

Coordenador-Geral de Imprensa e Comunicação Social

Antonio Lassance

Ouvidoria: <https://www.ipea.gov.br/ouvidoria>

URL: <https://www.ipea.gov.br>

Pesquisa e Planejamento Econômico (PPE)

Publicação quadrimestral com análises teóricas e empíricas sobre uma ampla gama de temas relacionados à economia brasileira. Estabelecida em 1971 sob o título Pesquisa e Planejamento, PPE é publicada em abril, agosto e dezembro.

CORPO EDITORIAL

Editor

Maurício Cortez Reis

Coeditores

Marco A.F.H. Cavalcanti

José Gustavo Feres

Danilo Santa Cruz Coelho

Membros

Lauro Ramos (Ipea)

Eduardo Fiuza (Ipea)

Alexandre Xavier Ywata de Carvalho (Ipea)

Daniel da Mata (Ipea)

Carlos Viana de Carvalho (PUC-RJ)

Eduardo Rios Neto (CEDEPLAR-UFMG)

José Raimundo Carvalho (CAEN-UFC)

Marcelo Portugal (UFRGS)

Marco Bonomo (Insper)

Mônica Viegas de Andrade (CEDEPLAR-UFMG)

Rafael Coutinho Costa Lima (UFPE)

Renata Narita (FEA-USP)

Ricardo Paes de Barros (Insper)

Roberto G. Ellery (UnB)

Sergio Firpo (Insper)

Vladimir Ponczek (EESP-FGV/SP)

Secretária Executiva

Thamires Fernandes da Silva

© Instituto de Pesquisa Econômica Aplicada – **ipea** 2024

Pesquisa e Planejamento Econômico v. 1 – n.1 – jun. 1971.

Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada, 1990 – v. – quadrimestral.

Título anterior: Pesquisa e Planejamento v. 1, n. 1 e 2, 1971

Periodicidade anterior: semestral de 1971–1975.

1. Economia – Pesquisa – Periódicos. 2. Planejamento Econômico – Brasil. I. Brasil. Instituto de Pesquisa Econômica Aplicada.

ISSN - 0 100-0551

CDD 330.05
33(81) (05)

As publicações do Ipea estão disponíveis para download gratuito nos formatos PDF (todas) e ePUB (livros e periódicos). Acesse: <https://www.ipea.gov.br/portal/publicacoes>

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério do Planejamento e Orçamento.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

SUMÁRIO

DETERMINANTES DA EVASÃO NO ENSINO SUPERIOR: UMA ABORDAGEM DE RISCOS COMPETITIVOS.....	9
Matheus Gomes Guilherme Hirata	
PROGRAMAS DE BONIFICAÇÃO A DOCENTES E RENDIMENTO ESCOLAR: UMA AVALIAÇÃO DO PROGRAMA ESCOLA DE VALOR NO ESTADO DA PARAÍBA	39
Gabriella da Silva Cavalcanti Joelson Oliveira Santos Ana Cláudia Annegues da Silva	
DETERMINANTES SOCIOECONÔMICOS E DEMOGRÁFICOS DAS MIGRAÇÕES NOS MUNICÍPIOS BRASILEIROS: 2000-2010.....	77
Luís Abel da Silva Filho Alexandre Gori Maia	
"OS AMIGOS DOS MEUS AMIGOS SÃO MEUS AMIGOS?": AS CONSEQUÊNCIAS DO CAPITAL SOCIAL PARA AS DIFERENÇAS SALARIAIS ENTRE HOMENS E MULHERES NO BRASIL	101
Fábio Lúcio Rodrigues Mércia Santos da Cruz Wallace Patrick Santos de Farias de Souza	
RETORNO SALARIAL E DOENÇAS CARDIOVASCULARES: EVIDÊNCIAS PARA O CASO BRASILEIRO	135
Victor Rodrigues de Oliveira Giácomo Balbinotto Neto Paulo de Andrade Jacinto Eduardo da Silveira	
DESIGUALDADE DE OPORTUNIDADES EDUCACIONAIS: INFLUÊNCIA DAS CIRCUNSTÂNCIAS DA INFÂNCIA NO DESEMPENHO ESCOLAR NO BRASIL.....	161
Sabrina Martins de Araújo Cláudia Malbouisson Wallace Patrick Santos de Farias Souza	

CONTENTS

DETERMINANTS OF DROPOUT IN HIGHER EDUCATION: A COMPETITIVE RISK APPROACH.....	9
Matheus Gomes Guilherme Hirata	
TEACHER BONIFICATION AND SCHOOL PERFORMANCE: AN EVALUATION OF “ESCOLA DE VALOR” PROGRAM IN THE STATE OF PARÁIBA, BRAZIL.....	39
Gabriella da Silva Cavalcanti Joelson Oliveira Santos Ana Cláudia Annegues da Silva	
SOCIOECONOMIC AND DEMOGRAPHIC DETERMINANTS OF MIGRATION IN BRAZILIAN MUNICIPALITIES: 2000-2010.....	77
Luís Abel da Silva Filho Alexandre Gori Maia	
“THE FRIENDS OF MY FRIENDS ARE THEY MY FRIENDS?”: THE CONSEQUENCES OF SOCIAL CAPITAL FOR WAGE DIFFERENCES BETWEEN MEN AND WOMEN IN BRAZIL	101
Fábio Lúcio Rodrigues Mércia Santos da Cruz Wallace Patrick Santos de Farias de Souza	
SALARY RETURN AND CARDIOVASCULAR DISEASES: EVIDENCE FOR THE BRAZILIAN CASE	135
Victor Rodrigues de Oliveira Giácomo Balbinotto Neto Paulo de Andrade Jacinto Eduardo da Silveira	
INEQUALITY OF EDUCATIONAL OPPORTUNITIES: INFLUENCE OF CHILDREN’S CIRCUMSTANCES ON SCHOOL PERFORMANCE IN BRAZIL.....	161
Sabrina Martins de Araújo Cláudia Malbouisson Wallace Patrick Santos de Farias Souza	

DETERMINANTES DA EVASÃO NO ENSINO SUPERIOR: UMA ABORDAGEM DE RISCOS COMPETITIVOS¹

Matheus Gomes²

Guilherme Hirata³

As altas taxas de evasão no ensino superior podem representar um custo para indivíduos e a sociedade ao gerar perda do potencial acúmulo de capital humano. Este artigo analisa os fatores associados à evasão no ensino superior brasileiro, utilizando um modelo de análise de sobrevivência com riscos competitivos. Empregam-se dados longitudinais dos ingressantes em 2010, acompanhando-os até 2017. Os resultados mostram que mulheres, brancos, ingressantes mais velhos e quem utiliza o Programa Universidade para Todos (Prouni) ou o Fundo de Financiamento Estudantil (Fies) possuem menor risco de evasão. Além disso, as diferenças entre categorias (gênero, racial, cursos etc.) aparecem já no primeiro ano e aumentam com o tempo.

Palavras-chave: evasão; ensino superior; riscos competitivos; análise de sobrevivência.

DETERMINANTS OF DROPOUT IN HIGHER EDUCATION: A COMPETITIVE RISK APPROACH

High college dropout rates can represent an individual cost as well as a cost for the society due to the potential loss in human capital accumulation. This paper analyzes the factors associated with Brazilian higher education dropout using a competing risk survival model. We use longitudinal data to analyze the 2010 student newcomers, using data from 2010 to 2017. The results show that women, whites, older students and students using Fies loan or Prouni scholarship present a lower dropout risk. Differences between categories within the same group (gender, racial, courses etc.) appear in the first year and increase over time.

Keywords: dropout; higher education; competing risks; survival analysis.

JEL: I23; I22; I28.

1 INTRODUÇÃO

A evasão no ensino superior representa um custo tanto para o indivíduo quanto para a sociedade. Para o indivíduo, os custos estão principalmente associados à perda potencial de capital humano. Ao evadir, o indivíduo provavelmente não trabalhará na área de formação cursada, e o conhecimento específico acumulado tende a perder-se ao longo tempo. Some-se a isso a ausência de retorno (financeiro ou não) aos recursos investidos e ao tempo dedicado aos estudos, além dos efeitos psicológicos negativos que a não conclusão do curso podem acarretar. Finalmente, há ainda a perda de experiência no mercado de trabalho que o indivíduo poderia ter acumulado ao optar pelo ingresso no ensino superior em vez de ofertar mão de obra no mercado de trabalho.

1. DOI: <http://dx.doi.org/10.38116/ppe52n3art1>

2. Pesquisador da IDados. E-mail: matheus.souza@idados.id.

3. Pesquisador da IDados. E-mail: guilherme@idados.id.

A sociedade também sente a evasão pelo não acúmulo de capital humano, fundamental para a produtividade do trabalho, a inovação e o desenvolvimento econômico. Há também o efeito por custo de oportunidade: boa parte dos alunos é financiada pelo setor público de alguma forma, seja via universidade pública, seja via financiamento das mensalidades no setor privado. O investimento realizado pela sociedade dissipa-se a cada aluno que evade: além de não acumular capital humano, ela deixa de recolher os frutos desse investimento caso tivesse sido aplicado de outra forma.

Entretanto, a evasão tem seu lado positivo, quando decorre de processo de autoconhecimento para o aluno (Stinebrickner e Stinebrickner, 2014). Em vez de insistir em algo para o qual não está apto e/ou estimulado, a evasão pode representar um novo começo, seja em outro curso, seja no mercado de trabalho. Nesse sentido, o custo para o indivíduo e a sociedade poderia ser ainda maior, caso o aluno concluísse o curso, e este não fosse utilizado para fim algum.

Este artigo analisa os fatores associados à evasão em cursos de ensino superior no Brasil. Utilizando-se de dados do Censo da Educação Superior, do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), órgão vinculado ao Ministério da Educação (MEC), a coorte de estudantes que ingressou em 2010 é acompanhada até 2017. Ao todo, são 2,1 milhões de alunos. Para realizar o estudo, empregou-se a abordagem de riscos competitivos em um modelo de análise de sobrevivência. Nesse sentido, essa abordagem leva em consideração o fato de que a evasão concorre com a possibilidade de outros eventos competitivos, como a transferência e a conclusão do curso. Na ocorrência de algum outro evento, a evasão não tem mais possibilidade de ocorrer. Portanto, o evento de interesse depende dos eventos competitivos. O método proposto trabalha com a função de interrupção (*failure function*) ou função de incidência, em vez de focar a função de sobrevivência. O interesse recai na probabilidade de a evasão ocorrer em dois anos, por exemplo, e não na probabilidade de estar matriculado após dois anos. O motivo é que não se sabe qual evento ocorrerá após dois anos de curso e que ainda há o risco de um evento competitivo ocorrer.

Os resultados sugerem diversos padrões. Em termos de características pessoais, mulheres e brancos possuem menor risco de evasão, condicional às demais características. Após quatro anos de curso, enquanto a probabilidade de evasão para mulheres alcança cerca de 30%, para os homens, chega a 34%. Ingressantes mais novos têm maior probabilidade de evasão que os mais velhos. Além disso, alunos que utilizam o Programa Universidade para Todos (Prouni)⁴ ou o Fundo de Financiamento Estudantil (Fies)⁵ apresentam menor probabilidade de evasão.

4. Programa do governo federal destinado a alunos que cursaram integralmente o ensino médio em escolas públicas; concede bolsas para o pagamento de mensalidades. Para mais detalhes, ver o *link* disponível em: <https://acessounico.mec.gov.br/>.

5. Programa do governo federal que concede empréstimos a juros baixos para pagamento das mensalidades, os quais são cobrados somente após a conclusão do curso. Para mais detalhes, ver o *link* disponível em: <http://sisfiesportal.mec.gov.br/index.php>.

Alunos com Fies possuem 23% de probabilidade de evasão após quatro anos, enquanto alunos com Prouni apresentam 25%; para os demais, a probabilidade é de 35% ou mais. Em relação aos cursos, observam-se poucas diferenças significativas de evasão entre as oito grandes áreas definidas pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), detalhadas adiante.

Grande parte da literatura sobre evasão no ensino superior refere-se aos Estados Unidos. Dos 27 estudos analisados por Haas e Hadjar (2020), dezenove utilizam dados daquele país. Os autores apontam que grupos historicamente desfavorecidos, definidos por questão de origem social e étnica, seguem trajetórias menos lineares. Para a Itália, Meggiolaro, Giraldo e Clerici (2017), utilizando dados da Universidade de Padova, reportam que nativos, residentes da cidade onde a universidade está localizada, e aqueles que ingressam no ensino superior logo após encerrar o ensino médio possuem menor probabilidade de evasão. Ambos os estudos mostram que mulheres têm menor probabilidade de evasão.

A restrição de crédito também parece ser um fator importante. Cardak e Vecchi (2015) argumentam que isso afeta especialmente os alunos com pior desempenho no ensino médio. Apesar de Johnson (2013) encontrar que as restrições a empréstimos têm pouco efeito sobre os alunos se graduarem, o autor argumenta que os subsídios são necessários para obter melhores resultados, com maior benefício se direcionado a estudantes de habilidades médias. Porém, até mesmo na presença de restrições de crédito, outros fatores, ligados principalmente a alunos de baixa renda, ainda se mostram importantes (Stinebrickner e Stinebrickner, 2008).

No Brasil, de modo geral, os estudos tratam de casos específicos. Andriola, Andriola e Moura (2006) analisam a evasão por intermédio das opiniões de docentes, coordenadores e discentes que se evadiram na Universidade Federal do Ceará (UFC). Segundo os alunos que evadiram entre 1999 e 2000, entre os principais motivos para a evasão, encontram-se o baixo conhecimento sobre o curso no momento de ingresso, a desmotivação com o estudo e a precariedade das condições físicas do curso. Lima Junior, Silveira e Ostermann (2012) analisam a evasão dos estudantes de física da Universidade Federal do Rio Grande do Sul (UFRGS), entre 1995 e 2009, mediante uma análise de sobrevivência. Segundo os autores, no curso analisado, não há diferenças significativas na ocorrência de evasão e diplomação entre sexo ou grau acadêmico; porém, as mulheres demoram mais para evadir. Ademais, quanto maior for a nota no vestibular, mais cedo os alunos tendem a se graduar. Um estudo com dados da rede privada da Fundação Escola de Comércio Álvares Penteado (Fecap), entre 2006 e 2009, mostra que reprovação e aumento nas mensalidades são fatores associados com evasão. Os resultados mostram ainda que participar do Prouni reduz as chances de desistência durante a graduação (Silva, 2013).

Mendes Junior (2014) estuda o caso de ações afirmativas da Universidade do Estado do Rio de Janeiro (UERJ). O autor analisa a progressão de estudantes cotistas e não cotistas que ingressaram em 2005, acompanhando-os até 2011. Os alunos cotistas apresentaram maiores taxas de graduação e menores taxas de evasão. O autor argumenta que esses alunos tendem a atribuir um alto valor ao curso em que ingressaram, pois apresentaram melhores taxas de progressão, apesar de menores coeficientes de rendimento médios.

Saccaro, França e Jacinto (2019) realizam um estudo mais abrangente, ao utilizar dados do Censo da Educação Superior, com o objetivo de investigar a trajetória nos cursos das áreas de ciência, matemática, computação, engenharia, produção e construção no período 2010-2014. Utilizando-se de métodos de análise de sobrevivência, os resultados encontrados mostram que a evasão foi maior nas instituições privadas e que alunos contemplados com apoio financeiro apresentaram maior taxa de conclusão.

Este artigo contribui para a literatura de quatro formas. Em primeiro lugar, trata do ensino superior, área com poucos trabalhos no Brasil. Em segundo lugar, utiliza dados para todo o Brasil, o que permite obter um panorama geral a respeito da evasão para todo o país. Em terceiro lugar, acompanha a trajetória de cada aluno por um longo período, reduzindo a quantidade de casos censurados, o que diminui a possibilidade de viés devido à ausência de informação. Finalmente, o artigo é um dos poucos a utilizar a abordagem de riscos competitivos para avaliar a evasão no Brasil.

Além desta introdução, este trabalho está dividido em mais quatro seções. A seção 2, a seguir, apresenta a metodologia empregada. Na seção 3, têm-se a descrição da fonte de dados e as estatísticas descritivas. Na seção 4, são apresentados os resultados e, na seção 5, seguem-se as conclusões.

2 METODOLOGIA

Há duas formas mais comuns de analisar-se a evasão. Uma consiste em utilizar uma estrutura de resultados binária ou multinomial com dados *cross-section*, em que se observa o *status* do estudante em determinado ponto no tempo. A outra conduz uma análise de sobrevivência que utiliza dados longitudinais, considerando tanto a incidência com que ocorre a evasão quanto o momento em que esta ocorre.

Neste artigo, optou-se pela análise de sobrevivência justamente porque tal método atribui um papel relevante ao período em que ocorre o evento, ao diferenciar os indivíduos nessa dimensão. Além disso, esse método permite considerar mais de dois resultados possíveis e excludentes, como os métodos multinomiais (Feng e Sass, 2012). Considerando-se que os estudantes do ensino superior encerram sua vida

acadêmica em determinado curso, concluindo-o, evadindo-se ou transferindo-se, a probabilidade de evasão depende da ocorrência desses eventos competitivos.

O evento de interesse para este estudo é a evasão. Seguindo Fine e Gray (1999) e Austin e Fine (2017), na presença de eventos competitivos, define-se a função de subdistribuição de risco (*subdistribution hazard*), daqui em diante denominada sub-risco, $h_j(t)$, que nada mais é que a taxa instantânea de ocorrência do evento j no período t para indivíduos que ainda não experimentaram o evento j , como:

$$h_j(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}(t \leq T < t + \Delta t, D=j | T \geq t \cup (T \leq t \cap D \neq j))}{\Delta t}, \quad (1)$$

em que $D = \{1, \dots, D\}$ é o conjunto de eventos possíveis, e T , o tempo em que ocorreu o evento. Note-se que a função de sub-risco se refere àqueles que ainda não experimentaram especificamente o evento de interesse, e não àqueles que não experimentaram nenhum evento – essa segunda função é dada pela função de risco de causa específica (*cause-specific hazard function*).⁶ Ou seja, a função de sub-risco refere-se à população em risco de experimentar o evento de interesse ou que já experimentou um evento competitivo. Nesse sentido, o grupo de referência para o qual é calculado o risco em dado ponto no tempo é maior na função de sub-risco.

Para estimar quais fatores estão associados ao risco de evasão, utiliza-se o modelo semiparamétrico (equação 2), em que o sub-risco $h_{ij}(t)$ do evento j no tempo t para o estudante i que ingressou no curso k da instituição m é dado por:

$$h_{ij}(t) = \exp(\beta_j X_{ikm} + \gamma_j C_k + \delta_j I_m) h_{0j}(t), \quad (2)$$

em que $h_{0j}(t)$ é o sub-risco base do evento j (o risco quando todas as covariadas são iguais a zero); X , um vetor de características dos estudantes; C , um vetor de características dos cursos; e I , um vetor de características das instituições. β_j, γ_j e δ_j são os vetores de coeficientes da regressão para o evento j , $j = \{1, 2, 3\}$, e o estudante que continua matriculado ao fim do período da amostra é considerado uma observação censurada à direita.⁷

A estimação para cada evento é realizada separadamente. Primeiramente, estima-se o risco de evasão com base na transferência e na conclusão como eventos competitivos, que é o principal interesse do artigo. Depois, calcula-se o risco de conclusão, admitindo-se evasão e transferência como eventos competitivos. Finalmente, isso também para transferência. Note-se que os coeficientes estimados, apesar de indicarem a direção do efeito, não se referem ao efeito quantitativo de cada variável sobre o risco de evasão. Por esse motivo, para obter uma interpretação

6. A função de risco de causa específica é dada por $h_j^{sc}(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob}(t \leq T < t + \Delta t, D=j | T \geq t)}{\Delta t}$.

7. Quando o tempo de ocorrência dos eventos de interesse (evasão/conclusão/transferência) é superior ao tempo registrado; no caso, oito anos.

mais intuitiva, os resultados são reportados em termos de razão de sub-risco (*subhazard ratio*). A razão, quando maior (menor) que 1, indica que o aumento de uma unidade na covariável está associado a um(a) aumento (redução) no risco de evasão para aqueles que ainda não experimentaram a evasão. Quando a variável é categórica – incluindo-se *dummies* –, a razão indica a diferença no risco entre determinada categoria e o grupo de referência. Percebe-se também na equação (1) que não se pode comparar a razão de sub-risco estimada para a evasão; por exemplo, a razão estimada para conclusão. As populações de referência em cada caso são diferentes.

Vale ressaltar que há uma peculiaridade a respeito da razão de sub-risco. No contexto deste artigo, por ser uma estimativa válida para quem ainda não evadiu, a estimativa da taxa instantânea de ocorrência do evento é válida também para quem já concluiu o curso ou se transferiu, ainda que a ocorrência desses eventos competitivos impeça a ocorrência da evasão. Como observado anteriormente, isso faz com que a estimativa seja menor em relação à taxa instantânea para quem não experimentou evento algum, ou seja, desconsiderando-se quem já concluiu, evadiu ou se transferiu – taxa estimada pelo modelo de causa específica.

Por sua vez, o uso da função sub-risco permite inferir a direção do efeito da covariável sobre a incidência do evento, dada pela função de incidência acumulada (CIF):

$$CIF_j(t) = 1 - \exp\left\{-\int_0^t h_j(t)dt\right\}. \quad (3)$$

Para alguns casos, será apresentada a CIF do evento para o qual a razão de sub-risco foi estimada.⁸ Essa função indica a taxa acumulada de incidência do evento ao longo do tempo. Isso é útil para comparar a situação de dois ou mais grupos (homens e mulheres, brancos e negros etc.). É possível, por exemplo, que alunos de cursos diurnos tenham inicialmente maior incidência de evasão no início do curso, ao passo que os alunos de cursos noturnos tendam a evadir mais no final. A observação de que alunos de um curso possuem maior ou menor probabilidade de evasão é complementada com a CIF. Note-se, no entanto, que, embora a razão de sub-risco reportada indique um aumento (se for maior que 1) ou diminuição (caso contrário) na CIF, o valor da razão não pode ser interpretado como o aumento na probabilidade de incidência.

3 DADOS, DEFINIÇÕES E AMOSTRA

Os dados utilizados na análise são os microdados do Censo da Educação Superior disponibilizados pelo Inep/MEC. O censo é anual e traz informações para

8. Pelo método Fine-Gray, o efeito sobre a CIF de um evento pode ser obtida por meio do efeito das covariadas sob esse evento (Geskus, 2015).

indivíduos, cursos e instituições de todas as matrículas no ensino superior do Brasil. São utilizados os dados de 2010 a 2017.

A principal variável de interesse é aquela que indica a situação do aluno no ano: cursando; matrícula trancada; desvinculado do curso; transferido para outro curso da mesma instituição de ensino superior (IES); formado; e falecido. O aluno evadido é aquele desvinculado do curso. Se a matrícula estiver trancada, considera-se que o aluno ainda está na situação *cursando*. Optou-se por essa definição porque nessa situação o aluno ainda pode retornar ao curso, e esse fato exclui a possibilidade de considerar o trancamento como um evento, que é uma situação definitiva.⁹ Finalmente, alunos falecidos são excluídos da amostra. Assim, há os três eventos de interesse: evasão; transferência para outro curso da mesma instituição – daqui em diante tratado somente como transferência;¹⁰ e conclusão.

O censo fornece ainda informações sobre sexo, raça e idade dos alunos; o curso e a instituição a que se está vinculado; o semestre de ingresso;¹¹ e se o aluno recebeu algum financiamento (Fies e outros) ou bolsa do Prouni. Para o curso, há informações sobre o turno em que as aulas ocorrem, o grau acadêmico conferido e a área geral do curso de acordo com as definições da OCDE. Finalmente, para as instituições, é possível saber a categoria administrativa e o tipo de organização acadêmica.

A base de dados fornece uma variável que identifica um estudante unicamente ao longo dos anos. Isso permite a construção da trajetória de cada aluno que ingressa no sistema. Vale ressaltar, no entanto, que os dados não são identificados. Isso significa que, se um mesmo indivíduo está matriculado em dois cursos, tal indivíduo, na prática, será tratado como dois alunos distintos, cada um com seu identificador único atrelado a apenas um curso. Para definir o universo de alunos ingressantes em 2010, utilizaram-se os alunos que apresentavam o ingresso em 2010 nos censos de 2010 a 2013. Isso é feito para que alunos vinculados a cursos de área básica de ingresso (ABI)¹² sejam incluídos na análise, tendo em vista que, na escolha da formação acadêmica, os vínculos são registrados com a data de ingresso original da ABI. Ou seja, um aluno que ingressou em ABI em 2010, cursou dois anos e passou para o curso de engenharia química em 2012 possui 2010 como ano de ingresso nesse

9. Ao considerar-se o trancamento como cursando, não se impede que o evento final (evasão, conclusão ou transferência) ocorra; permite-se que se admita a ocorrência do evento final quando este de fato ocorreu – ou seja, o aluno não tinha mais chances de retornar ao curso. Exemplificando: se um aluno trancou a matrícula em 2012; em 2014, sua situação apareceu como evadido. Para fins do estudo, ele evadiu nesse último ano.

10. Alunos que se transferiram ou mudaram de curso para outra IES são considerados evadidos, dado que recebem uma nova matrícula no censo e não podem mais ser acompanhados.

11. Até 2012, consideraram-se ingressantes no primeiro semestre os alunos que apresentavam de janeiro a junho na variável *dt_ingresso_curso*. De 2013 em diante, os alunos ingressantes no primeiro semestre exibiam o valor "01" para ingressantes no primeiro semestre e "07" para ingressantes no segundo semestre. Utilizou-se essa informação em censos posteriores a 2010, quando naquele ano a informação de ingresso era faltante.

12. ABI é quando o ingresso no curso possibilita que, após a conclusão de um conjunto básico de disciplinas – denominado ciclo básico –, o aluno escolha uma entre duas ou mais formações acadêmicas.

curso, mas concomitantemente estava registrado no sistema como ABI em 2010 e 2011, podendo ser confundido com aluno transferido. Com esse procedimento, foram selecionados 2.715.624 milhões de alunos ingressantes em 2010.

Para a construção da trajetória escolar de cada indivíduo em cada curso, é necessário resolver algumas inconsistências. Por exemplo, há alunos que aparecem com situação de conclusão de curso em determinado ano, mas voltam a aparecer no mesmo curso em anos posteriores com outra situação de matrícula. A correção da trajetória foi baseada no tratamento feito pelo Inep (2017) e seguiu as regras descritas a seguir.

- 1) Se, durante o período de análise, a data de ingresso do aluno foi alterada para um ano anterior a 2010, retirou-se esse aluno da amostra.
- 2) Se a data de ingresso foi alterada para um ano posterior a 2010, imputou-se a essa trajetória inicial a situação de evasão no ano em que houve alteração.
- 3) Se houve duas situações de finalização do curso (evadido, transferido para outro curso na mesma IES ou formado) em dois censos, retirou-se a última informação. Assim, o primeiro ano em que o vínculo foi finalizado é considerado como o evento analisado.
- 4) Se, em determinado ano do censo, a situação é de finalização do curso e, posteriormente, volta a ser de vinculação (cursando ou matrícula trancada), substitui-se a finalização por vinculação.
- 5) Se, em determinado ano, o aluno apresentar qualquer situação e, em todos os anos anteriores, não tiver nenhuma informação, substituem-se os anos anteriores pelo vínculo *cursando*.
- 6) Se o aluno estiver vinculado à IES em determinado ano e não houver nenhuma informação de vínculo posterior, atribuiu-se a situação de *evasão* no período seguinte.
- 7) Na presença de outras inconsistências, as observações foram retiradas da base de dados.

Após esses procedimentos, a amostra contém 2.419.415 alunos, na qual um aluno tem somente uma observação por curso, instituição e ano. Para cada aluno, há a informação de sua situação acadêmica de 2010 até o ano de ocorrência de algum evento ou até 2017 (último ano da amostra), o que ocorrer primeiro, e isso caracteriza um painel desbalanceado.

Vale notar que 24,8% dos alunos da amostra tiveram a trajetória corrigida. Essa alta porcentagem traz preocupações a respeito da influência dessas correções sobre o resultado final. É possível que haja correlação entre a ausência de informação da situação acadêmica e a qualidade das demais informações disponibilizadas no

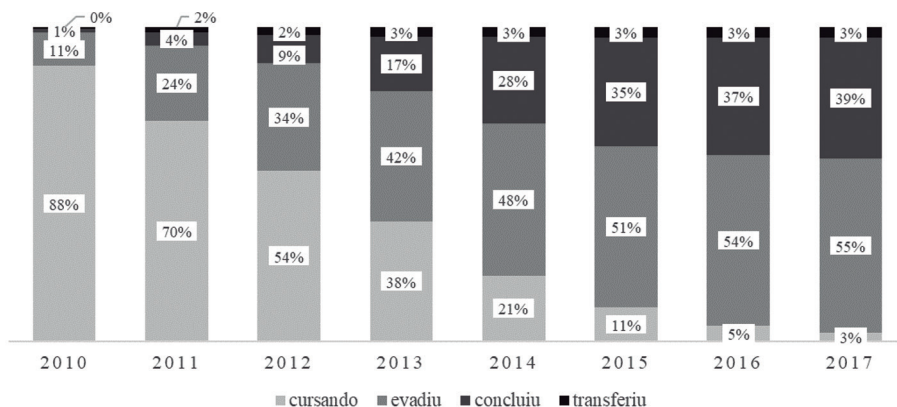
censo. Isso poderia levar a estimativas viesadas dos coeficientes de interesse. Como exercício de robustez, o mesmo modelo foi estimado ao excluir os alunos que tiveram a trajetória corrigida, e a tendência dos resultados manteve-se.

Ademais, cerca de 400 mil alunos (16,6% do 2,4 milhões) estavam matriculados em cursos de educação a distância (EaD). Para que a análise se concentre somente em alunos dos cursos presenciais, que possuem dinâmica diferente em relação aos cursos de EaD por natureza, a especificação principal exclui esses alunos da amostra. Novamente, estima-se uma especificação que inclui esses alunos com o objetivo de testar a robustez dos resultados, e os resultados não se alteram.

O gráfico 1 apresenta a distribuição anual dos ingressantes em 2010 por tipo de evento ocorrido até 2017. Observa-se que 11% dos ingressantes evadem logo no primeiro ano de curso. Ao fim do quarto ano, a taxa de evasão sobe para 42%. Cinco anos após o ingresso, somente 28% haviam concluído o curso em que ingressaram em 2010.

GRÁFICO 1

Distribuição dos alunos por tipo de evento ocorrido a cada ano: coorte 2010 (2010-2017)
(Em %)



Fonte: Censo da Educação Superior do Inep referente ao período 2010-2017. Disponível em: <http://portal.inep.gov.br/basicalevantamentos-acessar>. Acesso em: 13 jun. 2016.

Obs.: Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Elaboração dos autores.

Complementando a análise, a tabela 1 apresenta a distribuição das variáveis utilizadas no estudo tanto para o total (colunas *total* e %) quanto por tipo de evento ocorrido até 2017 (colunas *cursando*, *evadiu*, *concluiu* e *transferiu*). Observa-se que o ingressante típico¹³ de 2010 é mulher, possui 25 anos ou menos, não recebeu

13. Devido ao grande número de estudantes com dados faltantes sobre sua cor, não se incluiu essa característica no estudante típico.

nenhum tipo de financiamento ou bolsa e está matriculado na área de ciências sociais, negócios e direito e em uma universidade privada.

TABELA 1
Resultados das trajetórias dos alunos por indivíduo, curso e instituição (2010)

		Cursando (%)	Evadiu (%)	Concluiu (%)	Transferiu (%)	Total (1 mil)	%
Características do aluno							
Sexo	Masculino	3	59	34	4	921,4	46
	Feminino	2	52	43	3	1.096,3	54
Cor	Branco	3	50	44	3	521,0	26
	Não branco	3	52	42	2	300,7	15
	Não declarado ou dado faltante	2	59	35	4	1.195,9	59
Idade	25 anos ou menos	1	57	38	4	1.030,4	51
	26 anos ou mais	5	54	39	2	987,2	49
Semestre de ingresso	Primeiro semestre	2	54	41	3	1.324,1	66
	Segundo semestre	3	59	35	3	693,6	34
Financiamento	Pública	5	49	43	3	453,1	22
	Privada sem financiamento	2	64	30	4	981,6	49
	Privada (outro financiamento)	2	52	43	3	345,3	17
	Privada com Fies	4	37	56	2	116,3	6
	Privada com Prouni	2	38	57	3	121,3	6
Características do curso							
Turno	(Matutino, vespertino)	3	55	39	3	459,9	23
	Noturno	2	58	37	3	1.340,0	66
	Integral	4	43	49	4	217,8	11
Grau acadêmico	Bacharelado	3	55	38	3	1.386,6	69
	Licenciatura	3	54	40	3	331,2	16
	Tecnólogo	1	57	40	2	299,9	15
	Agricultura e veterinária	5	50	44	2	44,8	2
Cursos (OCDE)	Ciências, matemática e computação	3	62	31	4	157,8	8
	Ciências sociais, negócios e direito	2	57	38	3	845,8	42
	Educação	3	53	40	3	330,7	16
	Engenharia, produção e construção	5	56	34	6	260,1	13
	Humanidades e artes	3	57	36	4	59,6	3
	Saúde e bem-estar social	2	49	46	3	268,6	13
	Serviço	1	56	40	2	50,2	2

(Continua)

(Continuação)

		Cursando (%)	Evadiu (%)	Concluiu (%)	Transferiu (%)	Total (1 mil)	%
Características da instituição							
Organização acadêmica	Universidade	3	53	40	4	990,4	49
	Faculdade	2	59	38	2	706,5	35
	Centros universitários/ outros ¹	3	57	37	3	320,7	16
Total		55.016	1.118.506	778.624	65.550	2.017,7	-
%		3	55	39	3	100,0	-

Fonte: Censo da Educação Superior 2010-2017 do Inep. Disponível em: <http://portal.inep.gov.br/basicas-levantamentos-acessar>. Acesso em: 13 jun. 2016.

Elaboração dos autores.

Nota: ¹ Institutos federais de educação, ciência e tecnologia e centros federais de educação tecnológica (Cefets).

Destaca-se o fato de mais da metade das matrículas serem em cursos noturnos, o que sugere que boa parte dos alunos trabalha enquanto cursa a faculdade, informação inexistente no censo. Nota-se também que um em cada seis alunos se matriculou em curso na área de educação (pedagogia, principalmente), cursos que formam os futuros professores dos anos iniciais, área que deve demandar cada vez menos profissionais devido à transição demográfica.

A tabela 1 mostra que, para homens, negros¹⁴ e os que ingressaram com 25 anos ou menos, as taxas de evasão são maiores. Isso também ocorre para quem estava na rede privada que não tinha financiamento ou tinha algum financiamento que não o Fies e o Prouni. Alunos do período noturno também evadiram mais. Para as áreas gerais dos cursos, as taxas de evasão variam entre 49% e 57%. Os cursos da área de ciências, matemática e computação foram os que apresentaram as maiores taxas de evasão.

A tabela A.1, no apêndice A, complementa a análise ao apresentar a distribuição dos eventos por característica dos alunos. Entre evadidos, há ligeiramente mais mulheres; 57% são alunos com 25 ou menos; dois terços ingressaram no primeiro semestre; e 70% frequentavam o período noturno.

4 RESULTADOS

4.1 Risco de evasão

A tabela 2 apresenta os resultados do modelo de riscos competitivos aplicado aos estudantes ingressantes em 2010 para os três eventos possíveis: evasão, conclusão e transferência. Os valores reportados em cada coluna são as razões de sub-risco. Conforme discutido na seção 2, valores maiores que 1 indicam que a probabilidade de o respectivo evento ocorrer é maior para o grupo indicado pela *dummy* ou pela categoria da covariada em questão em relação ao grupo de referência. Todas as

14. Negros inclui pretos, pardos e indígenas. Brancos inclui brancos e amarelos.

estimações incluem todos os controles reportados na tabela 1, além de efeito fixo para a região de localização da IES (Norte, Nordeste, Sudeste, Sul e Centro-Oeste). O erro-padrão foi estimado utilizando-se o *cluster* no nível da instituição. A discussão a seguir foca a evasão; os resultados para conclusão, em geral, espelham os resultados para evasão. Os resultados para a conclusão e a transferência são mencionados oportunamente.

Condicional a todas as demais covariáveis, as mulheres possuem menor risco de evasão em comparação com os homens. Ou seja, condicional à área do curso, dependência administrativa da IES e raça, por exemplo, o risco de evasão para mulheres é 15% menor que o observado para homens, considerando-se a população em risco de experimentar evasão. Esse resultado é consistente com os encontrados por Meggiolaro, Giraldo e Clerici (2017), na Itália; por Haas e Hadjar (2020), para alguns países, principalmente Estados Unidos; e Saccaro, França e Jacinto (2019), no Brasil, para alguns cursos.

Além do gênero, a raça também está associada à evasão. Negros possuem maior risco de evasão que brancos. Isso pode estar associado em nível socioeconômico e restrição de crédito, variáveis não controladas nas estimações. No entanto, a diferença não alcança grande magnitude, 8%, sendo menor, por exemplo, que a diferença entre gêneros. Por sua vez, o fato de haver muitos alunos sem informação de raça limita a interpretação dos resultados. Nota-se que, para aqueles que não possuem informação sobre raça, há maior risco de evasão que a estimada para brancos. Se a maior parte desse grupo for de negros, a diferença entre brancos e negros está subestimada.

Em relação à idade, ingressar com 25 anos ou menos está associado a maior risco de evasão, dado que a razão de sub-risco é 1,21 para quem ingressa mais jovem. O risco de transferência para o jovem também é maior, de 79%. Isso indica que os mais jovens tendem a seguir uma carreira mais inconstante, o que pode estar associado à hipótese de autoconhecimento (Stinebrickner e Stinebrickner, 2014). Os mais jovens teriam expectativas menos alinhadas com a realidade dos cursos, e isso resultaria em mais evasão e transferência. É possível também que alunos mais velhos já tenham evadido anteriormente, o que aumenta as chances de acerto na escolha do curso seguinte.

TABELA 2
Modelo de riscos competitivos com três resultados possíveis e *cluster* no nível da IES (2010)

Variáveis		Razões de sub-risco		
		Evadiu	Concluiu	Transferiu
Características dos alunos				
Sexo	(referência: masculino)			
	Feminino	0,85***	1,27***	0,86***
Cor	(referência: brancos)			
	Não brancos	1,08***	0,92***	0,99
	Não declarado ou dado faltante	1,29***	0,78***	1,49***
Idade	(referência: 26 anos ou mais)			
	25 anos ou menos	1,21***	0,96**	1,79***
Semestre ingresso	(referência: segundo semestre)			
	Primeiro semestre	0,95***	1,30***	0,85***
Financiamento	(referência: públicas)			
	Privada sem financiamento	1,38***	0,76***	1,86***
	Privada com outros financiamentos	0,97	1,16***	1,60***
	Privada com Fies	0,59***	1,49***	1,36*
	Privada com Prouni	0,64***	1,64***	1,76***
Características do curso				
Turno	(referência: matutino e vespertino)			
	Noturno	1,07***	0,98	0,95
	Integral	0,79***	1,29***	0,98
Grau acadêmico	(referência: bacharelado)			
	Licenciatura	1,46**	0,46***	3,52***
	Tecnólogo	0,93***	1,57***	0,68***
Cursos (OCDE – geral)	Agricultura e veterinária	1,06*	0,90***	0,65***
	Ciências, matemática e computação	1,17***	0,74***	1,55***
	Educação	0,70**	2,37***	0,39**
	Engenharia, produção e construção	1,03	0,77***	2,01***
	Humanidades e artes	1,04	0,89**	1,56**
	Saúde e bem-estar social	0,96**	1,10***	0,93
	Serviço	1,06	0,94	1,07

(Continua)

(Continuação)

Variáveis	Razões de sub-risco			
	Evadiu	Concluiu	Transferiu	
Características da instituição				
(referência: universidade)				
Organização acadêmica	Faculdade	1,12***	0,97	0,38***
	Centro universitário ¹	1,08	0,95	0,66***

Elaboração dos autores.

Nota: ¹ Institutos federais de educação, ciência e tecnologia e Cefets.Obs.: Nível de significância: * $p < 0,10$; ** $p < 0,05$; e *** $p < 0,01$.

Infelizmente, não há informação sobre oferta de trabalho ou graduações cursadas e/ou concluídas anteriormente. O resultado, de acordo com idade, é consistente com a ideia de que os mais velhos ingressam no ensino superior quando já trabalham ou já possuem outro diploma: suas expectativas estariam mais alinhadas com seu perfil e com a realidade do curso. O fato de haver risco maior de evasão para cursos noturnos sugere que o custo de trabalhar e estudar está presente. No entanto, não há risco menor de conclusão, o que indica que o possível *trade-off* trabalhar-estudar é dominado pelo autoconhecimento. Se houvesse um efeito maior da restrição de crédito, esperar-se-ia um risco menor de conclusão.

Em relação a uma segunda graduação, como do Censo da Educação Superior não consta essa informação, um *insight* pode ser observado por meio do Censo Demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE)¹⁵, referente a 2010, observando-se quem estava cursando o ensino superior em 2010, e não somente quem ingressou em 2010.¹⁶ Por esse censo, 19% dos alunos que estavam cursando o ensino superior com mais de 25 anos estavam cursando a segunda graduação, contra 4% dos alunos mais novos. Portanto, apesar de o percentual ser bem maior que o dos novos, a parcela dos mais velhos que estão na segunda graduação ainda é modesta, apesar de não desprezível.

O risco de evasão é ligeiramente menor para quem ingressa no primeiro semestre, e há um risco maior de conclusão: quem ingressa no primeiro semestre tem 30% a mais de risco de concluir relativamente a quem ingressa no segundo semestre. O acesso ao ensino superior no início do ano é mais concorrido, uma vez que ocorre com o final do ano letivo, o que aumenta o número de candidatas. Isso faz com que as turmas do primeiro semestre tendam a ser compostas por alunos com habilidade mais alta, cujo desempenho no exame admissional é

15. Disponível em: https://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Resultados_Gerais_da_Amostra/Microdados/. Acesso em: 11 mar. 2016.

16. No Censo Demográfico, não é possível observar quando o indivíduo ingressou no ensino superior.

melhor que o observado para ingressantes do segundo semestre.¹⁷ Se a habilidade está relacionada com desempenho no curso, alunos com habilidade mais baixa teriam mais chance de retardar a conclusão ou acelerar a evasão, o que seria consistente também com a hipótese de que tais alunos perderiam menos em termos de perspectivas salariais que alunos de alta habilidade em caso de não conclusão (Hendricks e Leukhina, 2011).¹⁸ Uma explicação alternativa estaria ligada ao resultado reportado por Sampaio *et al.* (2011), os quais apontaram que, para ingressantes de 2003 a 2004 na Universidade Federal de Pernambuco (UFPE), a renda mais baixa está associada a uma menor probabilidade de evasão. Com menos opções, não valeria a pena aguardar o segundo semestre. Esse é um tópico que demanda mais estudos, e a falta de dados limita o que pode ser analisado no contexto deste artigo.

Um resultado para o qual ainda há pouca evidência empírica é o relacionado ao tipo de financiamento. Para analisar essa questão, utilizou-se uma variável categórica que divide os estudantes em cinco grupos, conforme apresentado na tabela, na qual a categoria de referência são os estudantes das IES públicas. Os alunos da rede privada que não recebem financiamento possuem risco de evasão quase 40% maior que o observado para alunos de instituições públicas. Por seu turno, usar financiamento equilibra o risco e pode até inverter o cenário. Para aqueles que recebem algum financiamento que não Fies ou Prouni, não há diferenças significativas para a rede pública. Ademais, estudantes que recebem Fies ou participam do Prouni obtêm razões de sub-risco significativas e menores que 1; isso indica que esses dois programas estão associados a uma redução do risco de evasão no ensino superior. Comparando-se as razões, o Fies é ligeiramente superior em evitar a evasão, o que pode estar ligado à percepção individual ou familiar do custo da evasão. Enquanto universidades públicas e o Prouni – para uma parcela da rede privada (aqueles com bolsa integral) – não demandam gastos monetários com mensalidades, o Fies é um empréstimo. Essa diferença pode influenciar a decisão de evasão por parte do aluno.

Vale notar também que alunos da rede privada com financiamento possuem maior risco de conclusão que alunos da rede pública; em especial, aqueles com Fies ou Prouni. Esse é um tópico que precisa ser mais bem investigado, uma vez que há grandes consequências para a sociedade. Muito se discute sobre os custos

17. Há universidades que formam duas turmas – uma para ingressar no primeiro semestre e outra para ingressar no segundo, e o ingresso é feito por ordem de colocação no vestibular.

18. Um resultado interessante é discutido por Ribas, Sampaio e Trevisan (2020). Os autores encontram que alunos com menores notas de ingresso no primeiro semestre tendem a trocar mais de curso que os alunos com as melhores notas no segundo semestre. Contudo, isso ocorre na margem entre quem ingressou no primeiro e no segundo semestre e possui notas semelhantes.

do Fies e do Proni, mas pouco se discute sobre o custo do aluno da universidade pública, que, de modo geral, possui nível socioeconômico mais elevado.

Sobre o grau acadêmico, é interessante notar que alunos em cursos tecnológicos tendem a evadir menos, o que estaria de acordo com a proposta do curso, mais curto e com formação mais específica que outras graduações. Isso pode estar associado com ingressantes cujos objetivos já estão mais consolidados. A razão de conclusão também favorece alunos nesse tipo de curso, no que concerne aos bacharéis.

Em relação aos diferentes cursos, a categoria omitida foi a do curso de ciências sociais, negócios e direito. O senso comum de que cursos na área de exatas são mais difíceis de concluir encontra respaldo parcial nas evidências. O maior risco relativo de evasão ocorre para cursos da área de ciências, matemática e computação; por sua vez, o risco de evasão na área de engenharia e afins não é maior que o observado para ciências sociais, negócios, direito, humanidades e artes. Ademais, os cursos na área de educação são os que apresentam menor razão de sub-risco de evasão.

Para as diferentes organizações acadêmicas, com a base de referência sendo as universidades, somente as faculdades apresentaram uma razão de sub-risco de evasão significativa maior que 1, além de não haver diferenças em termos de conclusão. Entretanto, o risco de transferência é menor tanto para faculdades quanto para centros universitários e institutos tecnológicos, o que é consistente com a menor variabilidade de oferta de cursos nesses institutos. Isso sugere que as transferências ocorrem para cursos não tão similares entre si, fato que reforçaria a hipótese de desalinhamento das expectativas do aluno.

Os resultados desta subseção documentam o perfil do aluno com maior risco de evasão: homens, negros, que ingressam com 25 anos ou menos no ensino superior em instituição privada sem financiamento. O perfil sugere que há um componente socioeconômico aliado à restrição de crédito. Isso já traz elementos suficientes para auxiliar no aprimoramento de políticas públicas relacionadas ao ensino superior. Contudo, além disso, os resultados sugerem que parte da evasão também ocorre por conta de diferenças entre a expectativa do aluno e a realidade enfrentada por ele após o ingresso e por conta das diferentes percepções do custo de evadir. Se, por um lado, é preocupante constatar que mais da metade dos ingressantes não conclui o curso escolhido, por outro lado, parte da evasão tende a ser benéfica ao não prolongar um *mismatch* inicial, condicional ao fato de ter ocorrido o ingresso no ensino superior. Quão benéfica depende de quanto a frustração de expectativas é capaz de explicar a evasão. Fatores como *background* familiar e até mesmo de cunho socioemocional – que podem ou não ser correlacionados com as covariáveis incluídas no modelo – não foram controlados neste estudo e podem influenciar em maior ou menor grau a probabilidade de evasão.

4.2 Robustez

Os resultados da subseção 4.1 excluem alunos ingressantes em cursos a distância, incluem alunos que tiveram sua trajetória escolar corrigida e trata-se dos ingressantes em 2010. Dois exercícios de robustez foram realizados para avaliar o efeito dessa seleção amostral nos resultados e um para observar se os efeitos se mantêm para alunos de outras coortes.

No primeiro exercício, foram incluídos na estimação os alunos que cursavam EaD. No segundo, eliminaram-se os alunos que apresentaram alguma inconsistência em sua trajetória. A tabela 3 reporta a razão de sub-risco de evasão para as duas especificações, além de apresentar novamente os resultados da tabela 2 para facilitar a comparação.

Observa-se que os resultados são muito semelhantes entre si. Nos dois exercícios, as estimativas pontuais são muito próximas, mantendo a direção da relação entre as covariáveis e o risco de evasão. Em alguns casos, houve perda de significância, mas nada que altere significativamente as conclusões anteriores. Adicionalmente, nota-se que não há diferença no risco de evasão para alunos de EaD em comparação com alunos de cursos presenciais diurnos (matutino e vespertino). A magnitude da razão de sub-risco é relativamente alta, mas não é estatisticamente significativa.

O terceiro exercício estima o mesmo modelo para coortes de 2010 a 2014 juntas, a fim de observar se o padrão observado para ingressantes de 2010 se mantém para outras coortes de alunos, ou se os resultados são uma especificidade da coorte de 2010. Como o último ano do painel é 2017, os ingressantes de 2014 apenas podem ser acompanhados por quatro anos consecutivos. Assim, nesse exercício, as coortes de 2010 a 2014 somente foram observadas por quatro anos, por exemplo, a coorte de 2010 é acompanhada até 2013; a coorte de 2011 é acompanhada até 2014 etc. Apesar do período mais curto, como visto no gráfico 1, boa parte dos alunos já evadiu até o quarto ano do curso.

TABELA 3
Análise de robustez: inclusão de alunos de EaD e exclusão de alunos com trajetória escolar corrigida (2010)

Variáveis		Razões de sub-risco de evasão		
		Original	Com EaD	Sem inconsistentes
Características dos alunos				
Sexo	(referência: masculino)			
	Feminino	0,85***	0,86***	0,83***
Cor	(referência: brancos)			
	Não brancos	1,08***	1,08***	1,09***
	Não declarado ou dado faltante	1,29***	1,26***	1,35***
Idade	(referência: 26 anos ou mais)			
	25 anos ou menos	1,21***	1,24***	1,22***
Semestre ingresso	(referência: segundo semestre)			
	Primeiro semestre	0,95***	0,97	0,97
Financiamento	(referência: públicas)			
	Privada sem financiamento	1,38***	1,40***	1,47***
	Privada com outros financiamentos	0,97	0,99	0,95
	Privada com Fies	0,59***	0,60***	0,43***
	Privada com Prouni	0,64***	0,67***	0,58***
Características do curso				
Turno	(referência: matutino e vespertino)			
	Noturno	1,07***	1,08***	1,07***
	Integral	0,79***	0,78***	0,75***
	EaD	-	1,20	-
Grau acadêmico	(referência: bacharelado)			
	Licenciatura	1,46**	1,39**	1,37
	Tecnólogo	0,93***	0,91*	0,95**
Cursos (OCDE – geral)	Agricultura e veterinária	1,06*	1,06	1,12***
	Ciências, matemática e computação	1,17***	1,16***	1,23***
	Educação	0,70**	0,72**	0,75
	Engenharia, produção e construção	1,03	1,03	1,08***
	Humanidades e artes	1,04	1,03	1,02
	Saúde e bem-estar social	0,96**	0,98	0,97
	Serviço	1,06	0,96	1,04
Características da instituição				
Organização acadêmica	(referência: universidade)			
	Faculdade	1,12***	1,09	1,12**
	Centro universitário ¹	1,08	1,03	1,10*
Total de observações		2.017.696	2.419.415	1.503.239

Elaboração dos autores.

Nota: ¹ Institutos federais de educação, ciência e tecnologia e Cefets.

Obs.: Nível de significância: * $p < 0,10$; ** $p < 0,05$; e *** $p < 0,01$.

A tabela A.2, no apêndice A, apresenta os resultados da estimação. Observa-se que os resultados permanecem os mesmos em termos de direção e significância dos coeficientes, com exceção do semestre de ingresso e do grau acadêmico. Logo, com esse exercício, os dados sugerem que as conclusões retiradas da coorte de 2010 podem ser generalizadas para os ingressantes da primeira metade da década de 2010.

4.3 CIF

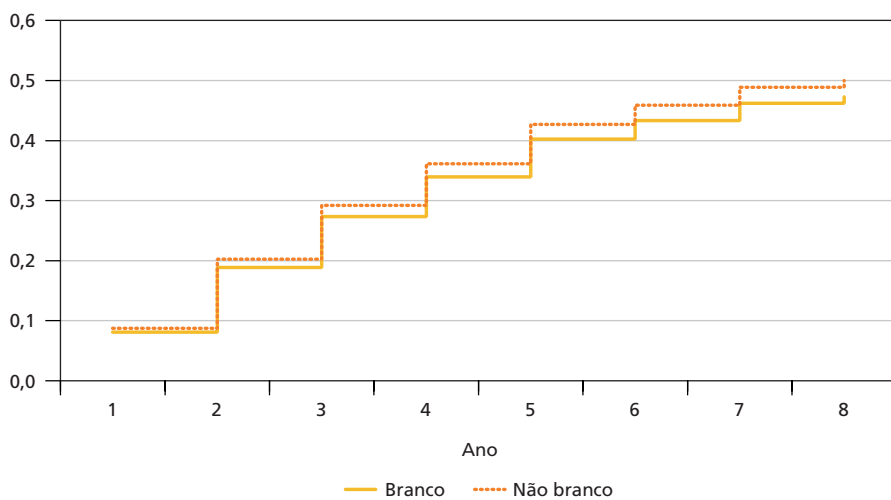
Nesta subseção, são apresentadas as CIFs para evasão. Conforme discutido na seção 2, a CIF traz a taxa acumulada de incidência do evento ao longo do tempo. Assim, é possível observar a evolução da probabilidade de evasão ao longo do tempo, bem como comparar grupos. Ao analisar um grupo específico, as covariáveis restantes são mantidas em seus valores médios. Note-se que a CIF não se refere à taxa instantânea de evasão analisada e discutida nas subseções anteriores.

O gráfico 2 apresenta as CIFs para sexo (mulheres *versus* homens), raça (negros *versus* brancos), idade (25 anos ou menos *versus* 26 anos ou mais), curso (ciências, matemática e computação *versus* outros) e tipo de financiamento, este último separado em dois gráficos para melhor visualização (no primeiro: pública, privada sem financiamento e privada com outros financiamentos; no segundo: pública e privada com o Fies ou o Prouni).

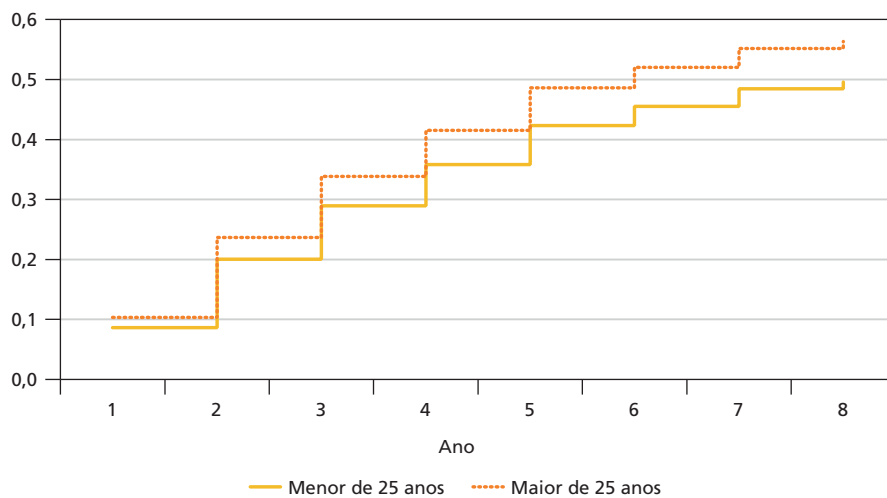
A primeira característica que chama atenção, já observada anteriormente, é a alta probabilidade de evasão em sete anos. Em diversos casos, a probabilidade de evasão alcança 50% após seis anos de ingresso no curso.

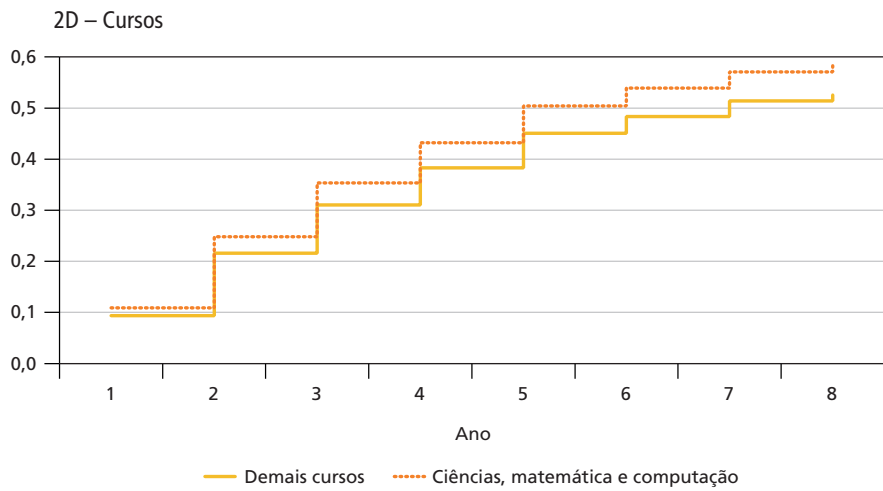
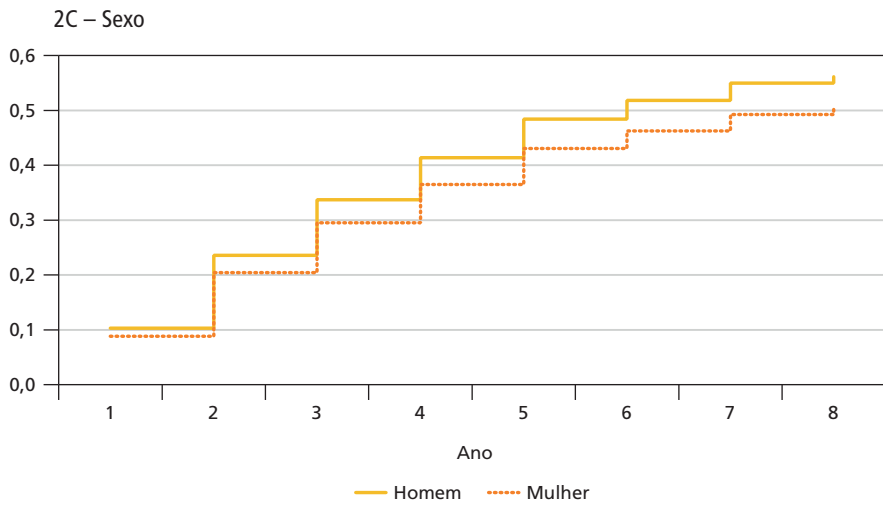
Ademais, nota-se que o maior salto na probabilidade de evasão ocorre entre o segundo e o terceiro ano de curso, o que representa metade do caminho para a graduação na maioria dos cursos. Esse resultado generaliza a evidência reportada por Saccaro, França e Jacinto (2019) para cursos na área de matemática, engenharia e afins. Outra característica comum aos casos apresentados é a probabilidade de evasão decrescente ao longo do tempo. Seja por apresentar resiliência, seja por superar restrições de crédito, alcançar a segunda metade do curso reduz o crescimento das chances de evasão.

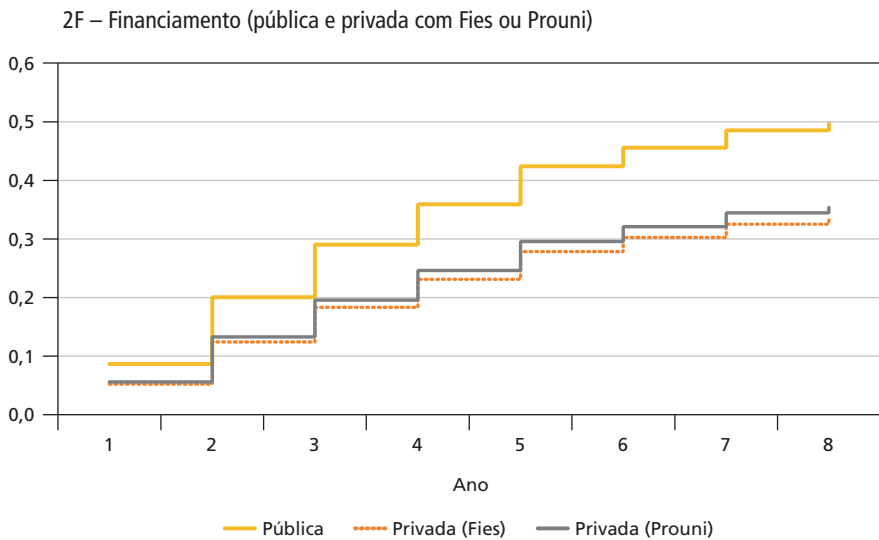
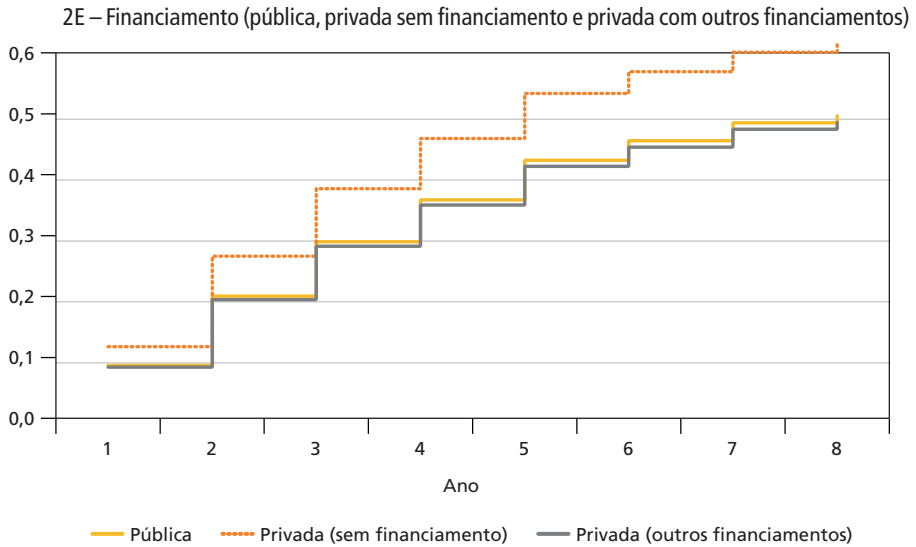
GRÁFICO 2
CIF separada por grupos (2010)
2A – Cor



2B – Idade







Elaboração dos autores.

Comparando-se as trajetórias nos grupos, nota-se que a diferença na probabilidade de evasão tende a aumentar com o tempo, aumento que varia entre os grupos. Mulheres, brancos e os que ingressam com menos de 26 anos possuem uma “inclinação” menos acentuada que homens, negros e os que ingressam com mais de 25 anos, respectivamente. Isso também ocorre na comparação entre alunos do curso de ciências, matemática e computação com os demais cursos –

a probabilidade de evasão é maior para o primeiro grupo desde o primeiro ano de curso. Assim, não há evidência de que o *timing* da evasão seja diferente para as categorias nos grupos analisados, ou seja, a evasão impacta a todos no mesmo período. A diferença é que alguns grupos são mais impactados que outros.

Isso também é válido na comparação por tipo de financiamento (gráficos 2E e 2F). Alunos da rede privada sem financiamento partem de um nível mais alto de probabilidade de evasão no primeiro ano, seguidos pelos alunos da rede pública e da rede privada com financiamento. Por sua vez, os estudantes que recebem Fies ou participam do Prouni apresentam menor probabilidade de evasão inicialmente. Os três grupos permanecem relativamente juntos ao longo do tempo, mas a distância entre eles torna-se cada vez maior. Após sete anos, enquanto alunos do Fies/Prouni apresentam probabilidade de evasão em torno de 35%, alunos da rede pública aproximam-se de 50%, e os da rede privada sem financiamento alcançam 60%.

Esta subseção documentou diversos padrões associados à probabilidade de evasão. Em particular, a evasão tende a ocorrer logo no início do curso, fato que é condizente tanto com a hipótese de restrição de crédito quanto com a de autoconhecimento. Todavia, a diferença por tipo de financiamento sugere que essas hipóteses podem valer para grupos distintos associados em nível socioeconômico. A restrição de crédito atuaria mais sobre os grupos mais vulneráveis, elegíveis ao Fies e Prouni, que são provavelmente os que ingressam mais velhos, enquanto a hipótese de autoconhecimento estaria mais presente nos alunos da rede pública e da rede privada com outros financiamentos.

5 CONCLUSÃO

Este estudo analisou os fatores que mais influenciam a evasão dos estudantes do ensino superior no Brasil. Para isso, acompanhou-se a trajetória da coorte ingressante de 2010 com base nos microdados do Censo da Educação Superior 2010 a 2017. Foi aplicado um modelo de análise de sobrevivência com abordagem de riscos competitivos.

Dos alunos que ingressaram em 2010, cerca de 60% não concluíram o curso após oito anos. Entender o que gerou esse cenário é fundamental tanto do ponto de vista individual quanto do ponto de vista da sociedade. O ponto de partida é a documentação do fenômeno para que se possa pensar em soluções que tornem as escolhas mais eficientes.

Os resultados mostram que mulheres, brancos, ingressantes mais velhos e alunos que utilizam o Prouni ou o Fies possuem menor risco de evasão. A evidência mostra ainda que a chance de evasão é maior na primeira metade do curso. Além disso, diferenças entre categorias de um mesmo grupo (gênero, racial, cursos etc.) aparecem já no primeiro ano e aumentam com o passar do tempo.

Se, por um lado, foram identificadas características que estariam associadas a um risco maior de evasão, por outro lado, permanecem em aberto o papel de fatores relevantes, como o nível socioeconômico e as opções de fora (mercado de trabalho) que podem concorrer com a frequência ao curso superior. Uma das limitações deste artigo é a ausência de informações que permitam avaliar o efeito de tais fatores. Nesse sentido, os resultados devem ser interpretados com cautela, sendo necessários mais estudos para corroborar ou não as evidências apresentadas. Outra característica ausente na análise é a política de cotas. Em 2010, ainda havia poucas universidades utilizando esse mecanismo na forma de ingresso.

Finalmente, se o motivo da evasão estiver associado ao autoconhecimento – o artigo apresenta alguma evidência consistente com essa hipótese –, a evasão pode ser benéfica no longo prazo, condicional ao aluno ter ingressado no ensino superior. O custo, financeiro e psicológico, de prolongar o *mismatch* poderia ser menor tanto para o aluno quanto para a sociedade caso a evasão ocorresse rapidamente. Isso sugere que a sociedade teria a ganhar se houvesse maior exposição dos alunos durante a educação básica aos cursos e às profissões, de tal modo que isso ampliasse o conjunto de informações que será usado para basear as escolhas ao fim do ensino médio.

REFERÊNCIAS

- ANDRIOLA, W. B.; ANDRIOLA, C. G.; MOURA, C. P. Opiniões de docentes e de coordenadores acerca do fenômeno da evasão discente dos cursos de graduação da Universidade Federal do Ceará (UFC). **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 14, n. 52, p. 365-382, 2006.
- AUSTIN, P. C.; FINE, J. P. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. **Statistics in Medicine**, v. 36, n. 27, p. 4391-4400, 2017.
- CARDAK, B. A.; VECCI, J. Graduates, dropouts and slow finishers: the effects of credit constraints on university outcomes. **Oxford Bulletin of Economics and Statistics**, v. 78, n. 3, p. 323-346, 2015.
- FENG, L.; SASS, T. R. **Competing risks analysis of dropout and educational attainment for students with disabilities**. Atlanta: Georgia State University, fev. 2012. (UWRG Working Papers, n. 48).
- FINE, J. P.; GRAY, R. J. A proportional hazards model for the subdistribution of a competing risk. **Journal of the American Statistical Association**, v. 94, n. 446, p. 496-509, 1999.
- GESKUS, R. B. **Data analysis with competing risks and intermediate states**. Boca Raton: Chapman & Hall; CRC Press, 2015.

HAAS, C.; HADJAR, A. Students' trajectories through higher education: a review of quantitative research. **Higher Education**, v. 79, n. 6, p. 1099-1118, 2020.

HENDRICKS, L. A.; LEUKHINA, O. **The return to college**: selection bias and dropout risk. [s.l.]: SED, 2011. (Meeting Papers, n. 311).

INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Metodologia de cálculo dos indicadores de fluxo da educação superior**. Brasília: Inep, 2017.

JOHNSON, M. T. Borrowing constraints, college enrollment, and delayed entry. **Journal of Labor Economics**, v. 31, n. 4, p. 669-725, 2013.

LIMA JUNIOR, P.; SILVEIRA, F. L.; OSTERMANN, F. Análise de sobrevivência aplicada ao estudo do fluxo escolar nos cursos e graduação em física: um exemplo de uma universidade brasileira. **Revista Brasileira de Ensino de Física**, v. 34, n. 1, p. 1-10, 2012.

MEGGIOLARO, S.; GIRALDO, A.; CLERICI, R. A multilevel competing risks model for analysis of university students' careers in Italy. **Studies in Higher Education**, v. 42, n. 7, p. 1259-1274, 2017.

MENDES JUNIOR, A. A. F. Uma análise da progressão dos alunos cotistas sob a primeira ação afirmativa brasileira no ensino superior: o caso da Universidade do Estado do Rio de Janeiro. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 22, n. 82, p. 31-56, 2014.

RIBAS, R. P.; SAMPAIO, B.; TREVISAN, G. Short-and long-term effects of class assignment: evidence from a flagship university in Brazil. **Labour Economics**, v. 64, p. 1-13, jun. 2020.

SACCARO, A.; FRANÇA, M. T. A.; JACINTO, P. de A. Fatores associados à evasão no ensino superior brasileiro: um estudo de análise de sobrevivência para os cursos das áreas de ciência, matemática e computação e de engenharia, produção e construção em instituições públicas e privadas. **Estudos Econômicos**, São Paulo, v. 49, n. 2, p. 337-373, abr.-jun. 2019.

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

SILVA, G. P. da. Análise de evasão no ensino superior: uma proposta de diagnóstico de seus determinantes. **Avaliação: Revista da Avaliação da Educação Superior**, v. 18, n. 2, p. 311-333, 2013.

STINEBRICKNER, R.; STINEBRICKNER, T. The effect of credit constraints on the college drop-out decision: a direct approach using a new panel study. **American Economic Review**, v. 98, n. 5, p. 2163-2184, 2008.

STINEBRICKNER, R.; STINEBRICKNER, T. Academic performance and college dropout: using longitudinal expectations data to estimate a learning model. **Journal of Labor Economics**, v. 32, n. 3, p. 601-644, 2014.

APÊNDICE A

TABELA A.1
Distribuição dos eventos por característica do aluno, curso e instituição (2010)

Variáveis		Total (1 mil)	Total (%)	Cursando (%)	Evadiu (%)	Concluiu (%)	Transferiu (%)
Características do aluno							
Sexo	Masculino	921,5	46	45	49	38	52
	Feminino	1.096,4	54	55	51	62	48
Cor	Branco	521,1	26	30	22	27	20
	Não branco	300,8	15	18	13	14	9
	Não declarado ou dado faltante	1.196,0	59	52	65	60	71
Idade	25 anos ou menos	1.030,5	51	45	57	48	68
	26 anos ou mais	987,4	49	55	43	52	32
Semestre ingresso	Primeiro semestre	1.324,2	66	59	67	76	65
	Segundo semestre	693,7	34	41	33	24	35
	Pública	453,2	22	30	19	14	19
	Privada sem financiamentos	981,7	49	35	60	50	56
	Privada com outros financiamentos	345,3	17	17	15	23	15
	Fundo de Financiamento Estudantil (Fies)	116,4	6	11	2	4	4
	Programa Universidade para Todos (Prouni)	121,3	6	7	4	9	6
Características do curso							
Turno	(Matutino e vespertino)	460,0	23	25	22	19	24
	Noturno	1.340,1	66	59	70	75	65
	Integral	217,8	11	16	8	6	11
Grau acadêmico	Bacharelado	1.386,8	69	80	67	47	72
	Licenciatura	331,2	16	14	16	22	17
	Tecnólogo	299,9	15	6	17	31	11

(Continua)

(Continuação)

Variáveis		Total (1 mil)	Total (%)	Cursando (%)	Evadiu (%)	Concluiu (%)	Transferiu (%)
Características do aluno							
Cursos (Organização para a Cooperação e Desenvolvimento Econômico – OCDE – geral)	Agricultura e veterinária	44,8	2	3	2	1	1
	Ciências, matemática e computação	157,8	8	7	9	8	10
	Ciências sociais, negócios e direito	845,9	42	41	43	43	33
	Educação	330,8	16	14	16	22	16
	Engenharia, produção e construção	260,2	13	16	12	5	23
	Humanidades e artes	59,6	3	2	3	4	4
	Saúde e bem-estar social	268,6	13	16	12	12	10
	Serviço	50,2	2	1	3	5	2
Características da instituição							
Organização acadêmica	Universidade	990,6	49	54	45	44	66
	Faculdade	706,6	35	32	37	39	18
	Centro universitário/ outros ¹	320,7	16	14	17	17	16

Fonte: Censos da Educação Superior do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), referente ao período 2010-2017. Disponível em: <http://portal.inep.gov.br/basicas-levantamentos-acessar>. Acesso em: 13 jun. 2016. Elaboração dos autores.

Nota: ¹ Institutos federais de educação, ciência e tecnologia e centros federais de educação tecnológica (Cefets).

TABELA A.2

Modelo de riscos competitivos aplicado à evasão com cluster no nível da instituição de ensino superior (IES) – ingressantes de 2010 a 2014

Variáveis	Razões de sub-risco	
	Evadiu	
Características dos alunos		
Sexo	(referência: masculino)	
	Feminino	0,88***
Cor	(referência: brancos)	
	Não brancos	1,14***
	Não declarado ou dado faltante	1,54***
Idade	(referência: 25 anos ou menos)	
	26 anos ou mais	1,27***
Semestre ingresso	(referência: segundo semestre)	
	Primeiro semestre	1,05***

(Continua)

(Continuação)

Variáveis	Razões de sub-risco	
	Evadiu	
Características dos alunos		
		(referência: públicas)
Financiamento	Privada sem financiamento	1,43***
	Privada com outros financiamentos	0,94
	Privada com Fies	0,61***
	Privada com Prouni	0,59***
Características do curso		
		(referência: matutino e vespertino)
Turno	Noturno	1,08***
	Integral	0,79***
		(referência: bacharelado)
Grau acadêmico	Licenciatura	1,18
	Tecnólogo	1,07***
	Agricultura e veterinária	1,06*
	Ciências, matemática e computação	1,17***
	Educação	0,91
Cursos (OCDE – geral)	Engenharia, produção e construção	1,08
	Humanidades e artes	0,98
	Saúde e bem-estar social	1,02
	Serviço	1,04**
Características da instituição		
Organização acadêmica	Faculdade	1,15***
	Centro universitário, instituto federal de educação, ciência e tecnologia e centro federal de educação tecnológica	1,16

Elaboração dos autores.

Obs.: * $p < 0,10$; ** $p < 0,05$; *** $p < 0,01$.

Originais submetidos em: dez. 2020.

Última versão recebida em: nov. 2021.

Aprovada em: nov. 2021.

PROGRAMAS DE BONIFICAÇÃO A DOCENTES E RENDIMENTO ESCOLAR: UMA AVALIAÇÃO DO PROGRAMA ESCOLA DE VALOR NO ESTADO DA PARAÍBA^{1,2}

Gabriella da Silva Cavalcanti³

Joelson Oliveira Santos⁴

Ana Cláudia Annegues da Silva⁵

O objetivo deste estudo consiste em avaliar o impacto do programa de bonificação Escola de Valor, do governo do estado da Paraíba, edição 2016, sobre alguns indicadores escolares. Foram aplicados os métodos de *propensity score matching* e diferenças em diferenças sobre os microdados do Sistema de Avaliação da Educação Básica (Saeb) e do Censo Escolar referentes aos anos de 2015 e 2017. Os resultados das estimações revelaram impactos positivos do programa de bonificação sobre a nota dos alunos das escolas premiadas, bem como sobre a taxa de aprovação e a taxa de abandono dos estudantes. Ademais, a significância estatística das *dummies* de interação entre a quantidade de premiações e o tempo sugere que o programa tende a apresentar efeito positivo e cumulativo sobre o desempenho escolar.

Palavras-chave: Prêmio Escola de Valor; pagamento a professores; avaliação de impacto; desempenho escolar; *accountability*.

TEACHER BONIFICATION AND SCHOOL PERFORMANCE: AN EVALUATION OF "ESCOLA DE VALOR" PROGRAM IN THE STATE OF PARAÍBA, BRAZIL

The objective of this study is to evaluate the impact of the Value School bonus program of the Government of Paraíba State, 2016 edition, on some school indicators. The methods of Propensity Score Matching and Differences in Differences were applied on Saeb microdata and School Census for the years 2015 and 2017. The results of the estimates revealed a positive impact of the bonus on the score of students from award-winning schools, as well as on the pass rate and dropout rate of students. Furthermore, the statistical significance of the interaction dummies between the number of awards and time suggests that the program tends to have a positive and cumulative effect on school performance.

Keywords: Value School Premium; payment to teachers; impact evaluation; scholar performance; *accountability*.

JEL: C31; I21; I28.

1. DOI: <http://dx.doi.org/10.38116/ppe52n3art2>

2. Gabriella Cavalcanti e Joelson Oliveira agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) as bolsas de mestrado. Ana Annegues agradece à Capes a bolsa de pós-doutorado.

3. Agente local de inovação do Serviço Brasileiro de Apoio à Micro e Pequena Empresa (Sebrae). *E-mail:* gabriellas.cavalcanti@gmail.com.

4. Doutorando em economia pelo Programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande do Sul (PPGE/UFRGS). *E-mail:* joelsonsantosrdp@hotmail.com.

5. Economista da Universidade Federal da Paraíba (UFPB). *E-mail:* annegues.ana@gmail.com.

1 INTRODUÇÃO

A importância da melhoria dos sistemas educativos na promoção de competências técnicas e socioemocionais aos indivíduos vem sendo amplamente reconhecida pela sociedade, assim como suas externalidades positivas à coletividade, como o aumento na produtividade do país ou a construção de cidadãos mais conscientes. Todavia, a melhoria desses sistemas vem se mostrando uma tarefa desafiadora para a sociedade brasileira. De acordo com o Programa Internacional de Avaliação de Alunos (Pisa), em 2018, o desempenho dos alunos brasileiros esteve abaixo da média dos alunos dos países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) em ciências (404 pontos, comparados à média de 489 pontos), em leitura (413 pontos, comparados à média de 487 pontos) e em matemática (384 pontos, comparados à média de 492 pontos). Tais resultados colocam o país na 66ª posição em ciências, na 57ª em leitura e na 70ª colocação em matemática, em um total de 79 nações. Dado que a amostra brasileira contou com 10.691 estudantes de 597 escolas, representando uma cobertura de 65% dos estudantes de 15 anos elegíveis à realização do teste, constata-se que a qualidade da formação de capital humano está distante da necessária a um país que almeja um alto nível de desenvolvimento sustentável.

Seguindo uma tendência internacional, a partir dos anos 1990, a expansão das avaliações externas de ensino intensificou o surgimento no país das políticas de responsabilização educacional. Bonamino e Sousa (2012) dividem a avaliação da educação brasileira em três períodos. O primeiro, de caráter diagnóstico, visava apenas à estimação dos resultados alcançados pelas escolas, porém sem a divulgação destes. No segundo, os resultados passam a ser entregues às escolas e divulgados à sociedade, constituindo uma política de responsabilização do tipo *low-stake*, dado que as consequências da divulgação seriam de caráter simbólico. No terceiro período, os resultados das avaliações começam a ser atrelados a consequências materiais, tais como incentivos monetários, vinculando o recebimento de bonificação ao cumprimento de metas pré-estabelecidas, o que denota uma política de responsabilização do tipo *high-stake*.

As políticas de responsabilização *high-stake* podem ser desenhadas com base em consequências materiais punitivas ou bonificatórias, as quais podem ser individuais ou coletivas. Até o momento, diferentemente do que ocorre em alguns países, o Brasil tem adotado apenas políticas bonificatórias, nas quais são concedidos incentivos salariais aos professores de disciplinas que são avaliadas pelo cumprimento de metas educacionais, no caso de políticas de incentivo individualizado, ou benefícios aos demais funcionários da escola, no caso de incentivos coletivos.

A análise empírica acerca da efetividade de tais políticas na promoção de melhorias no desempenho dos alunos em testes padronizados tem mostrado

resultados inconclusivos. Encontram-se resultados positivos (Lavy, 2002; Muralidharan e Sundararaman, 2011), mas também nulos ou negativos (Cowan e Goldhaber, 2018; Fryer Junior, 2011). No Brasil, poucos são os estudos empíricos na área (Furtado e Soares, 2018; Lépine, 2016; Oshiro, Scorzafave e Dorigan, 2015), e, tal qual o observado na literatura internacional, os resultados não apontam em direção a um consenso. Nesse sentido, a literatura sobre esse tema carece de mais testes empíricos, sobretudo de análises voltadas ao contexto brasileiro.

Com vistas a oferecer maiores contribuições ao debate dentro e fora da academia, surge a ideia presente neste trabalho, cujo objetivo é avaliar a efetividade do programa de bonificação Escola de Valor sobre indicadores educacionais, como a proficiência média em português e matemática, a taxa de aprovação escolar e a taxa de abandono entre os alunos das escolas contempladas. O programa consiste em uma iniciativa do governo do estado da Paraíba e foi instituído nos termos da Lei nº 9.879, de 13 de setembro de 2012. Tem por objetivo fomentar, selecionar e valorizar as práticas pedagógicas exitosas no processo de ensino e aprendizagem, premiando com um 14º salário todos os profissionais de educação em exercício nas escolas públicas estaduais de educação básica escolhidas por meio de um processo seletivo entre as escolas do estado.

Para atender ao objetivo proposto, o estudo lança mão de duas metodologias de avaliação de impacto de programas sociais, que são: o método de pareamento por *propensity score matching* (PSM) e o diferenças em diferenças (DID). O PSM busca construir um contrafactual observável do grupo de escolas premiadas selecionando as escolas não premiadas de acordo com o grau de similaridade quanto à propensão a serem premiadas. O objetivo é identificar a existência de um suporte comum entre tratamento e controle, com base em características observáveis dos grupos de tratamento e controle. Em seguida, por meio do cálculo de uma dupla diferença, comparam-se os resultados dos dois grupos antes e depois da intervenção, de modo a eliminar diferenças por não observáveis fixas no tempo entre as escolas do grupo de controle e as afetadas pelo tratamento.

Utilizam-se os microdados do Sistema de Avaliação da Educação Básica (Saeb) e do Censo Escolar referentes aos anos de 2015 e 2017 para as escolas públicas (municipais e estaduais) do estado da Paraíba, disponibilizados por meio do portal do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio de Teixeira (Inep).

Os resultados das estimações dão indícios de impacto positivo do programa de bonificação aos docentes da Paraíba sobre a proficiência média em português e matemática, taxa de aprovação e taxa de abandono escolar dos alunos do 5º e 9º anos do ensino fundamental das escolas premiadas em 2016, corroborando os resultados obtidos nos estudos de Furtado e Soares (2018) – em análise específica para o estado de Pernambuco –, Oshiro, Scorzafave e Dorigan (2015) e Lépine (2016) –

estes dois últimos trabalhos aplicados ao contexto do estado de São Paulo —, que encontraram efeitos positivos dos programas de bonificação desses estados sobre as notas médias dos alunos.

Além desta introdução, compõem o trabalho mais cinco seções. Na seção 2, efetua-se uma resenha acerca dos aportes teóricos que fundamentam as políticas de remuneração dos profissionais por desempenho na educação, bem como apontam-se alguns resultados empíricos presentes na literatura. A seção 3, por sua vez, destina-se à descrição do programa avaliado, enquanto a seção 4 versa sobre a estratégia empírica utilizada. A seção 5 analisa os resultados obtidos e, por fim, na seção 6, apontam-se as principais conclusões do trabalho.

2 PROGRAMAS DE PAGAMENTO DE BÔNUS AOS DOCENTES: DOS ASPECTOS TEÓRICOS À ABORDAGEM EMPÍRICA

Tomando por referência Blaug (1992⁶ *apud* Waltenberg, 2006), as linhas de pesquisa em economia da educação podem ser classificadas em dois campos principais: i) análises do valor econômico da educação, as quais tratam da importância de medidas agregadas de educação para o crescimento de um país ou região; e ii) aspectos econômicos dos sistemas educacionais. Waltenberg (2006) assinala que a segunda linha de pesquisa situa-se em uma investigação microeconômica dos sistemas educativos. Assim, nessa abordagem, a ênfase recai sobre “análises de custos e benefícios privados e sociais (...), análises de eficiência na alocação de recursos (...), definição de objetivos escolares e pós-escolares de sistemas educativos, avaliação de atingimento desses objetivos” (Waltenberg, 2006, p. 119).

Entre os aspectos abordados no segundo campo, a literatura econômica tem concentrado esforços na identificação dos principais determinantes da aprendizagem escolar e, conseqüentemente, na formulação e avaliação de políticas públicas que logrem níveis adequados de aprendizado aos estudantes. Uma das medidas, associada a essa linha de raciocínio, diz respeito à reformulação da estrutura de incentivos oferecida a um dos principais agentes do processo de aprendizado, o professor. Tais políticas atrelam parte da remuneração dos professores ao desempenho dos alunos em testes padronizados, caso do objeto de estudo deste trabalho.

Conforme salientam Alexandre, Lima e Waltenberg (2014), essas políticas não encontram fundamento teórico no campo da educação, mas se baseiam nos pressupostos teóricos oriundos da teoria dos incentivos.⁷ Tal estrutura de análise parte da premissa de que mecanismos de incentivos monetários por resultados harmonizam o objetivo dos agentes (professores) com o do principal (secretarias de educação etc.), logrando maior aprendizado dos alunos. Diante de tais

6. Blaug, M. *The methodology of economics*. 2. ed. Cambridge, Reino Unido: Cambridge University Press, 1992.

7. Ver Laffont e Martimort (2002).

considerações, nesta seção serão discutidos os aspectos teóricos que norteiam os programas de pagamento de bônus aos docentes, bem como os resultados encontrados na literatura acerca da temática.

2.1 Aspectos teóricos

Conforme mencionado previamente, programas de pagamentos de bônus a docentes baseiam-se nos pressupostos teóricos do modelo principal-agente, o qual é comumente utilizado na análise, no estabelecimento e na manutenção de contratos entre agentes econômicos envolvidos em assimetria de informação. Inicialmente proposto para caracterizar relações de trabalho individuais, apenas um principal e um agente no setor privado, variações do modelo foram desenvolvidas para adequar-se ao setor público e às peculiaridades do sistema educacional, bem como para apreciações envolvendo múltiplos agentes.

O modelo parte da premissa de que o principal depende de tarefas a serem executadas por um agente, que dispõe de vantagem informativa sobre seu próprio comportamento, tipo ou ambiente, não disponível ao principal. Tradicionalmente, a problemática pode ser exemplificada por meio do estabelecimento de um contrato, mediante o qual determinado principal induz um agente a realizar certa ação por ele desejada, como o empenho na execução de uma determinada tarefa. Partindo da premissa de assimetria de informação entre os agentes, o monitoramento e a avaliação dessa ação pelo principal, ou qualquer outro agente, pode não ser diretamente possível. Contudo, resultados indiretamente determinados pela ação do agente podem ser observados. Logo, o principal deverá incorporar ao contrato mecanismos que incentivem, indiretamente, o agente a agir da melhor forma possível do ponto de vista do principal.

Do ponto de vista do agente, duas restrições o envolvem ao se estabelecer o contrato: a restrição de participação e a restrição de compatibilidade de incentivo. Na primeira, o incentivo proposto pelo principal deve garantir um nível mínimo de utilidade para o agente, de modo que o incentivo supra a utilidade proveniente das demais possibilidades a ele disponíveis. Dado que o principal não é capaz de determinar diretamente o comportamento do agente, esse o influenciará a agir de maneira ótima, o que consiste na segunda restrição envolvendo o agente, a proposição de um contrato que induza o agente a agir de acordo com o desejado pelo principal. Assim, o principal deve garantir um nível mínimo de satisfação esperada ao agente, a fim de que o contrato de trabalho proposto seja aceitável. Portanto, o desafio do principal é o de definir incentivos que maximizem seu interesse e que, ao mesmo tempo, sejam atraentes para o agente.

Em trabalho que visa compreender a problemática da remuneração de professores à luz do modelo principal-agente, Alexandre, Lima e Waltenberg (2014)

denotam que, por meio desse aporte teórico, a remuneração ideal de professores passa a depender das relações probabilísticas entre diferentes níveis de desempenho de alunos e distintos níveis de esforço praticado por professores. Os autores enfatizam que o problema não reside no fato de a associação ser probabilística, mas no desconhecimento das funções de distribuição de probabilidades relevantes. Isto é, “não é possível saber com precisão qual é a probabilidade de que o esforço exercido pelo professor tenha sido alto, médio, baixo etc.” (Alexandre, Lima e Waltenberg, 2014, p. 48).

Portanto, é possível que um professor tenha total engajamento e esforço para melhorar o aprendizado de um determinado grupo de alunos sem que esse esforço se reverta efetivamente em bons resultados. Logo, apesar de agir conforme o desejável pelo principal (secretaria escolar), o professor não receberia seu bônus, ocasionando redução de motivação futura. O oposto (pouco esforço, bons resultados e recebimento de bônus) também poderia reduzir a motivação futura. Destarte, em ambos os casos, a legitimidade do programa poderia ser minada (Alexandre, Lima e Waltenberg, 2014).

Mesmo se supondo que o principal conheça com precisão as funções de distribuição de probabilidades relevantes, Alexandre, Lima e Waltenberg (2014) assinalam que, ainda assim, o principal incorreria em um dilema. Caso a remuneração esteja estreitamente atrelada ao desempenho dos alunos, os incentivos seriam maiores, contudo se incorreria na introdução de alta volatilidade à remuneração e, conseqüentemente, no alto grau de risco no sistema de remuneração. Todavia, o inverso ocasiona menor risco, mas incentivos débeis.

Embora tenham procurado inicialmente entender a problemática da remuneração de docentes à luz do modelo principal-agente para relações trabalhistas no setor privado, os autores apontam algumas especificidades de relações laborais em escolas públicas relevantes para a teoria dos incentivos e contratos.

A primeira especificidade diz respeito à ausência de concorrência⁸ e de preocupação com o lucro nas escolas públicas. Nesse caso, escolas privadas estão sujeitas a um ambiente concorrencial, o que lhes impele incentivos externos em busca de lucros para fins de sobrevivência. Tal constatação não é observada no setor público. Assim, se levarmos em consideração que parte do lucro potencial de uma escola privada relaciona-se com o desempenho (notas) de seus alunos em testes padronizados – dado que há uma correlação entre demanda maior e propensão

8. Tais considerações estão atreladas ao sistema educacional brasileiro. Diferentemente do que ocorre em alguns países, como os Estados Unidos, as políticas educacionais brasileiras de consequência material consistem em políticas de incentivo salarial, ao passo que a consequência material também pode relacionar-se à dispensa de professores e ao fechamento de escolas, o que ocasiona maior competitividade entre escolas públicas. Ademais, políticas que promovam a competitividade entre escolas públicas e privadas também estão presentes em alguns países. Para uma análise dessa temática aplicada à realidade polonesa, ver Bukowski e Kobus (2018).

a pagar mensalidades mais onerosas a escolas bem classificadas em *rankings* de desempenho de alunos – e admitindo-se que mecanismos de incentivo salarial sejam capazes de promover melhorias no desempenho cognitivo dos alunos, provavelmente escolas privadas tenham maior interesse em promover tais políticas (Alexandre, Lima e Waltenberg, 2014). A inexistência de pressão por lucro no setor público torna menos evidente o objetivo comum a ser alcançado pelos agentes envolvidos em tais esferas, de modo que um aspecto importante da bonificação no setor público é o de convergir os interesses dos agentes.

A segunda relaciona-se à multiplicidade de relações de agência. Ao tomarmos o setor público, observamos múltiplas relações entre agentes e principais com objetivos variados e, em alguns casos conflitantes, transcendendo à relação biunívoca presente no modelo básico trabalhado até aqui. Em virtude de objetivos não relacionados à educação, o desenho dos mecanismos de incentivos, bem como a desvirtuação de programas inicialmente bem desenhados, pode ocorrer ao se atender, por exemplo, às demandas de determinados grupos políticos.

Os autores assinalam que os agentes possuem motivações não monetárias ou intrínsecas, ao decidirem se dedicar a determinadas ocupações. Caso tais motivações sejam suficientemente fortes no setor público e, em especial, na educação pública, o salário esperado do professor poderá ser relativamente baixo, mas satisfaria a restrição de participação, pois seu esforço ao exercer a atividade não seria considerado oneroso.

A existência de multiplicidade de tarefas e objetivos é apontada por Alexandre, Lima e Waltenberg (2014) como uma das especificidades das escolas, públicas e privadas, que merece atenção. Há diferentes tarefas entendidas como relevantes na educação, de modo que o professor precisa empregar energia e tempo em tarefas distintas a fim de atingir cada uma delas. Logo, para incentivar docentes a se empenharem em todas as tarefas, seria preciso premiá-los por atingir todos os objetivos. Como alguns resultados são intangíveis ou imensuráveis, a remuneração por todos os objetivos alcançados se torna impraticável. Consequentemente, a utilização de objetivos mais tangíveis para fins de premiação é realizada. Conforme assinalam os autores, ocorre, porém, que diferentes principais das relações de agência do setor educacional podem divergir quanto à relevância dos objetivos. Assim, o êxito para uns pode não ser o mesmo para outros.

A quinta especificidade dos sistemas educativos diz respeito ao trabalho em equipe verificado nas escolas, dado que o conhecimento em determinada disciplina pode relacionar-se diretamente ao conhecimento adquirido em outras. Dessa forma, os autores indagam acerca do desenho de incentivos para que o trabalho em equipe possa ser estimulado na medida correta. Apontam que sistemas de remuneração individualizada possuem incentivos mais fortes quando

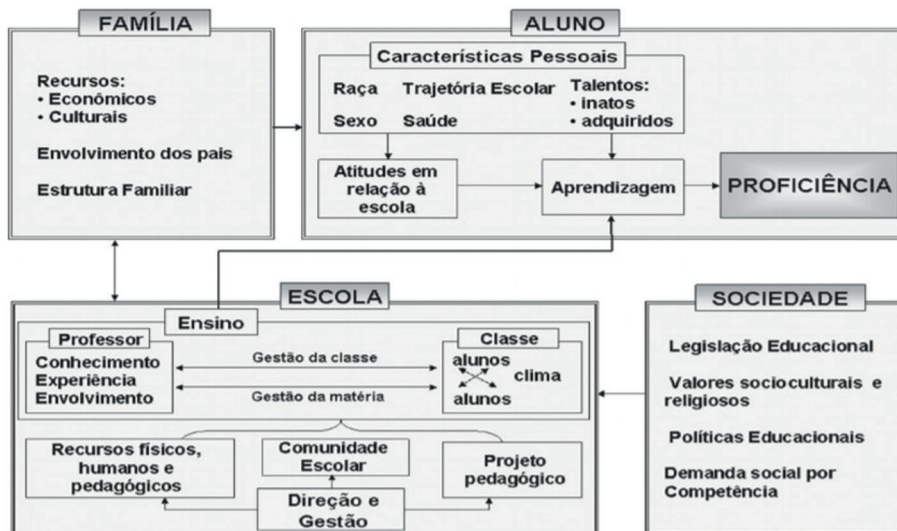
comparados a sistemas de remuneração coletiva, porém, em um contexto de trabalho em equipe, podem gerar sensação de injustiça e competitividade entre docentes, e, no limite, desmotivar professores a trabalharem em equipe ou prejudicar o ambiente de trabalho, prejudicando também o aluno. Em contrapartida, em sistemas de remuneração coletiva, a problemática do chamado *free-rider* pode surgir, isto é, aqueles indivíduos que não se esforçam adequadamente podem usufruir dos benefícios coletivos em decorrência do esforço dos demais.

Também se verifica certa dificuldade em mensurar o produto do processo educativo. O modelo principal-agente básico parte do pressuposto de que não é possível se observar diretamente o esforço, mas que algum resultado ou produto do esforço é observável. Como os objetivos e produtos da educação são variáveis, não se sabe com precisão qual o produto relevante do processo de educação. Conseqüentemente, a avaliação do desempenho de professores por si só é um obstáculo à transposição do modelo à esfera educacional (Alexandre, Lima e Waltenberg, 2014). Mesmo adotando um critério mais objetivo, tal qual a avaliação de desempenho medida por meio das notas dos alunos, a determinação exata da fração da nota obtida por um aluno decorrente do esforço de determinado professor é difícil, uma vez que a nota é reflexo de um conjunto de fatores, correntes e passados, escolares e não escolares, determinísticos e aleatórios (Waltenberg, 2006).

Ademais, a educação possui uma peculiaridade que a distingue de outras atividades: o próprio consumidor (aluno) é um insumo de produção. Waltenberg (2006) aponta que o desempenho do aluno depende da sua dotação de competências cognitivas inicial, bem como do esforço realizado por este durante o processo de educação, seja na escola, seja em casa. Bishop e Woessmann (2004) denotam que o esforço do aluno reflete sua motivação, tempo e envolvimento no aprendizado. Assinalam ainda que o esforço dos alunos é o insumo mais importante no processo educacional, ao exemplificarem que, dada uma razão de alunos por professor de 20 para 1, os alunos gastam cerca de vinte vezes mais horas estudando do que os professores ensinando.

Verifica-se que o aluno se encontra envolto em uma teia de relações sociais e institucionais. Dessa maneira, fatores extraescolares inerentes ao aluno e referentes ao contexto social em que este se insere, bem como fatores intraescolares, são os principais determinantes do desempenho dos alunos, seja esse desempenho mensurado por critérios objetivos ou não. Nesse sentido, Soares (2007) explicita um modelo conceitual, disposto na figura 1, que apresenta as várias inter-relações entre os fatores explicativos do aprendizado e destes com o resultado final, o desempenho cognitivo dos alunos.

FIGURA 1
Modelo conceitual



Fonte: Soares (2007, p. 141).

Obs.: Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Partindo de tal modelo conceitual, Soares (2007) denota que os fatores mais próximos do desempenho do aluno são suas características inatas ou já determinadas por sua história de vida. Além dessas, três outras estruturas influenciariam melhores ou piores desempenhos de alunos: a escola, a família e a sociedade.

Portanto, dado que a mensuração do produto da educação está sujeita a muitos determinantes e ruídos, o desenho dos incentivos de políticas de remuneração por desempenho e, conseqüentemente, a adaptação e transposição do modelo principal-agente à educação devem ser pensados com cuidado. Assim, diante dos requisitos para que os incentivos deem certo e das chances de que alguma falha se manifeste, não é de se estranhar que alguns programas falhem em lograr êxito, seja por critérios convencionais, que focam o desempenho dos estudantes em exames de proficiência, seja segundo critérios mais abrangentes de avaliação.

2.2 Aspectos empíricos: abordagens à realidade internacional e para o Brasil

Após a exposição dos elementos teóricos das políticas de remuneração por desempenho de professores, faz-se necessária uma breve exposição acerca dos resultados da implementação de tais políticas em alguns países e estados brasileiros, visto que

tal exposição pode subsidiar a análise dos resultados encontrados para o programa Escola de Valor. Ademais, não constitui nosso objetivo resenhar de maneira exaustiva a literatura que trabalha com tais questões, mas relatar alguns resultados encontrados em distintos países e estados.

As políticas de remuneração por resultados na educação datam do século XIX. Raple (1994) aponta que perdurou na Inglaterra, entre 1862 e 1897, uma política de pagamento por resultados que, para fins de avaliação, utilizava as informações das condições físicas da escola, um teste anual aplicado aos alunos e a frequência escolar. Apesar das reformulações sofridas pelo programa ao longo dos 35 anos de sua implementação, seus resultados não foram considerados positivos. Entre as principais críticas ao programa, Raple (1994) denota que estas se relacionaram aos testes e *scores* obtidos, que, por serem demasiado mecânicos, induziam práticas pedagógicas mecanicistas.

Lavy (2002), por seu turno, objetivou verificar o impacto de um programa que introduziu incentivos monetários para escolas e professores em função da melhoria nas notas de seus alunos, em Israel. Foram selecionadas 62 escolas secundárias para o programa em 1995, o qual ofereceu incentivos às escolas na forma de prêmios por desempenho, parte dos quais foram distribuídos para professores e funcionários da escola como pagamento por mérito e o restante usado para o bem-estar (melhorias das condições gerais de trabalho) da escola. A quantia total de recursos concedida foi determinada antecipadamente (cerca de R\$ 1,4 milhão) e foi distribuída entre o terço superior de um *ranking* de desempenho multidimensional.

Por meio da combinação de regressão descontínua, diferenças em diferenças e métodos de pareamento, excluindo-se os efeitos encontrados para a proporção de alunos que recebem certificados de conclusão, o autor encontra resultados positivos e estatisticamente significativos nos dois anos em que o programa foi avaliado, 1996 e 1997, para as escolas religiosas, e apenas em 1997 para as demais escolas. Os resultados também sugerem que, para todos os tipos de escola, os incentivos lograram ganhos significativos, especialmente entre os alunos de origem desfavorecida, os quais qualificaram-se o suficiente para obterem certificados de conclusão.

Resultados similares foram encontrados por Glewwe, Ilias e Kremer (2003), ao estudarem os efeitos de um programa de remuneração por resultados para professores na zona rural do Quênia, no qual os professores de 4ª a 8ª séries das escolas vencedoras recebiam o bônus. Os autores mostram que o programa obteve bons resultados de curto prazo, todavia resultados pífios de longo prazo,

sugerindo que, assim como o assinalado por Raple (1994), os professores participantes do programa podem ter praticado *gaming*.⁹

Muralidharan e Sundararaman (2011) apresentam outro exemplo de sucesso, desta vez no estado de Andra Pradexe, na Índia, na década de 2000. No programa, quinhentas escolas foram aleatoriamente selecionadas, as quais, sob forte esquema antifraude, receberiam o bônus: em média, 3% do salário anual médio do professor, linearmente ao aumento das notas dos alunos em testes aplicados em mais de uma data. Nesse caso, para uma escola receber o bônus, as notas dos alunos deveriam aumentar ao menos 5%. Ao final de dois anos do programa, os alunos das escolas selecionadas apresentaram desempenho significativamente melhor, quando comparados aos do grupo de controle, respectivamente, com aumento de 0,27 e 0,17 desvio-padrão nos testes de matemática e linguagem.

Apesar do primeiro registro de políticas de remuneração por desempenho de professores ser atribuído à Inglaterra, foi nos Estados Unidos que essas políticas adquiriram grande difusão. Logo, faz-se necessário elencar alguns resultados encontrados em estados estadunidenses. Por exemplo, Fryer Junior (2011) não encontrou evidências de que os incentivos pagos aos professores do estado de Nova York tenham aumentado o rendimento de seus alunos no período 2006-2010, tampouco indícios de que os incentivos tenham alterado substancialmente o comportamento de alunos e professores.

Cowan e Goldhaber (2018), por seu turno, estudaram os efeitos de uma política do estado de Washington que concede bônus financeiro a professores certificados pelo Conselho Nacional que atuem em escolas localizadas em áreas menos abastadas do estado. Por meio da utilização do método de regressão descontínua, os autores estimam que, nos primeiros seis anos do programa, a proporção de professores certificados em escolas periféricas aumentou de 0,7 a 1,6 ponto percentual (p.p.) por ano. Todavia, apesar da melhoria na contratação de profissionais qualificados, efeitos positivos do programa no desempenho dos alunos não foram verificados. Em relação ao estado do Tennessee, Swain, Rodriguez e Springer (2019) estimaram o efeito do pagamento de bônus para fixação de professores eficientes¹⁰ em escolas periféricas sobre o desempenho dos alunos. Seus resultados indicam que as escolas que ofereciam os bônus para fixação dos

9. No contexto de políticas educacionais, *gaming* pode ser entendido como comportamentos estratégicos intencionais, voltados apenas a aumentar a nota média das provas, sem necessariamente melhorar o aprendizado dos alunos de modo consistente. Os autores verificaram apenas mudanças nas atitudes dos professores em relação à alocação de tempo para treinamento para a prova, mudanças relativas à maior presença em sala de aula ou à utilização de novos métodos pedagógicos não foram observadas, o que pode explicar o aumento das notas por meios que burlem o sistema. Para mais informações, ver Glewwe, Ilias e Kremer (2003).

10. Definidos mediante a análise de um conjunto de observações, entre as quais frequência escolar e resultados de turmas em testes padronizados.

melhores professores lograram maiores ganhos em testes de anos subsequentes, especialmente nos exames estaduais de leitura.

Da análise desses casos, verifica-se que as evidências, para diversos contextos e nacionalidades, acerca da eficácia de programas de incentivo mediante pagamento por desempenho não são conclusivas. Com relação à literatura brasileira, poucos são os estudos que avaliaram a eficácia de tais políticas no contexto nacional. Todavia, alguns achados empíricos são verificados.

Oshiro, Scorzafave e Dorigan (2015) investigaram o efeito do programa de bonificação de professores e funcionários da Secretaria de Educação do Estado de São Paulo, implantado em 2008, sobre as notas dos alunos do 5º e 9º anos em testes padronizados. O programa estipula metas, estabelecidas com base no Índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp), para cada série das escolas, remunerando todos os profissionais (professores do ciclo avaliado e demais funcionários). O Idesp, por sua vez, é composto por dois outros índices: o Índice de Desempenho (ID), calculado com base nas notas do Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (Saresp), exame de avaliação externa paulista; e o Índice de Fluxo (IF).

A partir do Idesp, por meio do Índice de Cumprimento (IC), calcula-se a parcela de cumprimento da meta estabelecida pela escola. Adicionais de qualidade, presentes no Índice de Qualidade (IQ), que refletem a posição da escola em relação à média das demais e à sua meta de longo prazo, também são levados em consideração para efeito de pagamento¹¹ do bônus. Os autores apontam que a bonificação, de até 2,9 salários, varia conforme a parcela da meta cumprida e que apenas professores do ciclo que cumpriram a meta e funcionários com nível de absenteísmo abaixo de um terço a recebem. Trabalhando com técnicas de pareamento e de diferenças em diferenças, os resultados denotam indícios de impactos positivos, limitado ao ano de 2009, apenas para o 5º ano, mas negativos, por vezes estatisticamente significativos, para o 9º ano.

Lépine (2016) revisita os resultados encontrados por Oshiro, Scorzafave e Dorigan (2015), ao analisar o programa da Secretaria de Educação do Estado de São Paulo usando arcabouços de diferenças em diferenças e tripla diferença com efeitos fixos em um maior recorte temporal. As estimativas obtidas mostram que, em geral, o programa teve efeitos positivos, embora os ganhos de desempenho variem entre as séries e temáticas analisadas. Alunos do 5º ano apresentam ganhos positivos e consistentes em matemática e linguagem, enquanto os ganhos para os alunos do 9º ano são modestos. A autora denota que uma possível explicação para os resultados relaciona-se ao fato de que os alunos do 9º ano interagem com um

11. Oshiro, Scorzafave e Dorigan (2015) assinalam que o sistema paulista considera a soma dos indicadores (IQ + IC), limitado entre 0% e 120%.

número maior de professores, em comparação aos do 5º ano, o que pode dificultar a coordenação entre os professores e, conseqüentemente, o aprendizado dos alunos.

Furtado e Soares (2018), por sua vez, estimaram o impacto da bonificação educacional nas proficiências das disciplinas de língua portuguesa e matemática, entre 2008 e 2012, no estado de Pernambuco. Diferentemente de São Paulo, o bônus é pago para todos os funcionários das escolas estaduais que atingem o mínimo de 50% da meta preestabelecida, sendo que seu valor varia de acordo com o percentual de cumprimento das metas. Em um dos anos analisados pelos autores, 2012, 56% das escolas estaduais atingiram as metas, o que correspondeu a um bônus médio de R\$ 2.976,24 por funcionário das escolas.

Os resultados encontrados, apesar de positivos, pouco impactaram o desempenho dos alunos. Os testes de significância para o 5º ano validaram apenas os intervalos com distâncias maiores de tempo – para o Sistema de Avaliação Educacional de Pernambuco (Saepe), entre 2008 e 2012, e para Prova Brasil, entre 2007 e 2011. Os resultados para o 9º ano foram validados em todos os intervalos analisados, entretanto, os ganhos não foram sendo incrementados no decorrer da série histórica. Os autores assinalam que, provavelmente, a melhoria inicial induzida pelo bônus não se mantém constante ao longo do tempo. Conseqüentemente, caberia aos gestores da bonificação avaliar se o retorno de no máximo 5 p.p. justificaria a continuação dessa política (Furtado e Soares, 2018).

3 PROGRAMA ESCOLA DE VALOR

As iniciativas de bonificação a profissionais da educação no estado da Paraíba têm seu início em 2011, com a criação do Prêmio Educação Exemplar, instituído pela Medida Provisória nº 181, de 22 de agosto de 2011, e transformado na Lei nº 9.456, de 6 de outubro de 2011.

O Prêmio Educação Exemplar consistiu em uma iniciativa do governo do estado da Paraíba, por intermédio da Secretaria de Estado da Educação da Paraíba (SEE-PB), com o intuito de fomentar, selecionar e valorizar, por meio de premiação, as melhores práticas pedagógicas resultado de ações integradas e métodos docentes planejados que, tendo sido desenvolvidas por profissionais da educação lotados nas escolas públicas estaduais de educação básica, foram exitosas no enfrentamento dos desafios no processo de ensino e aprendizagem (Paraíba, 2011). Com esse fim, o programa se desdobrava em duas categorias: Gesto Exemplar e Professor Exemplar, com o intuito de premiar cem escolas e mil docentes, respectivamente, com o 14º salário. No entanto, a vigência do prêmio ficou restrita apenas ao ano de 2011.

O Prêmio Educação Exemplar constituiu-se uma ação anterior à criação do Sistema Estadual de Avaliação da Educação da Paraíba (Avaliando IDEPB), em 2012, e foi regulamentado em 2015. O Índice de Desenvolvimento da Educação

da Paraíba (IDEPB), assim como o Índice de Desenvolvimento da Educação Básica (Ideb), combina indicadores de desempenho e fluxo escolar, a fim de oferecer um diagnóstico da rede estadual de ensino da Paraíba. Além dos testes aplicados anualmente de português e matemática, compõem o índice os questionários socioeconômicos dos estudantes, sendo suas metas estabelecidas pelo Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora (CAEd/UFJF). Nesse sentido, cada unidade escolar deve, no seu projeto de intervenção pedagógica (PIP), planejar ações direcionadas para o alcance das metas estipuladas para o índice (Rodrigues e Silva, 2020).

Após a criação do Avaliando IDEPB, foram instituídas duas outras premiações: o Mestres da Educação e o Escola de Valor, que dão sequência às políticas de bonificação aos docentes e demais profissionais da educação na Paraíba. O Prêmio Escola de Valor constitui uma iniciativa do governo do estado da Paraíba instituída nos termos da Lei nº 9.879, de 13 de setembro de 2012. O programa se diferencia da iniciativa anterior do Prêmio Gesto Exemplar, por ter como objetivo avaliar as escolas públicas estaduais da educação básica nas diversas dimensões da gestão escolar, sendo elas: i) gestão pedagógica; ii) gestão participativa; iii) gestão de pessoas e liderança; e iv) gestão de infraestrutura: serviços e recursos (Cavalcante, 2019; Paraíba, 2012; 2013; 2014; 2015; 2016). Tais alterações no programa no decorrer de sua vigência podem ser melhor compreendidas por meio do exame do quadro 1.

QUADRO 1
Breve histórico do programa Escola de Valor e suas alterações (2011-2016)

Questão	2011	2012	2013	2014	2015	2016	
Objetivos	Destacar as experiências pedagógicas e administrativas do cotidiano curricular que contemplem os eixos norteadores das ações da SEE-PB: educação e cultura, educação e cidade, educação e campo, educação e trabalho, educação e comunicação, educação e tecnologia, educação e cidadania.	Valorizar as escolas públicas estaduais que se destacam no quesito pela competência da gestão administrativa e pedagógica, por iniciativas de experiências inovadoras e bem-sucedidas na melhoria contínua da escola. Reconhecer e dar visibilidade ao esforço empreendido pelos gestores e demais profissionais da educação etc.	Avallar as escolas públicas estaduais de educação básica nas diversas dimensões da gestão escolar, a saber: gestão pedagógica; gestão participativa; gestão de pessoas e liderança; e gestão de infraestrutura; serviços e recursos. Incentivar as escolas públicas estaduais de educação básica a desenvolverem a cultura da autoavaliação nas diversas dimensões da gestão escolar.	Avallar as escolas públicas estaduais de educação básica nas diversas dimensões da gestão escolar, a saber: gestão pedagógica; gestão participativa; gestão de pessoas e liderança; e gestão de infraestrutura; serviços e recursos. Incentivar as escolas públicas estaduais de educação básica a desenvolverem a cultura da autoavaliação nas diversas dimensões da gestão escolar.	Avallar as escolas públicas estaduais de educação básica nas diversas dimensões da gestão escolar, a saber: gestão pedagógica; gestão participativa; gestão de pessoas e liderança; e gestão de infraestrutura; serviços e recursos. Incentivar as escolas públicas estaduais de educação básica a desenvolverem a cultura da autoavaliação nas diversas dimensões da gestão escolar.	Avallar as escolas públicas estaduais de educação básica nas diversas dimensões da gestão escolar, a saber: gestão pedagógica; gestão participativa; gestão de pessoas e liderança; e gestão de infraestrutura; serviços e recursos. Incentivar as escolas públicas estaduais de educação básica a desenvolverem a cultura da autoavaliação nas diversas dimensões da gestão escolar.	Avallar as escolas públicas estaduais de educação básica nas diversas dimensões da gestão escolar, a saber: gestão pedagógica; gestão participativa; gestão de pessoas e liderança; e gestão de infraestrutura; serviços e recursos. Incentivar as escolas públicas estaduais de educação básica a desenvolverem a cultura da autoavaliação nas diversas dimensões da gestão escolar.
Critérios de participação	Envio do relato de experiência de gestão e cópias de documentos comprobatórios especificados no edital, além de posteriores vistórias às escolas.	Não houve necessidade de realização de inscrição, apenas o envio dos documentos comprobatórios exigidos.	Realização de inscrição por meio de preenchimento de formulário de inscrição, envio dos documentos comprobatórios e da avaliação com justificativa.	Realização de inscrição por meio de preenchimento de formulário de inscrição, envio dos documentos comprobatórios e da avaliação com justificativa.	Realização de inscrição por meio de preenchimento de formulário de inscrição, envio dos documentos comprobatórios e da avaliação com justificativa.	Realização de inscrição por meio de preenchimento de formulário de inscrição, envio dos documentos comprobatórios e da avaliação com justificativa.	Realização de inscrição por meio de preenchimento de formulário de inscrição, envio dos documentos comprobatórios e da avaliação com justificativa.
Pontuação	Não houve definição de pontuação específica para cada um dos critérios de avaliação neste ano do programa.	PPP: 0 a 2 pontos. Documentação comprobatória: 0 a 8 pontos.	PPP e relatório de execução do PIP: 0 a 1 ponto. Documentação comprobatória: 0 a 8 pontos. Autoavaliação com justificativa: não pontua, porém acarreta desclassificação, em caso de não apresentação.	PPP e relatório de execução do PIP: 0 a 1 ponto. Documentação comprobatória: 0 a 8 pontos. Autoavaliação com justificativa: 0 a 1 ponto e consiste em critério de desclassificação, em caso de não apresentação.	PPP e relatório de execução do PIP: 0 a 1 ponto. Documentação comprobatória: 0 a 8 pontos. Autoavaliação com justificativa: 0 a 1 ponto e consiste em critério de desclassificação, em caso de não apresentação.	PPP: a 1 ponto. Relatório de execução do PIP: 0 a 2 pontos. Documentação comprobatória: 0 a 6 pontos. Autoavaliação com justificativa: 0 a 1 ponto e consiste em critério de desclassificação, em caso de não apresentação.	
Critério de seleção ou premiação	Seleção, pela comissão julgadora, de cem escolas na categoria Gesto Exemplar.	Cumprimento de, no mínimo, 60% dos critérios estabelecidos no processo seletivo e apresentação da documentação comprobatória exigida.	Cumprimento de, no mínimo, 60% dos critérios estabelecidos no processo seletivo e apresentação da documentação comprobatória exigida.	Cumprimento de, no mínimo, 70% dos critérios estabelecidos, apresentação da documentação comprobatória e da autoavaliação com justificativa.	Cumprimento de, no mínimo, 70% dos critérios estabelecidos, apresentação da documentação comprobatória e da autoavaliação com justificativa.	Cumprimento de, no mínimo, 70% dos critérios estabelecidos, apresentação da documentação comprobatória e da autoavaliação com justificativa.	

Fonte: Paraíba (2011; 2012; 2013; 2014; 2015; 2016).

Elaboração dos autores.

Obs.: PPP – projeto político-pedagógico.

Inicialmente, o prêmio tinha como objetivos principais valorizar as escolas públicas destaque no quesito competência da gestão administrativa e pedagógica, por meio de suas experiências inovadoras e de sucesso na promoção da melhoria contínua escolar, e, assim, reconhecê-las e recompensá-las dando visibilidade ao esforço realizado pelos profissionais responsáveis pelo processo de aprendizagem e incentivo dos alunos pela busca do conhecimento (Paraíba, 2012). Entre os anos de vigência, de 2013 a 2016, passou a consistir também como objetivo do programa incentivar as escolas públicas estaduais de educação básica a desenvolverem a cultura da autoavaliação nas diversas dimensões da gestão escolar.

Desde que foi instituído, em todas as edições do programa, geralmente no início do mês de julho, a SEE-PB publica um edital que dispõe acerca das regras e condições estabelecidas para a inscrição das escolas aptas a concorrer ao prêmio, assim como sobre os critérios de seleção e a forma de concessão da premiação.

Exceto na primeira edição do Prêmio Escola de Valor, em 2012, quando não foi necessária a realização da inscrição pela instituição de ensino para estarem aptas a concorrerem ao prêmio, as escolas precisam preencher os requisitos, que consistem na manifestação de interesse em participar da seleção por meio de envio: do formulário de inscrição, do PIP, do relatório de execução do PIP no ano, dos demais documentos comprobatórios e da autoavaliação com justificativa (Paraíba, 2016). A autoavaliação com justificativa só começou a ser utilizada como critério de avaliação, e passou a ser a ela atribuído um peso na pontuação, a partir de 2014, sendo antes utilizada, no caso de não envio, apenas como critério de desclassificação da escola (Paraíba, 2013; 2014; 2016).

Após homologadas as inscrições, geralmente no mês de outubro, o dossiê enviado pelas escolas é analisado por uma comissão avaliadora do estado, constituída por profissionais da educação com formação de nível superior e/ou comprovação de conhecimento técnico acerca dos documentos a serem avaliados, sendo eles indicados por instituições de ensino superior e pela SEE-PB, que tomam como base um conjunto de critérios de seleção, atribuindo pontuações que variam de zero à nota máxima correspondente a cada critério.

Durante o período analisado de vigência do Prêmio Escola de Valor, a saber de 2011 (Prêmio Educação Exemplar) a 2017, houve mudança significativa nos critérios de avaliação utilizados para selecionar as escolas premiadas, assim como na importância ou no peso atribuído aos critérios na pontuação geral. Inicialmente, um peso maior na pontuação era atribuído ao PPP, aos resultados do Ideb, ao alcance das metas do IDEPB e aos indicadores educacionais do Censo Escolar; porém, a partir de 2016, passou-se a dar maior importância na avaliação ao PIP, ao relatório de execução do PIP, à autoavaliação com justificativa, além de outros indicadores de gestão escolar.

No edital nº 005/2016-GS, o PIP, o relatório de execução, os documentos comprobatórios e a autoavaliação com justificativa apresentados pela escola à SEE-PB durante o período de inscrição poderão totalizar pontuação de 0 a 10, considerando-se que a apresentação do PIP vale de 0 a 1 ponto, o relatório de execução do PIP garante pontuação de 0 a 2 pontos, a documentação comprobatória garante pontuação que varia de 0 a 6 pontos e a autoavaliação com justificativa, de 0 a 1 ponto (Paraíba, 2016).

Em relação ao PIP, ele deve conter necessariamente ações pedagógicas alinhadas com o considerado nas avaliações de português e matemática da Prova Brasil, visando à obtenção de melhores desempenhos em tais avaliações e ao atingimento das metas previstas para cada escola no IDEPB para o referido ano. Além disso, o PIP precisa apresentar estratégias para redução da evasão escolar, melhoria do desempenho escolar dos alunos, minimização da violência escolar, além de incentivar discussões sobre direitos humanos, diversidade, sustentabilidade, inclusão etc.

Quanto ao relatório de execução do PIP, este deve estar alinhado ao Plano de Metas do IDEPB do ano e estar devidamente respaldado com provas documentais que comprovem os resultados obtidos pela escola naquele período, enfatizando, principalmente, o crescimento dos resultados no IDEPB em relação ao ano imediatamente anterior e a participação mínima dos estudantes na avaliação do IDEPB no ano. De acordo com o edital nº 005/2016-GS, ainda é avaliada a consistência das práticas pedagógicas em relação aos objetivos propostos e os resultados de rendimento escolar dos alunos.

Inicialmente, na edição de 2013 do programa de incentivo, eram selecionadas as escolas que atendiam ao mínimo de 60% dos critérios estabelecidos no processo seletivo e que apresentavam a documentação comprobatória exigida. A partir de 2014, esse percentual subiu para o cumprimento de 70% dos critérios estabelecidos. Segundo o edital nº 005/2016-GS, atualmente, são selecionadas e, conseqüentemente, premiadas com o Prêmio Escola de Valor as escolas públicas da rede estadual que atendam a, no mínimo, 70% dos critérios, isto é, atinjam pontuação maior ou igual a 7,0, além da apresentação do PIP, do relatório de execução do PIP, dos documentos comprobatórios que atendam aos critérios antes estabelecidos no edital e da autoavaliação com justificativa.

Dessa forma, o recebimento do Prêmio Escola de Valor pelas escolas inscritas não está vinculado apenas ao desempenho da escola nas avaliações externas, como acontece em programas de bonificação implementados em outros estados, como Pernambuco e São Paulo. Ademais, conforme afirmam Lazear (2001), Neal e Schanzenbach (2010) e Reback (2008), pelo fato de não se tratar de um programa de responsabilização e bonificação de professores com o objetivo de melhoria do desempenho dos alunos em avaliações externas, reduz-se o viés

de mudança de distribuição do desempenho dos estudantes, isto é, a tendência de os professores concentrarem seus esforços nos melhores alunos, visto que são os com maior potencial de atingir a meta desejada.

No que diz respeito à premiação, o Escola de Valor gratifica com o equivalente a uma remuneração mensal a qual recebe o profissional, caracterizando um 14º salário para todos os profissionais de educação em exercício nas escolas públicas estaduais de educação básica selecionadas no processo seletivo.

A indisponibilidade de informações nos meios de publicidade oficiais, a falta de divulgação das pontuações e de transparência da SEE-PB, de uma maneira geral, dificultaram o acesso aos resultados dos processos seletivos do Escola de Valor, trazendo limitações a este estudo. Como dito anteriormente, o programa existe desde 2011 (com outra denominação) e, desde sua criação, os critérios de seleção das escolas premiadas mudam constantemente, ano a ano. Assim, não é possível tratar de forma homogênea as premiações em anos subsequentes ao primeiro ano de vigência do programa, uma vez que, com as particularidades de cada seleção, é como se o Escola de Valor englobasse diferentes programas dentro de um. Por essa razão, este artigo analisa o impacto nas escolas premiadas em 2016. Além disso, esse ano em particular conta com informações mais completas para essa edição do prêmio.

Outra questão que impacta a estratégia de mensuração é que, com as mudanças nos critérios de premiação ano a ano, a relação de escolas premiadas se altera a cada premiação, fazendo com que várias escolas transitem entre os grupos das premiadas e não premiadas. Dentro de um grupo de controle, é necessário diferenciar as escolas que porventura já foram premiadas, sob o risco de tratar igualmente como grupo de controle escolas já tratadas em premiações anteriores. Como forma de incorporar essas diferenças, a estratégia aqui utilizada leva em conta o número de vezes que a escola foi premiada antes de 2016. Técnica semelhante foi empregada por Rocha e Soares (2010), os quais consideraram o tempo de exposição de municípios brasileiros ao Programa Saúde da Família (PSF). A seção 4 explica com mais detalhes a técnica utilizada.

4 METODOLOGIA E BASE DE DADOS UTILIZADAS

4.1 Método de análise

Diferentemente do que ocorre em um experimento, no qual os grupos de tratados e controle são selecionados por um processo aleatório, garantindo que ambos sejam semelhantes, nos quase-experimentos, tais grupos surgem de uma intervenção governamental específica (Wooldridge, 2013).

Como na maioria dos programas de bonificação por desempenho escolar, no Prêmio Escola de Valor, as escolas premiadas não são escolhidas aleatoriamente, mas de acordo com critérios de elegibilidade, de modo que existe uma potencial correlação entre o recebimento do prêmio e as características não observáveis das escolas, o que implica a existência de um viés de autosseleção. Ademais, no caso do Escola de Valor, já na concorrência ao prêmio, tem-se a manifestação de interesse da escola em participar do processo seletivo por meio de sua inscrição, o que torna essas escolas positivamente selecionadas. Dessa forma, é de se esperar que as escolas que possuem diretores e gestores mais motivados concorram ao prêmio e tenham, conseqüentemente, uma probabilidade maior de serem contempladas com a bonificação, o que, por sua vez, também condiciona o rendimento escolar dos alunos. O objetivo do trabalho consiste em identificar se o recebimento do prêmio impacta de forma positiva o rendimento dos alunos no período seguinte e, em decorrência da forma de seleção ao prêmio, se uma diferença de média simples entre as escolas premiadas e não premiadas não forneceria o impacto real do programa.

O problema fundamental da avaliação de programas sociais consiste na construção do contrafactual do grupo tratado pela política, isto é, verificar as mesmas unidades escolares quando não sofreram intervenção do programa, o que não é possível, dado que são situações mutuamente excludentes. Portanto, o desafio crucial da avaliação é encontrar um grupo de escolas representativo da situação de não tratamento, isto é, um grupo que constitua um bom contrafactual do grupo tratado (Foguel, 2012b). Nesse caso, para minimizar os problemas de seleção, a estratégia empírica aqui utilizada para testar o impacto do Escola de Valor consiste na combinação de dois métodos de avaliação de impacto: o método de pareamento por PSM e o método DID. O PSM visa selecionar, entre as escolas não premiadas, aquelas que teriam probabilidade de ser, dadas algumas características observáveis. Selecionadas essas escolas, o método DID busca mensurar a diferença de rendimento escolar entre escolas premiadas e não premiadas atribuível ao recebimento da bonificação.

Os métodos de pareamento se baseiam em duas hipóteses fundamentais: a seleção nos observáveis, ou ignorabilidade, e a existência de suporte comum entre os grupos de tratamento e controle. A hipótese de seleção em observáveis consiste em considerar um vetor X de características observáveis de ambos os grupos, a fim de controlar os fatores que influenciam o resultado potencial e a participação no programa, visando garantir que o resultado obtido pelo grupo de controle seja um bom predictor do que aconteceria ao grupo de tratado na ausência de tratamento. Assim, controlando por X , o resultado potencial torna-se independente de T (variável de tratamento), isto é, independência condicional em X entre T , Y_1 (indivíduo i na situação de tratado) e Y_0 (indivíduo i na situação de não tratado), como visto na equação (1).

$$(Y_1, Y_0) \perp T | X, \quad (1)$$

em que T é ignorável se $\Pr(T = 1 | Y_1, Y_0, X) = \Pr(T = 1 | X) \equiv p(X)$, sendo $p(X)$ o *propensity score*.

Por sua vez, a hipótese de suporte comum visa garantir que a região do vetor X que comporta as características observáveis de tratados também englobe as características dos controles, de maneira que não se possa identificar a que grupo o indivíduo pertence, comparando apenas as características de X , como se pode representar de maneira formal pela equação (2).

$$0 < \Pr(T_i = 1 | X_i) < 1. \quad (2)$$

Admitindo tais hipóteses, o estimador de pareamento encontra, para cada indivíduo do grupo de tratado, indivíduos mais próximos no grupo de controle em termos das características X . No entanto, deve-se atentar para a sensibilidade do estimador à escolha do vetor X , pois, se X for composto por muitas variáveis, haverá uma dificuldade em atender à hipótese de suporte comum; se, por sua vez, ele contiver poucas variáveis, a hipótese de ignorabilidade tenderá a ser rejeitada (Pinto, 2012).

Uma forma de contornar tal dilema é parear os grupos utilizando uma função de X que resume toda a informação inclusa nesse vetor por meio da técnica de PSM. A função ou o escore de propensão ($p(X_i)$) é dada(o) pela probabilidade condicional de receber o tratamento, dado o vetor X no período pré-tratamento, como representada(o) na equação (3).

$$p(X_i) = \Pr(T_i = 1 | X_i). \quad (3)$$

Se respeitadas as hipóteses de ignorabilidade e suporte comum e se conhecido o escore de propensão, então se pode admitir que

$$Y_i(0) \perp T_i | X_i \Rightarrow Y_i(0) \perp T_i | p(X_i). \quad (4)$$

Sendo assim, os indivíduos estarão pareados com base na similaridade quanto à propensão de serem submetidos ao tratamento. Neste estudo, buscou-se construir um contrafactual observável do grupo de escolas premiadas por meio do pareamento em características observadas, encontrando um suporte comum, tornando os grupos de escolas premiadas e não premiadas os mais homogêneos possíveis, de modo a minimizar o viés de seleção nas características observáveis.

Considerando-se os critérios de elegibilidade do programa, os quais influenciam na seleção para receber a bonificação e no resultado potencial, as variáveis observáveis do período pré-tratamento para construção do contrafactual são apresentadas no quadro 2. A justificativa da utilização de tais variáveis para a seleção do grupo de controle se encontra na relação entre elas e a probabilidade de participação no programa de bonificação. Espera-se que as escolas que possuem diretores mais

experientes, escolhidos de forma legítima e democrática, professores mais qualificados, maior corpo de funcionários, dotadas de melhor infraestrutura escolar e nível socioeconômico, assim como as que têm menor média de alunos por turma, maior carga horária e duração do ano letivo, tenham uma gestão administrativa e pedagógica mais eficiente, condicionando um melhor desempenho escolar de seus alunos e, conseqüentemente, uma probabilidade maior de serem premiadas.

Além de variáveis ligadas diretamente a alguns dos critérios mostrados na seção anterior, foram incluídas outras características que podem ser relevantes à concessão do prêmio às escolas. São elas: se a escola se encontra ou não em uma zona urbana; e um índice que mede a situação socioeconômica da escola. Conforme já discutimos, diante da falta de informação acerca da real pontuação que cada escola obteve nas premiações, bem como acerca dos pesos dados aos critérios supostamente utilizados, é razoável supor que a premiação esteja ligada não somente às variáveis estritamente escolares. É possível que escolas localizadas em zonas rurais tenham precedência no recebimento da bonificação, tendo em vista se tratar de escolas de difícil acesso ou carentes de infraestrutura. Escolas que apresentem alunos em situação de maior vulnerabilidade social podem ter maiores chances de serem premiadas, ainda que não atendam inteiramente aos critérios estritamente educacionais. A preocupação do governo em compensar o baixo nível socioeconômico e/ou a falta de estrutura de escolas mais afastadas dos grandes centros pode ser um fator importante para definir as escolas que receberão as bonificações.

QUADRO 2
Características observáveis no período pré-tratamento

Variáveis	Descrição	Fonte
infraestrutura_escolar ¹	Índice de Infraestrutura Escolar	Censo Escolar
inse	Índice de Nível Socioeconômico	Censo Escolar
urbana	1, caso a escola esteja localizada na zona urbana, 0 caso seja na zona rural	Censo Escolar
carga_horaria	Quantidade de horas-aula na escola	Censo Escolar
media_alunos	Média de alunos por turma na escola	Censo Escolar
duracao_ano	Quantidade de dias de duração do ano letivo	Censo Escolar
num_funcionários	Quantidade de funcionários por escola	Censo Escolar
escolha_diretor	1, caso diretor tenha sido escolhido por concurso público, processo seletivo e/ou eleição, 0 caso contrário	Saeb
experiencia_diretor	Tempo de experiência do diretor da escola na função	Saeb
docentes_superior	Percentual de docentes com ensino superior	Censo Escolar

Fonte: Inep. Disponível em: <https://www.gov.br/inep/pt-br/ acesso-a-informacao/dados-abertos/microdados>. Acessos em: 10 nov. 2018 e 20 mar. 2019.

Elaboração dos autores.

Nota: ¹ O Índice de Infraestrutura Escolar foi criado a partir do método de componentes principais como método corretivo para o problema de colinearidade apresentado quando do pareamento, visto que tal grupo de variáveis *dummy* utilizado para a criação do índice (escola com fornecimento de merenda, água, energia, presença de esgoto) descrevem um mesmo componente: a estrutura da escola.

Para estimar o escore de propensão $p(X)$, recomenda-se utilizar modelos de probabilidade, por exemplo, o modelo *logit* e o modelo *probit*. Contudo, para definir os indivíduos mais próximos com base no pareamento com escore de propensão, é necessário estabelecer previamente o algoritmo de pareamento. Existem diversos algoritmos utilizados na literatura empírica, sendo os principais: vizinho mais próximo, caliper, kernel, polinomial etc. (Pinto, 2012).

Por meio da estimação de um modelo *logit*, o contrafactual para cada escola do grupo de tratamento foi construído usando as escolas do grupo de não tratados que estão em uma vizinhança em torno do escore de propensão da escola tratada. O grupo de controle será formado pelas escolas estaduais e municipais do Estado. Com isso, os grupos de tratados e controles foram identificados por meio de uma variável *dummy*, sendo o primeiro escolas estaduais da Paraíba premiadas no programa Escola de Valor ($D = 1$), e o segundo, escolas municipais e estaduais que não participaram do programa ($D = 0$).

Posteriormente, foram utilizados testes de diferença de média para verificar se, após o PSM, os grupos de tratamento e controle apresentaram proximidade em termos de média das variáveis observáveis, dado o nível de significância estatística admitido. Em seguida, foi testada, por meio da análise gráfica, a existência de suporte comum pela sobreposição entre as distribuições dos escores de propensão em cada grupo. Dessa forma, constatada a semelhança na média e havendo sobreposição das características observáveis, tem-se a validade das hipóteses de ignorabilidade e suporte comum, minimizando, assim, o viés de seleção nas observáveis.

No tocante ao método DID, uma das principais vantagens do método é a capacidade de lidar com o viés de seleção ligado às características não observáveis dos indivíduos, invariantes no tempo. A principal hipótese do método DID requer que os grupos de tratamento e controle apresentem trajetórias paralelas da variável dependente nos períodos anteriores ao tratamento. A intuição por trás dessa suposição é simples: se as trajetórias são parecidas no período anterior ao programa, parece razoável concluir que a evolução do grupo de controle após o programa represente o que ocorreria com o grupo de tratados na situação de não tratamento. Desvios na trajetória da variável dependente entre os grupos pós-programa confirmam os efeitos causais da intervenção (Foguel, 2012a).

A forma convencional de apresentar o estimador de DID é por meio do cálculo de uma dupla diferença de médias da variável dependente. Seja $T = 0,1$, em que 1 = tratados e 0 = não tratados, e $t = 0,1$ os períodos antes e depois da implementação do programa, respectivamente, o estimador de DID pode ser representado segundo a equação (5).

$$\beta_{DD} = E[T = 1, t = 1] - E[T = 1, t = 0] - \{E[Y|T = 0, t = 1] - E[Y|T = 0, t = 0]\}. \quad (5)$$

Inicialmente, calcula-se a diferença de média entre os grupos em cada ponto no tempo e , em seguida, é calculada a diferença entre essas duas diferenças no tempo, ou faz-se a diferença de média entre os tempos 0 e 1 para cada grupo e , em seguida, faz-se a diferença dessas duas diferenças entre os grupos. Esse cálculo pode ser resumido no modelo empírico a seguir.

$$Y_{it} = X_{it}\alpha + \lambda T_{it} + \rho t_{it} + \beta T_{it}t_{it} + \epsilon_{it}, \quad (6)$$

em que t é uma variável binária que assume valor 0 para o período pré-tratamento e 1 para o período posterior ao tratamento; T é uma variável binária que assume valor 1, se o indivíduo é tratado, e 0, caso contrário; e o vetor X é composto por uma série de variáveis de controle. O termo de erro é dado por ϵ_{it} . Sob a hipótese de que $E[\epsilon|X, T, t] = 0$, o efeito causal do programa (diferença das diferenças) condicional a X é estimado pelo coeficiente β da interação tratado x tempo $T_{it}t_{it}$.

Conforme argumentado na seção 2, a mudança nos critérios de premiação ano a ano pode originar um trânsito de escolas entre os *status* de tratado e não tratado, de maneira que, embora em 2016 uma escola em particular esteja no grupo de controle, esta pode ter sido contemplada com a bonificação em uma edição anterior do Escola de Valor. Para contornar essa fonte de viés, estimamos um modelo semelhante ao da expressão (6), entretanto medimos o tratamento por meio de variáveis *dummies* que denotam o número de vezes em que as escolas já foram premiadas antes da edição 2016 do Escola de Valor, desde o seu início, em 2011.

$$y_{it,k} = \beta_{0,k} + \beta_{1,k}T + \beta_{2,k}d_{1it} + \beta_{3,k}d_{2it} + \beta_{4,k}d_{3it} + \beta_{5,k}d_{4it} + \beta_{6,k}d_{5it} + \delta_{1,k}d_{1it}T + \delta_{2,k}d_{2it}T + \delta_{3,k}d_{3it}T + \delta_{4,k}d_{4it}T + \delta_{5,k}d_{5it}T + X'_{it}\alpha_i, \quad (7)$$

em que $y_{it,k}$ são os indicadores de desempenho escolar comumente utilizados na literatura, como a proficiência média da escola em português e matemática, a taxa de aprovação escolar e a taxa de abandono escolar, referentes ao 5º e 9º anos do ensino fundamental. T é a *dummy* para o tempo, com $T = 0$ para o período anterior à premiação realizada em 2016 (2015) e $T = 1$ para o período seguinte (2017). α_i refere-se ao efeito fixo das escolas. $\delta_1, \delta_2, \delta_3, \delta_4$ e δ_5 representam o efeito médio do tratamento sobre os tratados para as interações da variável *tempo* com as *dummies* para a quantidade de vezes em que uma escola recebeu a premiação: uma vez, duas vezes, até cinco vezes. O somatório de $\delta_1, \delta_2, \delta_3, \delta_4$ e δ_5 será o efeito total do Escola de Valor nas escolas estaduais.

A combinação dos métodos de pareamento e diferenças em diferenças – proposta por Heckman, Ichimura e Todd (1998) – permite que algumas das hipóteses de cada um dos métodos possam ser substituídas por hipóteses mais fracas. Por exemplo, a hipótese de seleção nos observáveis estabelece que, condicional ao vetor de variáveis observáveis X , não pode existir nenhum fator não observável que influencie simultaneamente a decisão de participar ou não no tratamento e os

resultados potenciais. Ao combinar o pareamento com o método DID, podemos permitir que fatores não observáveis que sejam constantes ao longo do tempo influenciem simultaneamente a decisão de participar e os resultados potenciais, mesmo controlando pelo vetor de variáveis observáveis X (Pinto, 2012). A seleção do grupo de controle com base no PSM torna a hipótese de trajetórias paralelas mais plausível, uma vez que são analisados grupos de tratamento e controle semelhantes, ao menos em termos de fatores observáveis.¹²

Uma forma de se chegar ao estimador que combina PSM e DID é estimar o modelo empírico de (7), ponderando a regressão pelos escores de propensão. Hirano, Imbens e Ridder (2003) mostram que, com isso, é possível obter uma amostra balanceada de tratados e controles. O impacto do tratamento é fornecido pelo cálculo das diferenças em diferenças sobre as unidades de tratamento e controle pareadas dentro do suporte comum. Para dois períodos no tempo $t = 1, 2$, a estimativa DID para cada unidade de tratamento é calculada por

$$\hat{D}_{1,2} = (y_{i2} - y_{i1}) - W(T, X) \sum_{j \in H_M(i)} (y_{j2} - y_{j1}), \quad (8)$$

em que w é o peso que é dado à unidade j de controle, pareada à unidade i de tratamento. Um problema desse estimador é que nem sempre os pesos correspondem a 1. Uma forma de contornar isso é normalizar os pesos para 1 para as unidades tratadas e, para as unidades de controle, os pesos serão $\frac{\hat{P}(X)}{1-\hat{P}(X)}$. Assim,

$$W(PMM, X) = PMM + (1 - PMM) \frac{\hat{P}(X)}{1-\hat{P}(X)}. \quad (9)$$

A partir da estimação dos coeficientes dos tempos de exposição, será possível verificar também como o *timing* de premiação explica as diferenças de resultado entre os tratados. Além disso, sabe-se que o programa apresenta também uma natureza de incentivo, o que faz com que ele demore algum tempo para surtir efeito, fato que se tentou controlar introduzindo diferentes horizontes de premiação.

4.2 Dados e estatísticas descritivas

Foram utilizados os microdados do Saeb e do Censo Escolar referentes aos anos de 2015 e 2017 para as escolas públicas (municipais e estaduais) do estado da Paraíba, disponibilizados por meio do portal do Inep.

Os indicadores de impacto que poderiam ser escolhidos objetivando avaliar o impacto do programa Escola de Valor nas escolas premiadas poderiam ser diversos. O programa poderia impactar o aumento da eficiência da gestão administrativa e pedagógica escolar, o incentivo às escolas no desenvolvimento de práticas inovadoras de ensino e aprendizagem, o aumento do esforço dos professores, o melhor desempenho dos alunos, assim como a redução da exposição dos jovens à

12. Para mais informações sobre a combinação de pareamento com DID, ver Gertler *et al.* (2011) e Khandker, Koolwal e Samad (2010).

criminalidade, da evasão escolar etc., dados os objetivos elencados pelo programa de bonificação.

Contudo, visando utilizar indicadores confiáveis e mensuráveis, serão utilizadas a proficiência média dos alunos no Saeb nas disciplinas de língua portuguesa e matemática, a taxa de aprovação escolar e a taxa de abandono escolar referentes às séries/aos anos finais da primeira e segunda fases do ensino fundamental (5º e 9º anos). Esses indicadores foram escolhidos devido ao relacionamento direto que apresentam com o objetivo fim do programa, a saber, a melhora do desempenho escolar. Cabe salientar que, pelo fato de o bônus estar atrelado às metas do IDEPB – uma avaliação interna –, o fato de o Saeb consistir em uma avaliação externa permite comparar os resultados do estudo com escolas de outros estados e evita alguns tipos de *gaming*. A literatura sugere a não utilização de indicadores locais, como forma de evitar a manipulação de resultado por parte da escola que está sendo avaliada. Sendo assim, procuramos utilizar como variável de impacto um indicador nacional, como as notas do Saeb, bastante utilizado na literatura de economia da educação e que mede exclusivamente o rendimento escolar. Além disso, serão utilizadas como variáveis de controle as covariáveis extraídas do questionário socioeconômico do Saeb e dos microdados do Censo Escolar.

A tabela 1 reporta as estatísticas descritivas das variáveis, antes e depois do programa, para o conjunto das escolas da rede pública paraibana, independentemente do tempo de exposição ao programa e da participação no Escola de Valor.

TABELA 1
Estatísticas descritivas das variáveis utilizadas antes e depois da intervenção

Variáveis	Pré-tratamento (2015)		Pós-tratamento (2017)		Geral (2015-2017)	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
nt_media5A	189,04	16,24	193,35	20,10	191,31	18,50
nt_media9A	233,36	14,71	238,70	17,46	236,02	16,36
tx_aprovacao5A	87,15	17,67	86,88	16,67	87,02	17,21
tx_aprovacao9A	85,36	12,36	86,07	11,86	85,71	12,11
tx_abandono5A	2,99	8,78	2,67	7,43	2,84	8,18
tx_abandono9A	7,13	8,61	6,55	7,99	6,84	8,31
escola_urbana	0,44	0,50	0,47	0,50	0,45	0,50
num_funcionarios	22,12	22,74	23,50	22,64	22,77	22,70
duracao_ano	315,02	17,72	311,61	16,20	313,40	17,10
media_alunos	17,70	6,55	18,14	6,73	17,91	6,64
carga_horaria	281,83	99,69	282,15	99,77	281,98	99,72
experiencia_diretor	3,13	1,42	2,77	1,55	2,94	1,50

(Continua)

(Continuação)

Variáveis	Pré-tratamento (2015)		Pós-tratamento (2017)		Geral (2015-2017)	
	Média	Desvio-padrão	Média	Desvio-padrão	Média	Desvio-padrão
diretor_indicacao	0,69	0,46	0,76	0,43	0,73	0,44
docentes_superior5A	63,28	39,80	67,68	37,72	65,31	38,91
docentes_superior9A	72,06	36,65	74,99	35,18	73,44	35,99
socioeducativa	0,01	0,08	0,01	0,09	0,01	0,09
indigena	0,01	0,08	0,01	0,09	0,01	0,08
profissionalizante	0,03	0,02	0,02	0,12	0,02	0,15
aee	0,12	0,33	0,16	0,37	0,14	0,35
agua	0,49	0,50	0,51	0,50	0,50	0,50
esgoto	0,98	0,15	0,98	0,12	0,98	0,14
energia	0,99	0,07	0,99	0,06	0,99	0,07
merenda	1	0	0,99	0,03	0,99	0,02
internet	0,42	0,49	0,51	0,50	0,46	0,50
lab_informatica	0,40	0,49	0,40	0,49	0,40	0,49
lab_ciencia	0,05	0,21	0,05	0,23	0,05	0,22
biblioteca	0,20	0,40	0,21	0,41	0,20	0,40
acessibilidade	0,22	0,41	0,24	0,43	0,23	0,42
inse	3,11	0,75	2,33	0,74	2,73	0,84
outro_pgm	0,17	0,37	0,08	0,27	0,13	0,33

Fonte: Inep. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados>. Acessos em: 10 nov. 2018 e 20 mar. 2019.

Elaboração dos autores.

Os indicadores educacionais das escolas públicas (municipais e estaduais) apresentaram uma melhora após a implementação do Escola de Valor, com a elevação da média das notas em português e matemática, o aumento da taxa de aprovação, bem como a redução da taxa de abandono ou evasão escolar do 5º e 9º anos do ensino fundamental.

Entretanto, nota-se que houve no período um crescimento do percentual de escolas urbanas, do número de funcionários das escolas, da carga horária escolar, do percentual de docentes com nível superior, além de uma melhora significativa na infraestrutura escolar no sentido do crescimento do percentual de escolas com abastecimento de água filtrada, acesso à internet, biblioteca, equipadas para a acessibilidade e com turmas de atendimento educacional especializado, o que pode ser responsável por parte da melhora dos indicadores educacionais citados. Em contrapartida, percebe-se que houve uma redução na duração do ano letivo, na experiência média dos diretores e um aumento no percentual destes selecionados por indicação política, além de uma queda no índice socioeconômico das escolas e na participação das unidades escolares em outros programas governamentais.

5 RESULTADOS

5.1 Pareamento

A estimação do escore de propensão foi realizada por meio de um modelo *logit*. A grande disponibilidade de dados do Saeb e do Censo Escolar permitiu testar, inicialmente, um grande número de características observáveis pré-tratamento da escola. Contudo, foi observado que a medida que o número de variáveis crescia, diminuía a eficiência do pareamento, no sentido de encontrar um suporte comum entre escolas tratadas e não tratadas. Desse modo, optou-se pela utilização de um modelo com um número menor de covariadas.

Foram testados diversos algoritmos comumente utilizados para definir os indivíduos mais “próximos” dos tratados e julgou-se como mais adequado o pareamento por escore de propensão baseado em kernel. O algoritmo de kernel foi o escolhido em razão da obtenção do grupo de controle mais parecido com o grupo de tratamento em termos das características observáveis, conforme os resultados dos testes de médias e de suporte comum apresentados na sequência.

Na edição de 2016, das 558 escolas que se candidataram ao prêmio, 82 foram selecionadas. Sendo assim, a amostra obtida foi composta por 82 escolas estaduais premiadas (tratadas), em 2016, e 462 escolas estaduais e municipais do grupo de controle selecionado no pareamento. A tabela 2 apresenta os testes de balanceamento e de médias das características observáveis no período pré-tratamento para o grupo de tratados e para o grupo de controle após o PSM.

TABELA 2
Teste de balanceamento entre tratados e controle no período pré-tratamento

Variáveis	Pareamento – antes (U)/após (M)	Média		Teste t	
		Tratado	Controle	t	p > t
infraestrutura_escolar	U	0,70	0,38	11,38	0,00***
	M	0,64	0,65	-0,17	0,86
inse	U	3,43	3,09	3,02	0,00***
	M	3,42	3,34	0,34	0,74
escola_urbana	U	0,88	0,43	8,19	0,00***
	M	0,88	0,86	0,1	0,92
carga_horaria	U	272,48	282,00	-0,86	0,39
	M	258,38	258,58	-0,01	0,99
media_alunos	U	22,01	17,62	6,05	0,00***
	M	23,67	23,31	0,26	0,80
duracao_ano	U	331,72	314,71	8,68	0,00***
	M	330,33	327,15	0,66	0,51

(Continua)

(Continuação)

Variáveis	Pareamento – antes (U)/após (M)	Média		Teste t	
		Tratado	Controle	t	p > t
num_funcionarios	U	47,42	21,65	10,28	0,00***
	M	51,00	51,89	-0,15	0,88
diretor_indicacao	U	0,65	0,70	-0,73	0,47
	M	0,63	0,62	0,03	0,98
experiencia_diretor	U	2,81	3,14	-1,59	0,11
	M	2,92	2,93	-0,03	0,98
docentes_superior5A	U	69,51	63,21	0,97	0,33
	M	81,62	80,65	0,15	0,88
docentes_superior9A	U	85,23	71,76	2,72	0,01***
	M	87,19	86,92	0,07	0,94

Elaboração dos autores.

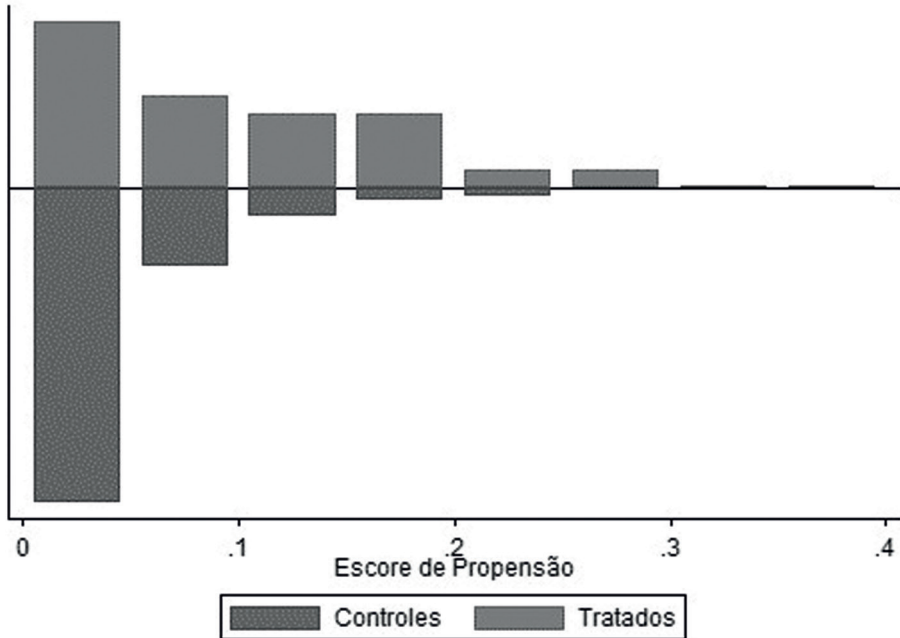
Obs.: *** Estatisticamente significante a 1%.

Ao aplicarmos testes de balanceamento, podemos verificar que se rejeita a hipótese nula de que as médias das variáveis observáveis entre os grupos de controle e de tratamento antes do pareamento são iguais ao nível de 5% de significância estatística em apenas algumas características, desrespeitando as condições de seleção nas observáveis e de suporte comum. No entanto, após o pareamento, percebe-se, por meio dos resultados dos testes de média, que o grupo de controle e o grupo de tratamento tornam-se estatisticamente iguais ao nível de 5% de significância nas características observáveis. Isso quer dizer que, no ano anterior à intervenção da política, os grupos de tratamento e controle são semelhantes quanto às características observáveis.

No gráfico 1, observa-se a existência de suporte comum pela sobreposição do vetor de variáveis observáveis nos grupos de tratamento e de controle após o pareamento, minimizando o viés de seleção nas observáveis. Além disso, o respeito a tal hipótese do PSM permite o relaxamento da hipótese de trajetórias paralelas, condição da estimação do efeito médio do tratamento por meio do método DID.

GRÁFICO 1

Teste de suporte comum das características observáveis de tratados e controles no período pré-tratamento



Elaboração dos autores.

Obs.: Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

5.2 Estimações

Nesta subsecção são apresentados os resultados estimados, bem como uma discussão a respeito do impacto do programa Escola de Valor sobre os indicadores educacionais, resultante da aplicação do método DID, ponderado pelo escore de propensão obtido via PSM.

5.2.1 Impacto sobre a proficiência média

A tabela 3 apresenta os resultados da estimação do impacto do programa de bonificação sobre a proficiência média em português e matemática do 5º e 9º anos do ensino fundamental das escolas premiadas.

Os coeficientes presentes nas tabelas correspondem às interações entre as *dummies* do número de vezes em que as escolas foram escolhidas para receberem a premiação e a *dummy* de tempo (períodos pré e pós-tratamento), e representam o efeito médio do tratamento sobre os tratados (ATT). Esses serão os parâmetros

de interesse para verificar se o Escola de Valor teve efeito sobre os indicadores educacionais.

Visando controlar para outros fatores que poderiam também impactar as notas das escolas, foram incluídas variáveis de controle já utilizadas na estimação do escore de propensão. Também se estimou o modelo DID por meio de duas especificações: um modelo usual de mínimos quadrados ordinários (OLS) e um modelo incluindo efeitos fixos (FE).

TABELA 3
Estimação do impacto do programa Escola de Valor sobre a proficiência média dos alunos

Variáveis	5º ano		9º ano	
	Nota média		Nota média	
	OLS	FE	OLS	FE
δ_1	5,20 (5,24)	7,23** (3,15)	4,25 (6,98)	7,04 (7,21)
δ_2	-4,22 (6,34)	-1,96 (4,25)	-6,19 (8,87)	-8,38 (6,66)
δ_3	-10,24 (7,51)	-7,15 (4,76)	9,39 (6,97)	10,99*** (2,79)
δ_4	14,27 (9,59)	11,42*** (3,19)	-7,17 (7,04)	-2,30 (3,42)
δ_5	0 -	0 -	4,77 (7,41)	-0,87 (3,62)
Controles	Sim	Sim	Sim	Sim
Nº de observações	781	781	709	709
F	9,62***	-	6,62***	-
R ²	0,29	0,40	0,27	0,47

Elaboração dos autores.

Obs.: 1. Estatisticamente significantes: ** a 5%; *** a 1%.

2. Erros-padrão robustos entre parênteses.

Como se pode observar da tabela 3, apenas os resultados das estimações segundo o modelo de efeito fixo dão indícios da existência de impacto positivo do programa de bonificação aos docentes da Paraíba sobre a proficiência média dos alunos do 5º e 9º anos do ensino fundamental. A significância estatística das *dummies* de interação de exposição ao programa e tempo sugere que o programa tende a gerar um efeito positivo cumulativo sobre o desempenho escolar, em vista de uma defasagem entre a premiação e o efetivo impacto desta sobre a nota dos alunos. Em outras palavras, escolas com um maior histórico de premiação tendem, em média, a apresentar melhores resultados em termos de desempenho escolar, relativamente às escolas contempladas menos vezes com as bonificações e as não premiadas.

Semelhante aos resultados dos estudos de Furtado e Soares (2018), Lépine (2016) e Oshiro, Scorzafave e Dorigan (2015), que encontraram efeitos positivos dos programas de bonificação nos estados de Pernambuco e São Paulo sobre as notas médias dos alunos, as evidências deste estudo revelaram um grau de efetividade do programa de bonificação na Paraíba no tocante à melhoria da qualidade da educação, mesmo com diferentes estratégias empíricas. Em contrapartida, as evidências aqui encontradas de um efeito cumulativo da premiação ao longo dos anos nas escolas se contrapõem aos resultados de Furtado e Soares (2018) e Oshiro, Scorzafave e Dorigan (2015), os quais apontaram para uma diminuição expressiva dos efeitos dos programas de bonificação com o passar do tempo.

5.2.2 Impacto sobre a taxa de aprovação

Um dos objetivos do Escola de Valor consiste em premiar as escolas que mostraram êxito no tocante ao processo de ensino e aprendizagem, o qual pode ser medido de forma imediata pelo maior índice de aprovação dos alunos ou redução da repetência. A tabela 4 mostra os resultados obtidos sobre esses indicadores.

TABELA 4
Estimação do impacto do programa Escola de Valor sobre a taxa de aprovação dos alunos

Variáveis	5º ano		9º ano	
	Taxa de aprovação		Taxa de aprovação	
	OLS	FE	OLS	FE
δ_1	2,05 (5,04)	-0,96 (2,50)	10,34** (4,46)	-0,57 (2,12)
δ_2	-2,07 (6,77)	2,95 (2,99)	-13,17** (6,01)	1,76 (2,56)
δ_3	-6,81 (7,05)	-4,34 (4,40)	-1,91 (5,20)	-2,77 (4,30)
δ_4	11,94* (6,15)	9,38** (4,68)	3,90 (3,34)	-4,53 (9,48)
δ_5	0 -	0 -	-0,95 (3,32)	0 -
Controles	Sim	Sim	Sim	Sim
Nº de observações	830	830	795	817
F	36,66***	-	4,92***	-
R ²	0,24	0,14	0,19	0,20

Elaboração dos autores.

Obs.: 1. Estatisticamente significante: * a 10%; ** a 5%; *** a 1%.

2. Erros-padrão robustos entre parênteses.

Como pode ser visto, a maioria dos coeficientes não apresentou significância estatística em ambos os modelos. Ou seja, as taxas de aprovação não responderam significativamente ao programa de bonificação. Por sua vez, as escolas premiadas em quatro anos apresentaram resultados positivos estatisticamente, corroborando a tese de que os efeitos das bonificações tendem a surgir quanto maior for a participação da escola no programa. Embora a *dummy* de dois anos de premiação tenha apresentado um efeito negativo nas taxas de aprovação pelo método OLS, com a inclusão dos efeitos fixos das escolas (que representam um modelo mais completo), o efeito tende a ser nulo.

Considerando que o resultado OLS se verifique na prática, uma provável explicação, e que pode revelar uma faceta adversa dos programas de bonificação, está no desenho do programa voltado à remuneração coletiva em detrimento da bonificação individual. Esse mecanismo de incentivo coletivo pode gerar o chamado problema do *free-rider*, no qual alguns funcionários ganham a remuneração adicional graças ao esforço dos demais, sem, no entanto, contribuir com um maior nível de esforço. Isso pode gerar desmotivação nos funcionários que de fato se esforçam e, conseqüentemente, prejuízos ao trabalho em equipe.

5.2.3 Impacto sobre a taxa de abandono

Conforme citado anteriormente, entre os critérios de elegibilidade do programa está a apresentação de estratégias para a redução da evasão escolar, de maneira a atingir patamares mínimos de abandono. Dessa forma, foi estimado o efeito do bônus sobre este último indicador.

Da análise dos resultados apresentados na tabela 5, pode-se inferir que o programa de bonificação por desempenho apresentou efeito positivo sobre a taxa de abandono escolar para as escolas premiadas em 2016, especialmente para aquelas que já foram premiadas em anos anteriores. Ou seja, houve diminuição da evasão escolar nas escolas premiadas nesta edição e que receberam o bônus anteriormente em outra(s) edições, comparativamente às escolas que não foram premiadas (controles).

TABELA 5
Estimação do impacto do programa Escola de Valor sobre a taxa de abandono dos alunos

Variáveis	5º ano		9º ano	
	Taxa de abandono		Taxa de abandono	
	OLS	FE	OLS	FE
δ_1	-3,09 (1,90)	-4,49** (1,85)	10,34** (4,46)	-0,60 (1,44)
δ_2	4,68* (2,59)	6,14*** (2,30)	6,14*** (1,82)	3,74* (1,94)
δ_3	0,05 (2,72)	-1,13 (2,82)	-1,02 (2,80)	0,92 (1,86)
δ_4	0,27 (2,23)	0,57 (3,15)	3,90 (3,34)	0,43 (1,92)
δ_5	0 -	0 -	0 -	-4,45** (2,09)
Controles	Sim	Sim	Sim	Sim
Nº de observações	830	830	795	795
F	4,97***	-	4,92***	-
R ²	0,17	0,21	0,19	0,18

Elaboração dos autores.

Obs.: 1. Estatisticamente significantes: * a 10%; ** a 5%; *** a 1%.

2. Erros-padrão robustos entre parênteses.

6 CONSIDERAÇÕES FINAIS

Este estudo se insere na discussão empírica acerca da efetividade das políticas de responsabilização educacional, mais especificamente, dos programas de bonificação aos docentes. Como pano de fundo, o estudo avaliou o impacto do Escola de Valor, programa de bonificação aos docentes do estado da Paraíba, que gratifica com um 14º salário os profissionais de educação das escolas públicas estaduais de educação básica escolhidas por meio de um processo seletivo, cujos critérios de seleção passam por fatores como infraestrutura e gestão escolar. Os indicadores de resultado foram: a proficiência média em português e matemática, a taxa de aprovação escolar e a taxa de abandono.

Para atender ao objetivo proposto, utilizou-se a técnica de PSM aliada ao método DID. A partir do PSM, buscou-se a construção do melhor contrafactual do grupo das escolas tratadas por meio do pareamento em termos das características observadas no período pré-tratamento, construindo um suporte comum entre os

grupos de tratamento e controle. Por meio do método DID, compararam-se os resultados ao longo do tempo, visando eliminar as diferenças atribuíveis a não observáveis fixas no tempo. O efeito médio de tratamento sobre os tratados na abordagem do PSM + DID resultou da regressão do modelo DID ponderado pelo escore de propensão obtido no PSM.

Devido à inconstância dos critérios de seleção das escolas premiadas pelo programa durante sua vigência e a ausência de informações relevantes do programa, escolheu-se avaliar o impacto do Escola de Valor em um ano específico, 2016 (ano da intervenção), sendo 2015 o ano pré-intervenção e 2017 o pós-intervenção. Para tanto, utilizaram-se os microdados do Saeb e do Censo Escolar referentes aos anos de 2015 e 2017 para as escolas públicas (municipais e estaduais) do estado da Paraíba, disponibilizados por meio do portal do Inep. Levou-se em conta no modelo o histórico de premiações das escolas, substituindo a *dummy* de tratamento do DID por uma série de *dummies* para as quantidades de vezes em que as escolas receberam as bonificações.

A estimação do impacto do programa sobre as variáveis de resultado apontou para o impacto positivo do programa de bonificação sobre a nota dos alunos das escolas premiadas, assim como sobre a taxa de aprovação e a taxa de abandono dos estudantes, em conformidade com outros resultados encontrados na literatura nacional. Ademais, a significância estatística das *dummies* de interação entre a quantidade de premiações e o tempo sugeriram que o programa tende a apresentar efeito positivo e cumulativo sobre o desempenho escolar. É importante salientar que, nesse sentido, tais resultados se contrapõem aos encontrados por Furtado e Soares (2018) e Oshiro, Scorzafave e Dorigan (2015), dado que, ao analisarem, respectivamente, programas de bonificação dos estados de Pernambuco e São Paulo, os autores concluíram que o efeito dos programas tende a diminuir expressivamente com o passar do tempo.

Entretanto é preciso deixar algumas ressalvas. Como aponta a literatura, o desenho do programa, caracterizado por sistema de remuneração coletiva, pode gerar o problema do *free-rider*, causando uma desmotivação do trabalho em equipe e, em última instância, dos professores em relação à melhora do desempenho escolar. Outra questão referente ao desenho do programa deriva da utilização de múltiplos critérios e objetivos, de forma que os profissionais da educação precisam se esforçar para atingir diversas metas para alcançar a premiação, causando uma possível perda de foco no objetivo do programa. Ademais, mesmo que o mecanismo de incentivo seja bem desenhado, é possível que haja desvios em relação ao seu objetivo inicial, dada a dificuldade de alinhar os interesses das secretarias de educação e dos funcionários avaliados. O fato de o bônus estar condicionado às notas da avaliação interna do IDEPB, e não às notas do Saeb, uma avaliação externa, garante maior

confiabilidade aos resultados obtidos nesse estudo, visto que evita a ocorrência de potenciais vieses, como a ocorrência de *gaming*.

Por fim, deve-se salientar que a indisponibilidade de informações e a consequente ausência de transparência no tocante ao processo no qual se dá as premiações acarretou uma série limitações ao estudo, inviabilizando, inclusive, a aplicação de outros métodos de avaliação de impacto que poderiam conferir maior robustez à análise. Nesse sentido, torna-se imperiosa a cobrança ao poder público por maior transparência na execução das políticas públicas. A disponibilização de informações é de suma importância, uma vez que estas são o principal insumo para a realização de análises econômicas, as quais podem subsidiar a criação de políticas públicas e o aprimoramento das já existentes.

REFERÊNCIAS

ALEXANDRE, M. R.; LIMA, R. S. P. de; WALTENBERG, F. D. Teoria econômica e problemas com remuneração de professores por resultados. **Cadernos de Pesquisa**, v. 44, n. 151, p. 36-61, jan.-mar. 2014.

BISHOP, J. H.; WOESMANN, L. Institutional effects in a simple model of educational production. **Education Economics**, v. 12, n. 1, p. 17-38, abr. 2004.

BONAMINO, A.; SOUSA, S. Z. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. **Educação e Pesquisa**, v. 38, n. 2, p. 373-388, abr.-jun. 2012.

BUKOWSKI, P.; KOBUS, M. The threat of competition and public school performance: evidence from Poland. **Economics of Education Review**, v. 67, p. 14-24, dez. 2018.

CAVALCANTE, M. do S. S. **Gestão e avaliação da educação**: o Avaliando IDEPB e o Prêmio Escola de Valor – o projeto de formação humana no contexto da hegemonia neoliberal. 2019. 397 f. Tese (Doutorado) – Universidade Federal da Paraíba, João Pessoa, 2019. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/15307>. Acesso em: 20 dez. 2020.

COWAN, J.; GOLDHABER, D. Do bonuses affect teacher staffing and student achievement in high poverty schools? Evidence from an incentive for national board certified teachers in Washington State. **Economics of Education Review**, v. 65, p. 138-152, ago. 2018.

FOGUEL, M. N. Diferenças em diferenças. *In*: MENEZES-FILHO, N. (Org.). **Avaliação econômica de projetos sociais**. 1. ed. São Paulo: Dinâmica Gráfica e Editora, 2012a. p. 85-110.

FOGUEL, M. N. Modelo de resultados potenciais. *In*: MENEZES-FILHO, N. (Org.). **Avaliação econômica de projetos sociais**. 1. ed. São Paulo: Dinâmica Gráfica e Editora, 2012b.

FRYER JUNIOR, R. G. **Teacher incentives and student achievement**: evidence from New York City public schools. Cambridge, Estados Unidos: Harvard University; NBER, nov. 2011.

FURTADO, C. S. do V.; SOARES, T. M. O impacto da bonificação educacional em Pernambuco. **Estudos em Avaliação Educacional**, v. 29, n. 70, p. 48-76, jan.-abr. 2018.

GERTLER, P. J. *et al.* (Ed.). **Impact evaluation in practice**. 2. ed. Washington: World Bank, 2011.

GLEWWE, P.; ILIAS, N.; KREMER, M. **Teacher incentives**. Cambridge, Estados Unidos: NBER, abr. 2003. (Working Paper n. 9671).

HECKMAN, J. J.; ICHIMURA, H.; TODD, P. Matching as an econometric evaluation estimator. **Review of Economic Studies**, v. 65, n. 2, p. 261-294, abr. 1998.

HIRANO, K.; IMBENS, G. W.; RIDDER, G. Efficient estimation of average treatment effects using the estimated propensity score. **Econometrica**, v. 71, n. 4, p. 1161-1189, jul. 2003.

KHANDKER, S. R.; KOOLWAL, G. B.; SAMAD, H. A. (Ed.). **Handbook in impact evaluation**: quantitative methods and practices. Washington: World Bank, 2010.

LAFFONT, J.-J.; MARTIMORT, D. (Ed.). **The theory of incentives**: the principal-agent model. Princeton: Princeton University Press, 2002.

LAVY, V. Evaluating the effect of teachers' group performance incentives on pupil achievement. **Journal of Political Economy**, v. 110, n. 6, p. 1286-1317, dez. 2002.

LAZEAR, E. P. **Paying teachers for performance**: incentives and selection. Stanford: Stanford University, 2001. Mimeografado.

LÉPINE, A. G. **Essays in the economics of education in Brazil** – ensaios em economia da educação no Brasil. 2016. 119 f. Tese (Doutorado) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2016.

MURALIDHARAN, K.; SUNDARARAMAN, V. Teacher performance pay: experimental evidence from India. **Journal of Political Economy**, v. 119, n. 1, p. 39-77, fev. 2011.

NEAL, D.; SCHANZENBACH, D. W. Left behind by design: proficiency counts and test-based accountability. **The Review of Economics and Statistics**, v. 92, n. 2, p. 263-283, maio 2010.

OSHIRO, C. H.; SCORZAFAVE, L. G.; DORIGAN, T. A. Impacto sobre o desempenho escolar do pagamento de bônus aos docentes do ensino fundamental do estado de São Paulo. **Revista Brasileira de Economia**, v. 69, n. 2, p. 213-249, abr.-jun. 2015.

PARAÍBA. Secretaria do Estado da Educação. **Edital nº 2/2011-GS-Prêmio Educação Exemplar**. João Pessoa: SEE, 2011. 10 p.

PARAÍBA. Lei nº 9.879, de 13 de setembro de 2012. Institui, no âmbito do Poder Executivo estadual, os Prêmios Mestres da Educação e Escola de Valor e dá outras providências. **Diário Oficial do Estado da Paraíba**, João Pessoa, n. 15.039, p. 1, 14 set. 2012. Seção 1.

PARAÍBA. Secretaria do Estado da Educação. **Edital nº 3/2013-GS-Prêmio Escola de Valor**. João Pessoa: SEE, 2013. 10 p.

PARAÍBA. Secretaria do Estado da Educação. **Edital nº 4/2014-GS-Prêmio Escola de Valor**. João Pessoa: SEE, 2014. 9 p.

PARAÍBA. Secretaria do Estado da Educação. **Escolas contempladas no Prêmio Escola de Valor, edição 2015 (Edital nº 14/2015-GS)**. João Pessoa: SEE, 2015. 7 p.

PARAÍBA. Secretaria do Estado da Educação. **Edital nº 5/2016-GS-Prêmio Escola de Valor**. João Pessoa: SEE, 2016. 11 p.

PINTO, C. C. X. Pareamento. *In*: MENEZES-FILHO, N. (Org.). **Avaliação econômica de projetos sociais**. 1. ed. São Paulo: Dinâmica Gráfica e Editora, 2012. p. 111-144.

RAPLE, B. A. Payment by results: an example of assessment in elementary education from nineteenth century Britain. **Education Policy Analysis Archives**, v. 2, n. 1, p. 1-21, jan. 1994.

REBACK, R. Teaching to the rating: school accountability and the distribution of student achievement. **Journal of Public Economics**, v. 92, n. 5-6, p. 1394-1415, jun. 2008.

ROCHA, R.; SOARES, R. R. Evaluating the impact of community-based health interventions: evidence from Brazil's Family Health Program. **Health Economics**, v. 19, n. S1, p. 126-158, set. 2010.

RODRIGUES, J. da S.; SILVA, A. F. da. As políticas de *accountability* na rede estadual de ensino da Paraíba e o Prêmio Escola de Valor. *In: EPEN – REUNIÃO CIENTÍFICA REGIONAL NORDESTE DA ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM EDUCAÇÃO*, 25., 2020, Salvador, Bahia. **Anais...** Salvador: Anped, 2020. Disponível em: http://anais.anped.org.br/regional/sites/default/files/trabalhos/20/7045-TEXT0_PROPOSTA_COMPLETO.pdf. Acesso em: 20 dez. 2020.

SOARES, J. F. Melhoria do desempenho cognitivo dos alunos do ensino fundamental. **Cadernos de Pesquisa**, v. 37, n. 130, p. 135-160, jan.-abr. 2007.

SWAIN, W. A.; RODRIGUEZ, L. A.; SPRINGER, M. G. Selective retention bonuses for highly effective teachers in high poverty schools: evidence from Tennessee. **Economics of Education Review**, v. 68, p. 148-160, fev. 2019.

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

WOOLDRIDGE, J. M. **Introdução à econometria: uma abordagem moderna**. 4. ed. São Paulo: Cengage Learning, 2013.

Originais submetidos em: abr. 2020.

Última versão recebida em: abr. 2021.

Aprovada em: abr. 2021.

SOCIOECONOMIC AND DEMOGRAPHIC DETERMINANTS OF MIGRATION IN BRAZILIAN MUNICIPALITIES: 2000-2010¹

Luís Abel da Silva Filho²

Alexandre Gori Maia³

Brazilian economic transformations are responsible for the dynamics of the population of its territory. There are many studies that attribute a substantial part of migration decisions to the labor market, without, however, considering other possible determinants. In this sense, this paper aims to analyze the determinants of the Brazilian migratory dynamics, considering multiple dimensions of analysis. Literature is reviewed and empirical treatment is given by the construction of the Migration Effectiveness Index (MEI) and the use of fixed effects panel data models, with municipalities as units of analysis. The results show that labor market variables, such as occupation, work in industry, commerce and services, explain to a greater extent the participation of migrants in the municipalities of the country. Regarding emigration, it is the negative effects of these variables that motivate people to leave their municipality. In this sense, it is, therefore, the labor market that shows the greatest relation with the migratory dynamics in Brazil.

Keywords: migration dynamics; infrastructure; human capital; job market.

DETERMINANTES SOCIOECONÔMICOS E DEMOGRÁFICOS DAS MIGRAÇÕES NOS MUNICÍPIOS BRASILEIROS: 2000-2010

As transformações econômicas brasileiras são responsáveis pela dinâmica da população ao largo de seu território. Muitos são os estudos que atribuem ao mercado de trabalho parte substancial das decisões de migração, sem, contudo, considerar outros possíveis determinantes. Nesse sentido, este artigo tem como objetivo analisar os condicionantes da dinâmica migratória brasileira, considerando-se múltiplas dimensões de análises. Revisa-se a literatura e o tratamento empírico é dado a partir da construção do Índice de Eficácia Migratória (IEM) e do uso de modelos de dados em painel de efeitos fixos, tendo os municípios como unidades de análise. Os resultados mostram que as variáveis de mercado de trabalho, tais como: ocupação, trabalho no setor da indústria, comércio e serviços, explica em maior proporção a participação de migrantes nos municípios do país. Já em relação à emigração, são os efeitos negativos dessas variáveis que motivam a saída de pessoas de seu município. Nesse sentido, é, pois, o mercado de trabalho que apresenta a maior relação com a dinâmica migratória no Brasil.

Palavras-chave: dinâmica migratória; infraestrutura; capital humano; mercado de trabalho.

JEL: J0; J15; J61.

1. DOI: <http://dx.doi.org/10.38116/ppe52n3art3>

2. Professor at the Department of Economics at the Regional University of Cariri (Urca). E-mail: abeleconomia@hotmail.com.

3. Full professor at the Institute of Economics at the State University of Campinas (Unicamp). E-mail: gore@eco.unicamp.br.

1 INITIAL CONSIDERATIONS

This paper aims to study the issues that impact the migratory dynamics, comparing the variables that affect immigration (entry of people) and emigration (exit of people) in Brazilian municipalities. The basic assumption taken for the analysis is that the socioeconomic development and the population dynamics of the Brazilian population were marked by sharp differences in their consolidation. Regional economic concentration; indicators of disparate housing conditions; inequality in the labor market; disparities in national human capital; unemployment rates with substantial differences between regions; different activity conditions among the economically active national population; substantially higher income differentials among those employed in different regions of the country, among other issues, marked the population movement in search of better socioeconomic conditions in economically more dynamic regions (Sjaastad, 1980; Lee, 1980; Martine, 1990; Camarano and Abramovay, 1998; Pacheco, 1998; Guimarães Neto, 1997; Diniz, 2001; Araújo, 2000; Brito, 2006).

The socioeconomic context of the country was an inducer of migratory dynamics, especially in years of low economic growth and with a high incidence of climatic volatility in regions of agrarian economies, which caused the need to migrate as a strategy to overcome the challenges imposed by local conditions. Regional differences and the search for job opportunities, faced with a low economic growth and low labor availability in some regions, motivated a migratory movement in search of occupational insertion, in a context of limited employment possibilities in the regions source (Sjaastad, 1980; Lee, 1980; Martine, 1990).

The regional socioeconomic dynamics was the first characteristic observed as a factor of attraction of migrants in the classic studies of Sjaastad (1980), Lee (1980) and Martine (1987). The expected income in the region of destination substantially higher than the income earned in the region of origin is a significant factor in the process of expulsion and attraction of labor. Even knowing that factors that are implicit in the decision to migrate have strong impacts, the personal desire to improve socioeconomic conditions is decisive in the final decision (whether or not to migrate).

In this sense, the spatial mobility of the population is intrinsically related to social mobility. Leaving the origin for another destination is the search for better working conditions, with which it is expected that there will be greater possibilities of social mobility. Considering studies by Maia and Quadros (2009) and Maia (2013) it is possible to observe that social mobility through work has become an important condition for social improvement in the Brazilian economy in the 2000s. In addition, recent studies converge to the logic that has been set especially in the

theoretical models on migration for many years, affirming that the social mobility of the migrant happens through work (Lima and Vale, 2001; Santos Junior, 2002).

The historical-structural approach proposes that the decision to migrate comes from the desire for better remuneration and work conditions offered at the place of destination in relation to the place of origin. Todaro (1980) considered that the decision to migrate is a decision based on factors of an economic nature and is directly influenced by the possibilities of better socioeconomic insertion conditions of individuals. Based on this principle, the observable individual characteristics (race/color, sex, schooling, among others) and unobservable (determination, motivation, entrepreneurship, among others) analyzed in the migration decision are only highlighted in more recent empirical studies. But the possibility of social ascension of the individual is what most explains the decision of the migration.

Over the years, the migratory movement has been extremely high, with the most diverse destinations among the regions of the country, but with specific foci in regions where economic development was potentially high. For many decades, the Brazilian Southeast has been, in both absolute and relative terms, the most attractive region of intraregional and interregional migrants. In contrast, the Northeast showed significant expulsion in the last fifty years of the 20th century (Netto Júnior, Porto Junior and Figueiredo, 2008), but with marked changes in recent years (Oliveira and Jannuzzi, 2005).

The low economic dynamism of some regions is shown to be a high factor to explain the migratory processes (Singer, 1980; Taylor, 1999; Borjas, 1996; Pacheco and Patarra, 1997). The dynamics of the population is driven by economic dynamics, which is the main attraction factor for migrants in Brazil and in the world (except for the few countries where migrations are motivated by wars, climatic catastrophes, religious persecutions, among others), including that the decision is individual (Singer, 1980; Sawyer, 1984; Cançado, 1999; Santos and Moreira, 2006; Mata et al., 2007; Cambota and Pontes, 2012; Freguglia and Menezes-Filho, 2012).

The migratory dynamics in Brazil is a result of the socioeconomic characterization of its regions. Regional productive concentration; low level of human capital; substantially high socioeconomic disparities are major determinants of national migration dynamics. Before the Brazilian economic situation and the above mentioned factors, studying the Brazilian migratory dynamics is substantially relevant from the point of view of the dynamics of human capital in search of opportunities for occupational insertion. Besides, the Brazilian migratory flows are high.

In recent years, short-distance migration has predominated. Improvement in national economic indicators, such as: economic growth; regional productive decentralization; reduction of regional and economic inequalities of the population; greater formalization in the labor market, generally across the country, have ended

up impacting on the reduction of long-distance flows and accentuating short-distance migration in the national territory. Municipalities of economic growth became important recipients of internal migrants. In order to study the Brazilian migratory dynamics, considering inter-municipality migration is important to consider the short-term flows in the national migratory dynamics in a temporal cut (Demographic Censuses of 2000 and 2010) where the short-distance flows in the national migratory dynamics stand out.

Based on the context presented, this paper aims to analyze the Brazilian socioeconomic dynamics and the socioeconomic development factors involved in the attraction and repulsion of in-migration and out-migration, respectively, in the municipalities of the country. The specificity of this research lies in the joint analysis of the inter-municipality⁴ and individual socioeconomic factors that impact the migration decision. From this perspective, the development of the work is based both on a historical-structuralist approach and on the maximization of individual well-being.

Thus, the intention is to observe the socioeconomic differences and the profile of the migrants, as well as their socioeconomic insertion in the Brazilian municipalities, since the economic divergence of the country is extremely high. However, the analysis intends to answer some questions related to population displacement, which are: what are the relations between the municipality socioeconomic conditions and the migratory dynamics? How do migrants enter socioeconomically in places of destination?

Considering the decision to migrate as being influenced by socioeconomic and individual factors as diverse as possible, the magnitude of the economic variables influences the population dynamics (infrastructure, labor market, unemployment rate, wages, among others). In addition, it is worth pointing out that it is possible to measure the influence that each of the individual characteristics (socioeconomic and demographic) used can exert on decision-making and migratory processes. Besides, it should be noted that the socioeconomic insertion of migrants in the municipalities of the country, based on the last two Demographic Censuses, allows us to observe both the migratory dynamics and the economic and social context resulting from the socioeconomic transformations experienced in Brazil. From this perspective, the purposes of this study are established.

The paper is divided into five sections, besides these initial considerations, the second section presents the methodological procedures used; in the third section, there is empirical evidence of migration off the national territory; soon after, in

4. Intermunicipality migration is one in which a person changes from one municipality to another. It is an intermunicipality migrant that person who migrates from a municipality belonging to a state to another municipality belonging to another state, that is, interstate; or one that migrates from a municipality from one state to another municipality in the same state, that is, intra-state. Thus, intermunicipality migration happens from one municipality to another, whether of the same or another state.

the fourth section, the results of the Migration Effectiveness Index (MEI) are presented; in the fifth section, the results obtained from the panel data model; in the sixth section, the final considerations are made.

2 METHODOLOGICAL PROCEDURES

This paper analyzes the impacts of socioeconomic variables on migratory dynamics: attraction and repulsion in Brazilian municipalities, based on microdata from the Demographic Census of Brazil in the years 2000 and 2010. In addition, information is used here about heads of households in the two years under analysis. The fixed date migration was the one used here. In Brazil it is considered the one in which the migrants are less than five years old in the current municipalities at the moment the Census goes to the field. That is, if the person answered that resided in another municipality on 7/31/1995 and 7/31/2005, in the 2000 and 2010 census,⁵ respectively, this person is a fixed date migrant. Likewise, a fixed-date migration was used in order to verify the effect of short-distance migration in the last Demographic Census, as well as to capture the effects of Brazilian socioeconomic transformations and their impacts on migration dynamics in recent years, excluding the other types of migration.

The variables used throughout the paper are named in the following.

- 1) Households in the destination with responsible in-migration.
- 2) Households at the origin with those who emigrated.
- 3) Households with a water supply service through a general network.
- 4) Households with sewage system by general sewage or rainwater network, septic tank and rudimentary septic tank.
- 5) Households with garbage collection by cleaning service or placed in a cleaning service bucket.
- 6) Households with electricity supply by distributing company or other sources.
- 7) Households in which the responsible person claimed to be literate.
- 8) Households in which the responsible person claimed to have completed high school.
- 9) Households in which the responsible person has complete higher education.
- 10) Households in which the reference person answered to be occupied in the reference week of the survey.

5. The Demographic Census in Brazil is held every 10 years, from the year 1950 to the year 2010, the last year. In that period, only one Census was not decennial, that of 1991.

- 11) Households in which the person responsible for the household responded to be a contributor to an official social security institute in the main job or other job.
- 12) Households in which the person responsible responded to be engaged in sectors of economic activities of industry, commerce or services (ICS).
- 13) Households in which the person in charge claimed to be poor at work. That is, to have remuneration lower than ¼ minimum wage.

2.1 MEI

In order to analyze the Brazilian municipal migratory dynamics, an important index used by the international and national literature is the MEI. This was created and indicated by the United Nations (UN, 1970) and suggests a way of comparing the internal migratory dynamics of a country, without risks of fragility of the indicator over time. With this, one can make an internal analysis of the migrations in different time clippings, maintaining the robustness of the results. Thus, a dynamic analysis of the migrations in the territory is considered, considering the emigrations and immigration of each unit of analysis (municipalities, states or regions; exits and entrances). With the index, it is possible to observe the municipalities (unit of analysis of this paper) for their capacity of retention, evasion and migratory turnover. Therefore, it is necessary to build the Brazilian migratory matrix, presented here at municipality levels. The migratory matrix of this paper is made from the following mathematical denomination. Let A be any migratory matrix.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} \end{pmatrix}, \quad (1)$$

in which

- A_{ij} = exit of the migrants from the municipality i to the municipalities j in the period t under analysis;
- $\sum a_{1j}$ = total number of people who emigrated from the municipality i to j municipalities in Brazil; and
- $\sum a_{i1}$ = the total number of people who are in-migration in the municipality i and who left the j municipalities in Brazil.

From this definition, it is possible to build a flow analysis for the Brazilian migrants responsible for the households and then build the MEI. The values assumed by the index vary between - 1,00 and 1,00. The mathematical formula is given by the following expression.

$$MEI = \left[\frac{(I_i - E_i)}{(I_i + E_i)} \right], \quad (2)$$

in which I_i and E_i are, respectively, the in-migration residing in the municipality i and the out-migration of the municipality i in the two Censuses under analyzes.

With the elaboration of the index, it is worth mentioning that the national literature (Oliveira, 2011) and international literature (Menezes, 2003) interpret the values approximately as follows: $-1,00 < -0,13$ are considered areas of migratory loss; $\geq -0,13 \leq 0,12$ are attributed as areas of migratory turnover; and $> 0,12 \leq 1,00$ are classified as a migratory retention area.

The use of the index has the purpose of presenting, from maps, the Brazilian municipalities that present a high population loss (in the case of Brazil, the outflows occur mainly due to economic reasons. That is, low economic dynamics of the municipalities which induces the outflows movement in search of work); the areas which attract and create labor constantly, being, therefore high in entrance and exit. These municipalities are those of temporary economic dynamism, without, however, establishing a dynamic economy base capable of attracting and maintaining the work force, configuring, this way, as municipalities where there is migratory turnover; the retention areas, in the Brazilian case, can be classified as those which attract the work force for considerable periods of time. That is, they are municipalities located in areas of constant economic growth.

2.2 Panel data model

The second part of this paper makes use of the panel data model by the fixed effects approach. Municipalities are used as cross-sectional units. The municipalities in 2010 were reconciled to the year 2000, in such a way that the new dismembered municipalities were grouped to their municipalities of origin, as stated in the 2000 Demographic Census, so that the data matrix could be reconciled. Thus, the analysis panel had 5,507 municipalities in two periods, totaling 11,014 records.

The panel data model consists of associating time series and cross section data. This results in the joining of information from several units of monitored analysis over time. In this sense, the composition of the units is given from $i = 1, 2, 3 \dots, N$ units at any time interval in which $t = 1, 2, 3 \dots, T$ periods of time used in time series in i units of observation.

Thus, the units of analysis and time are crucial to the construction of the data panel to be used in studies of this nature. Hence, the mathematical representation takes the following form.

$$Y_{it} = \alpha_i + X_{it}\beta + \varepsilon_{it}, \quad (3)$$

in which α_i assumes the specific effects of the cross-sectional units, and these are considered constant over time under study; and, ε_{it} assumes the error term not specified in the model. The α_i factor can be controlled by the fixed or random effects

approach. The first, fixed effects, assumes that this factor can be correlated to the characteristics of the municipalities X_{it} . The second, random effects, assumes that this factor is not related to the characteristics of the municipalities. Fixed effects approach were opted in this paper, since the decision to migrate (attractiveness, for example), would also be associated to the socioeconomic characteristics of the municipalities.

Thus, the correlation between the characteristics of the units and the random variables used in the panel data model used in this study is sought. The participation of migrants in municipality i can be explained by circumstantial variables of basic household infrastructure, such as: energy service, water, sewage and garbage collection; variables of municipality human capital, such as: participation of literate people and participation of people with at least high school; participation of persons with complete higher education; and labor market variables, such as: occupancy rate (participation of persons employed in the Economically Active Population – EAP in the country), social security contribution and participation of employees in ICS by municipality.

3 BRAZILIAN MIGRATORY DYNAMICS: SOCIOECONOMIC AND DEMOGRAPHIC ASPECTS

The limitations of geographic space, working conditions, concentrated income, among others, have provided crucial problems in Brazilian socioeconomic development. Political and economic issues were strongly influenced by the urban dynamics registered by the concentration of production and the fluidity with which the population dynamics intensified. The structure of the labor market always had a strong influence on the Brazilian migratory movement, especially at the boom of the industrialization process, as well as the advance of the national agricultural frontier (Todaro, 1980; Wood, 1982; Martine, 1990). On the one hand, substantial changes in the productive structures strongly affected the inclusion of labor, especially the one with low level of schooling and reduced professional experience in the great industrial development centers of the country. On the other hand, the search for work in the agricultural frontier regions absorbed for many years the population surplus not absorbed by the activities related to industry.

Studies of this nature confirm the hypothesis of the historical-structural model that determines migration (Singer, 1980). It is understood that the decision to migrate is due to the deficient capacity of the installed productive structures and negatively affect the economy of the regions that expel the population. Substantial changes in the regional productive dynamics reinforce a migratory movement in search of better conditions of existence in other regions.

According to Ramalho (2005), the regional economic disparities are in a country with a high population distributed in sharply disparate territory, as well as the concentration of monetary income, which determines the migratory movement. Changes in regional productive structures change the direction of migratory flows. In this case, it is the search for better working conditions, mainly with a higher remuneration than the one obtained at the origin, which determines the destination. In the intermunicipality migratory dynamics, issues related to household infrastructure and the possibility of access to human capital training services are also relevant in the decision to migrate.

These inequalities in regional productive structures have a marked impact on the regional distribution of the Brazilian population's monetary income. The more economically developed regions are the ones that absorb the most part of the work force, and with the best salaries practiced, which encourages interregional and intra-regional migration in these spaces (Freguglia, Menezes-Filho and Souza, 2007). Even in view of the better performance assumed in recent years in productive structures at regional scales in the country, Freguglia and Menezes-Filho (2012) verified the high effect of wage differentials on regional scales influencing the migratory movement motivated by higher labor income. However, a relatively small part of the migrant population is influenced by labor market dynamics in sectors of high technological intensity and substantial capacity to absorb skilled labor. The majority of the migrant population occupies jobs in more traditional sectors.

Thus, some studies in Brazil have sought to list determinants of migratory processes, with regional income disparities as conditioning factors, without, however, analyzing socioeconomic issues of a structural nature in the destination of migrants (Ferreira and Diniz, 1995; Menezes and Ferreira-Júnior, 2003; Netto Júnior, Porto Junior and Figueiredo, 2008). In this context, questions related to regional income inequality, through the spatial allocation of the work force by migratory processes, justify a series of empirical studies.

In addition, there is considerable production of migratory selectivity⁶ in international studies (Borjas, 1987; 1998; Chiswick, 1978; 1999) and national (Ramalho, 2005; Silva, Silva-Filho and Cavalcanti, 2016) that attribute to migration the loss of qualified or entrepreneurial work force, since factors such as better remuneration have significant importance in the decision to migrate. Thus, the entrepreneurs, who are part of the migratory contingent, seek to live in urban areas of greater economic relevance and with greater possibilities of social ascension.

6. By this model, it is understood that the immigrants are part of the population with individual characteristics more aggressive, enthusiastic, enterprising and motivated to look for better opportunities. Therefore, they are a positively selected sample.

Freguglia, Menezes-Filho and Souza (2007) and Freguglia and Menezes-Filho (2012) showed that migrants from economically less developed regions seek better conditions for occupational insertion. For the authors, there are salary differentials between regions and this is one of the incentives to migration in the studies conducted. Given this, positive selection may not happen precisely because of the greater qualification or determination in the individual view, but because of the conditions offered by the labor market in the destination, contrary to some of the works that elect migratory selectivity as a phenomenon that accentuates the regional income disparity in regional scale.

In addition, Brazilian agribusiness cities have also become important poles of opportunity: these cities receive substantial participation of migrants over the years in search of work. They are important destinations for the workforce and meet the needs of unskilled labor, especially for direct activities in the countryside. Thus, it is not possible to overlook the important role of agribusiness in the Brazilian migratory dynamics, especially with the advancement of the national agricultural frontier and the consolidation of countryside activities in the attraction of agroindustries in the years following the advance of grain production in the country (Guimarães and Leme, 2002; Juttel, 2007).

The impacts of the expansion of the national agricultural frontier on population dynamics are still recorded in a few municipalities in the Brazilian Northeast. This dispersion of the national migratory dynamics is accentuated by the modernization and agricultural occupation in the less populated areas of the national territory (Todaro, 1980; Wood, 1982; Martine, 1990; Camarano and Abramovay, 1998; Guimarães and Leme, 2002; Juttel, 2007). According to Brito (2006) the Midwest still stands out in the migratory dynamics of the 2000s, as well as the North of the country that has experienced significant growth in migrations in recent years (Jakob and Jakob, 2015).

The Midwest, with the development of large-scale agribusiness, has become an important nucleus of attraction for migrants throughout the last decades of the twentieth century (Martine and Camargo, 1984; Camarano and Abramovay, 1998) and the beginning of the 2000s (Guimarães and Leme, 2002; Brito, 2006; Juttel, 2007). So, can structural issues impact the Brazilian migratory dynamics? Is there a correlation between the level of municipality socioeconomic development and the attraction of people across the national territory?

The national economic situation in the years 2000, with regional productive decentralization, although small; reduction of socioeconomic inequalities across Brazilian territory; emergence of poles of economic growth in the medium cities of the country, becomes an important reason to study the intermunicipality migration dynamics. In the last Demographic Census, long-distance migration has been

giving way to a short-distance migratory movement in the country (Vasconcellos and Rigotti, 2005). In these aspects, the following section seeks to contemplate the purpose of the paper to present the determinants of migration in a context of socioeconomic performance of the municipalities of Brazil.

3.1 Determinants of migration in Brazilian municipalities

Brazilian migratory dynamics is defined by the intensity and volume of migratory flows. The patterns of classification and the determinants of people's mobility in the country assume the most diverse motives over the years. Issues related to the economic development of some Brazilian regions and factors of climatic nature took on a substantial dimension in the analysis of the migratory dynamics in the second half of the twentieth century (Martine and Camargo, 1984; Camarano and Abramovay, 1998; Aydos, 2010). According to Myrdal (1956), the main determinants of demographic dynamics are related to their inequalities in internal economic geography. In this interpretation, economic issues are one of the main determinants of population mobility.

It is possible to emphasize that from the 1940s to the 1980s, the Brazilian migratory dynamics was mainly based on issues related to climatic factors (Camarano and Abramovay, 1998) – exits from the Northeast – and related to the socioeconomic development of the country (Martine and Camargo, 1984). The construction of access roads and the labor market in industrial sectors that emerged in the regions of greater economic importance were crucial to the promotion of the internal population movement.

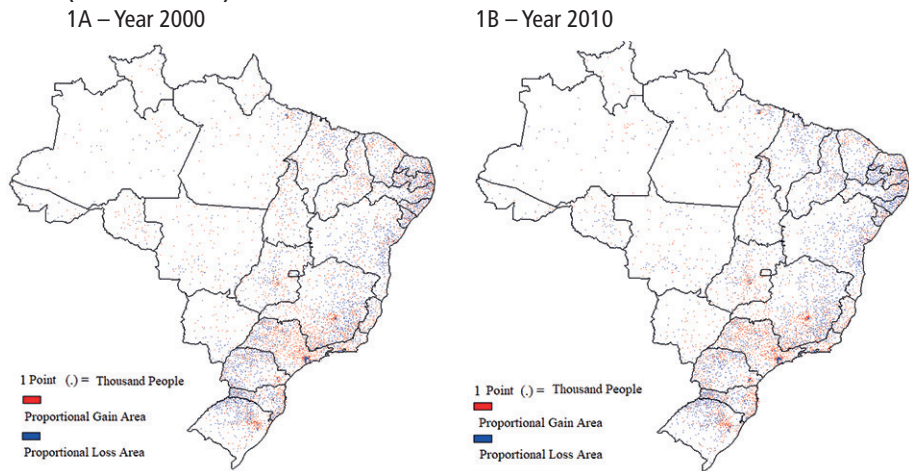
In recent years, especially after the policies of productive deconcentration and the reduction of Brazilian regional inequalities, new migratory flows and new directions have been recorded (Vasconcellos and Rigotti, 2005; Lima and Braga, 2013; Gama and Machado, 2014). The new areas of potential economic development in the Midwest and North of the country (Jakob and Jakob, 2015), in addition to the movements registered to the Northeast, especially return migration, recent phenomena marked the new phase of Brazilian internal migratory processes (Guimarães and Leme, 2002; Juttel, 2007; Justo et al., 2010; Silva, Silva-Filho and Cavalcanti, 2016).

Figure 1 shows the balance of migrants in Brazilian municipalities, considering the fixed date migration and the clipping by heads of household in the years 2000 and 2010. From the results, it is possible to perceive that the great majority of municipalities, especially in the North and in the Northeast had net loss of people, considering only the movement of entry and exit, without considering here the vegetative growth.⁷

7. Difference between birth rate and mortality in municipalities.

National socioeconomic transformations, besides reducing the intensity of migratory flows, also reduce the relative participation of migrants. As recorded by Vasconcellos and Rigotti (2005), the strong concentration of migrants in the 2000 census occurred among those who were in the range of up to 250 kilometers away from the municipality of origin. That is, migration happens with reduced distance for most of the recorded flows. This may result from the various needs in the mobility of the population: work, study, or other subjective issues of the migrants.

FIGURE 1
Net migration (fixed time migration) of heads of households in Brazilian municipalities
(2000 and 2010)



Source: Demographic Census data from Brazil (2000 and 2010).

Authors' elaboration.

Obs.: Figure whose layout and texts could not be formatted due to the technical characteristics of the original files (Publisher's note).

In both 2000 and 2010, most of the Brazilian municipalities that had positive balances are close to municipalities with net loss of migrants. That is, considering the distance of the flows, this can reflect the intermunicipality migration, in the majority, within the same state or to near states, since the distance of the flows reduces over the years. The results reflect only the balance, without a deeper analysis of the flows (not the central objective of this article). This may be due to the fact that household heads change more often because of the need to work or to provide better opportunities for families, such as the possibility of study for their children.

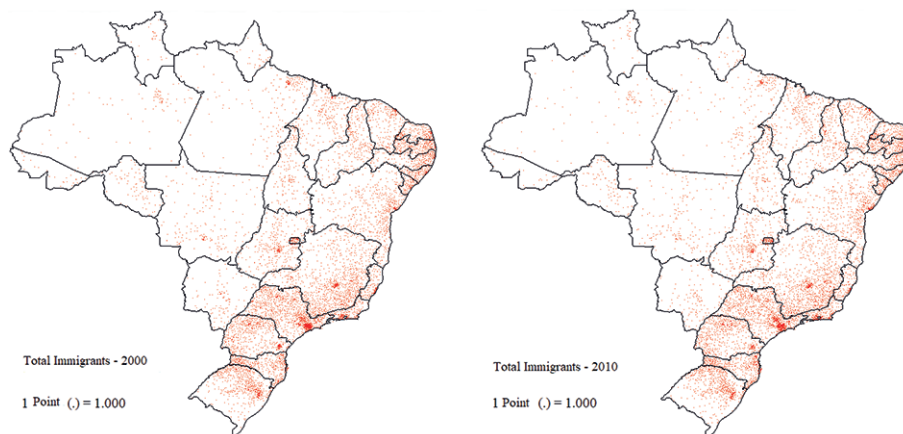
Figure 2 shows the concentration of migrants in the metropolitan areas of the country in the two years studied. The coastal areas of the Northeast and the capitals of the Southeast and South of Brazil have the highest absolute concentration of migrants in both years.

FIGURE 2

Total Brazilian migrants with a fixed date (2000 and 2010)

2A – Year 2000

2B – Year 2010



Source: Demographic Census data from Brazil (2000 and 2010).

Authors' elaboration.

Obs.: 1. Each [.] map point equals 1,000 people.

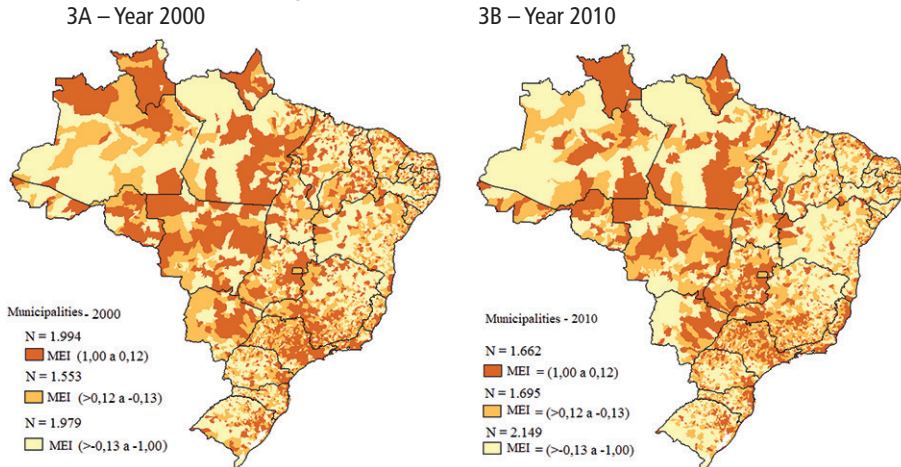
2. Figure whose layout and texts could not be formatted due to the technical characteristics of the original files (Publisher's note).

From the MEI, it is possible to classify areas such as retention, evasion and turnover in the macro-regional context. The migration efficacy indexes that can be spatially observed in figure 3 show that in the year 2000, many municipalities, mainly in the Brazilian Northeast, were characterized as areas of population evasion, despite a high return recorded in the migratory dynamics of the region (Oliveira and Jannuzzi, 2005; Queiroz and Santos, 2011; Justo et al., 2010; Queiroz and Baeninger, 2013). These results can find empirical support for the fact that regional economic development, even with productive deconcentration and reduction of regional disparities, is still a concentrated development in areas of greater capacity to absorb economic activities, given, in particular, its potential level of infrastructure (Silva Filho, Silva and Queiroz, 2015).

Thus, metropolitan municipalities and regions with high potential for development present the best results in the MEI. As can be seen, in the metropolitan municipalities and in the medium-sized cities of the Northeast, the value of the indicator accuses them as areas of population retention, with less developed areas such as high migratory turnover or evasion areas in the year 2000. Furthermore, the Southeast region, especially the state of São Paulo, Espírito Santo and southern of Minas Gerais, present indexes related to retention and/or migratory turnover, with few municipalities classified as areas of population evasion. On the other hand, the north of Minas Gerais and the whole backwoods of Bahia stand out in areas of municipalities with evasive populations.

In the year 2000, the concentration of municipalities in the Midwest and the North as a migratory retention area is extremely high. In the North, as highlighted by Jakob and Jakob (2015), five of the region's capitals had population growth above 3%, and only one capital outside this region obtained these results. Besides, the high process of regional agro-industrialization has empirical support in the literature to justify the intense migratory movement to the Midwest region (Guimarães and Leme, 2002; Juttel, 2007), as well as the permanence of the family migrants in the states. Goiás and Mato Grosso stand out in the regional migratory retention (Juttel, 2007). The expansion of the agricultural frontier coupled with the region's development programs, based on the development of agribusiness, has an important impact on population attraction and retention.

FIGURE 3
MEI in Brazilian municipalities (2000 and 2010)



Source: Demographic Census data from Brazil (2000 and 2010).

Authors' elaboration.

Obs.: Figure whose layout and texts could not be formatted due to the technical characteristics of the original files (Publisher's note).

In 2010, as can be seen from the map, although the Midwest region holds the highest values assumed by the MEI, other municipalities in other regions of the country have assumed values classified as units of population retentions. The results may be associated with the improvement in socioeconomic development indicators of Brazilian municipalities. By the map, from the right of the figure, it is possible to see that evasion areas expand in 2010. In the Northeast, evasion municipalities are higher in 2010, when compared to the year 2000. Besides this region, the number of municipalities in the South considered as areas of population loss also expands.

The number of municipalities classified with at least some migratory turnover rises throughout the national territory. The intensity and reduction of migration flows, as well as the length of stay in cities, given the labor market turnover, may explain the pattern of recent migration. Thus, a high turnover in the labor market such as that presented by Brazil can accentuate the spatial movement of the population.

4 DESCRIPTIVE STATISTICS AND THE PANEL DATA MODEL

In this section, the approach will be made from the descriptive statistics of the variables used in this study, as well as the results of the panel data models, by the fixed effects approach with transformation *within* (Wooldridge, 2006, p. 300). Here, it is sought to explain the participation of in-migration and out-migration in Brazilian municipalities. Based on the assumption that there are many socioeconomic and demographic determinants of the migratory dynamics in the country, variables are used in multiple dimensions. The objective is to establish a cause and effect relationship between socioeconomic factors (cause) and migration (effect) in Brazilian municipalities.

4.1 Descriptive statistics of variables

From the descriptive statistics, it is possible to confirm that the household infrastructure variables (households with electricity supply, water supply, sewage and garbage collection) show that there was a substantial improvement between the first and last year under analysis. In 2010, electricity supply approached universalization and 97% of Brazilian households responded to electricity, according to the 2010 Census. In addition, water supply, garbage collection and sanitary sewage had a higher coverage rate in 2010, compared to the year 2000, although there are still substantial possibilities for improvement in the supply of such basic services (table 1).

TABLE 1
Description of variables and municipality average values in Brazil (2000 and 2010)

Variables	Year		Description of the variables – participation by municipality 2000 and 2010
	2000	2010	
<i>p_imigrante</i>	0.11	0.08	Percentage of households in the destination headed by in-migration.
<i>p_emigrante</i>	0.10	0.08	Percentage of households at the origin with responsible out-migration.
<i>p_agua</i>	0.58	0.69	Percentage of households with water supply service by general network.
<i>p_esgoto</i>	0.75	0.86	Percentage of households with sewage system by general sewage or rainwater network; septic tank and rudimentary tank.
<i>p_lixo</i>	0.53	0.70	Percentage of households with garbage collection per cleaning service or placed in a cleaning service bucket.
<i>p_energia</i>	0.86	0.97	Percentage of households with electricity supply by distribution company or other sources.
<i>p_alfabetizado</i>	0.72	0.79	Percentage of households in which the responsible person claimed to be literate.
<i>p_segrau</i>	0.08	0.15	Percentage of households in which the responsible person responded to have completed high school.
<i>p_superior</i>	0.02	0.05	Percentage of households in which the responsible person has completed higher education.
<i>p_ocupado</i>	0.79	0.74	Percentage of households in which the reference person responded that was working in the reference week of the survey.
<i>p_previdencia</i>	0.11	0.12	Percentage of households in which the head of household responded to pay social security contributions in the main job or in another job.
<i>p_ICS</i>	0.30	0.32	Percentage of households in which the person responsible responded to be occupied in sectors of economic activities of ICS.
<i>p_pobretrab</i>	0.68	0.57	Percentage of households in which the responsible declared to be poor at work. That is, to have remuneration lower than 1/2 of minimum wage.

Source: Demographic Census data from Brazil (2000 and 2010).

Authors' elaboration.

With regard to human capital (formal education of heads of households), the participation of literate people with high school and complete secondary education also improves in the period between the censuses. However, the participation of heads of Brazilian households with complete higher education is still low, being 2% in 2000 and 5% in 2010.

In the labor market (occupancy rate, social security coverage of the employed, engaged in activities of commerce, industry and services and labor poverty), the relative participation of employed persons is reduced from the first to the last year, rising from 79% to 74% of the workforce. However, poverty at work is reduced from 68% to 57%, from the first to the last year, which can result from higher income through conditional income transfer programs. In addition, there is a slight increase in the participation of employed persons in ICS, to the detriment of other

activities, as well as the participation of employed persons insured by official social security institutes (table 1).

4.2 Results and discussions: the panel data model

By analyzing the panel data model, the results indicate that if the participation of high school household heads increases by one percentage point (p.p.) in relation to those with first degree, a reduction of 0.117 p.p. in immigration in the municipality. Moreover, if the percentage of heads of households with higher education is varied by one p.p., a reduction of 0.04 p.p. in the participation of in-migration in the municipality is expected. The variation of one p.p. in the *p_alfabetizado* participation implies a variation of 0.066 p.p. in the participation of in-migration in the municipalities. The level of literacy in the municipality is directly related to the participation of in-migration at municipality levels. The results suggest that immigration is associated with locations where there is a greater participation of people enrolled.⁸ But in those places with high qualification rates, the share of immigration is lower. Probably because migrants would have a more difficult insertion in the labor market. The coefficient of the *p_superior* variable is less expressive, but signals in the same direction as the high school.

With regard to the immigrant, it is worth mentioning that all variables of household infrastructure, except *p_esgoto*, were statistically significant at levels of 0.001, 0.01 and 0.05. In home infrastructure variables, *p_agua* and *p_lixo* showed positive signs and *p_energia* negative sign to explain the participation of in-migration by municipality of the country. In other words, the municipalities with the highest coverage of basic services are also those that tend to attract in-migration.

Furthermore, the variation of a percentage point in the participation *p_ocupado* implies in the variation of 0.06 p.p. in the participation of in-migration in the Brazilian municipalities. It can be interpreted that the occupation rate at the municipality levels is directly related to the participation of in-migration. Thus, the local labor market is important indicator of the migratory dynamics of the municipality. As the occupation rate increases by one p.p., there is a tendency to register a positive variation in the participation of in-migration at the municipality level.

The economic activity sector also presented a positive signal for the coefficient. This means that the variation of one percentage point in occupied *p_ICS* implies a variation of 0.06 p.p. in the participation of intermunicipality in-migration in Brazil. That is, there is a cause and effect relationship between the employed

8. It is necessary to note that there may be an endogenous effect on the variable education and immigration. Since it is not possible, from the fixed-date migration, to work with this lagged variable, it is used here as a way of inferring by a cause-effect relationship that may be questionable.

in these sectors of activities (cause) and the participation of in-migration (effect) that are directly related. Otherwise, the variation of one point in the poverty in the work $p_pobretrab$ implies in the reduction of the participation of in-migration in the municipalities in 0.02 p.p. That is, there is a positive correlation between labor market indicators and in-migrant participation, as well as a negative correlation between labor poverty and intermunicipality in-migration in the country.

TABLE 2
Panel data model for participation of migrants and in-migration in Brazilian municipalities (2000 and 2010)

Variables	2000 and 2010		2000 and 2010	
	In-migrants		Out-migrants	
	Estimate	$Pr > t $	Estimate	$Pr > t $
Intercept	-0.015	0.600	0.035	0.267
p_agua	0.019	0.001***	0.014	0.007**
p_esgoto	-0.006	0,264	0.020	0.002**
p_lixo	0.009	0.064*	-0.003	0.646
$p_energia$	-0.015	0.011*	0.048	0.000***
$p_alfabetizado$	0.066	0.000***	-0.014	0.344
p_segrau	-0.117	0.000***	0.031	0.105
$p_superior$	-0.037	0.315	-0.116	0.004**
$p_ocupado$	0.060	0.000***	-0.008	0.358
$p_previdencia$	-0.012	0.083	0.005	0.483
p_ICS	0.055	0.000***	-0.034	0.000***
$p_pobretrab$	-0.022	0.008**	0.058	0.000***

Source: Demographic Census data from Brazil (2000 and 2010).

Authors' elaboration.

Obs.: Significance – 0 '***' 0.001 '***' 0.01 '**' 0.05', 0.1 ' ' 1.

Regarding the participation of out-migration, the data show that, of the 11 variables used, 5 did not present statistical significance. However, they were not removed from the model so as not to cause specification bias. That is, the difficulty of predicting the factors of people evasion in the municipalities of the country is greater. Besides, the other variables presented significance at 0.001% and 0.01%.

The variable $p_pobretrab$ showed positive sign. That is, the variation of a percentage point in the poverty in the work can imply in the variation of 0.06 p.p. in the participation of out-migration. With this, the coefficients indicate that the greater the municipality poverty, the greater the records of people who

emigrated from that municipality. In this way, it is possible to suppose that there is a direct relation between poverty and people evasion in the Brazilian municipalities. In addition, the variables: complete higher education; employed, or working in the commerce or services industry, had negative signals, signaling that negative variations of a point in these variables imply negative variations, in percentage points, of the participation of out-migration in the municipalities of the country. That is, the exit of people from the municipalities may be related to the low participation of the employed or workers in the sectors of ICS, as well as people with higher level of schooling.

5 FINAL CONSIDERATIONS

The objective of this paper was to make a study about the socioeconomic factors that impact on the Brazilian intermunicipality migratory dynamics. The Brazilian Demographic Census of the years 2000 and 2010 show changes in the socioeconomic and migratory dynamics in the country, which is the motivation for an analysis of the intermunicipality migratory dynamics and the use of fixed date migration and heads of households as the chosen form for analysis.

By the MEI it is possible to notice that there is a substantial change in relation to the evasive areas between 2000 and 2010, as well as between the areas of migratory turnover. The economically less dynamic regions are responsible for the higher incidence of population evasion, and the labor market has support in the areas of migratory turnover. As regards areas of population retention, they are concentrated in economically developed and economically important regions, both in 2000 and 2010.

It is possible to perceive that there is population concentration coming from the migrations in the Brazilian metropolitan regions, as well as in the capitals of the states and in the Federal District. They are, therefore, areas of population attraction, given the best indexes of job opportunity, arising from the more accentuated economic dynamics of these municipalities.

The results of the panel data model showed that the percentage of in-migration' participation is positively affected by the occupation and work in the industry, services and commerce, being these variables with higher coefficients and with greater explanation powers in the model. That is, it is possible to affirm that there is a direct relationship between labor market performance, particularly in the more dynamic activities related to the ICS sectors, with the participation of in-migration in the municipalities that present the best indicators. The percentage of out-migration' participation in the municipalities can be explained mainly by the negative variation in the total occupation and by sector.

The results found, point to several determinant fronts of the Brazilian intermunicipality migratory dynamics. It was not possible to verify the isolated effect of the variables on the migratory dynamics of the heads of households. This is related to issues of opportunity for migrants and the selection is made by criteria differentiated between them. As next steps for future research, it is recommended to decompose the factors that influence the intermunicipality immigration dynamics of heads of households.

REFERENCES

- ARAÚJO, T. B. de. Nordeste, Nordestes: que Nordeste? In: ARAÚJO, T. B. de. **Ensaio sobre o desenvolvimento brasileiro: heranças e urgências**. Rio de Janeiro: Fase, 2000. p. 165-196.
- AYDOS, M. R. Migrações internas no Brasil contemporâneo: reflexões teóricas e analíticas dos principais fluxos interestaduais 1930-2008. In: SEMANA DE PLANEJAMENTO URBANO E REGIONAL, 16., 2010, Rio de Janeiro. **Proceedings...** Rio de Janeiro: Ippur/UFRJ, 2010.
- BORJAS, G. J. Self-selection and the earnings of immigrants. **The American Economic Review**, v. 77, n. 4, p. 531-553, 1987.
- BORJAS, G. J. (Ed.). **Labor economics**. New York: McGraw-Hill, 1996.
- BORJAS, G. J. **The economic progress of immigrants**. Cambridge, United States: NBER, Apr. 1998. (Working Papers, n. 6506).
- BRITO, F. O deslocamento da população brasileira para as metrópoles. **Estudos Avançados**, v. 20, n. 57, p. 221-236, 2006.
- CAMARANO, A. A.; ABRAMOVAY, R. Êxodo rural, envelhecimento e masculinização no Brasil: panorama dos últimos cinquenta anos. **Revista Brasileira de Estudos de População**, Rio de Janeiro, v. 15, n. 2, p. 45-65, 1998.
- CAMBOTA, J. N.; PONTES, P. A. O papel da migração interna na convergência de rendimentos do trabalho no Brasil, no período de 1994 a 2009. **Revista EconomiA**, Brasília, v. 13, n. 1, p. 131-147, Jan.-Apr. 2012.
- CANÇADO, R. P. Migrações e convergência no Brasil: 1960-1991. **Revista Brasileira de Estudos de População**, Rio de Janeiro v. 53, n. 2, p. 211-236, Apr.-June 1999.
- CHISWICK, B. R. The effect of Americanization on the earnings of foreign-born man. **Journal of Political Economy**, v. 86, n. 5, p. 897-921, 1978.
- CHISWICK, B. R. Are immigrants favorably self-selected? **The American Economic Review**, v. 89, n. 2, p. 181-185, 1999.

DINIZ, C. C. **A questão regional e as políticas governamentais no Brasil**. Belo Horizonte: Cedeplar/UFMG, 2001. (Texto para Discussão, n. 159).

FERREIRA, A. H. B.; DINIZ, C. C. Convergência entre as rendas *per capita* no Brasil. **Revista de Economia Política**, v. 15, n. 4, p. 497-516, 1995.

FREGUGLIA, R. da S.; MENEZES-FILHO, N. A. Inter-regional wage differentials with individual heterogeneity: evidence from Brazil. **The Annals of Regional Science**, v. 49, n. 1, p. 17-34, 2012.

FREGUGLIA, R. da S.; MENEZES-FILHO, N. A.; SOUZA, D. B. de. Diferenciais salariais inter-regionais, interindustriais e efeitos fixos individuais: uma análise a partir de Minas Gerais. **Estudos Econômicos**, São Paulo, v. 37, n. 1, p. 129-150, Jan.-Mar. 2007.

GAMA, L. C. D.; MACHADO, A. F. Migração e rendimentos no Brasil: análise dos fatores associados no período intercensitário 2000-2010. **Estudos Avançados**, São Paulo, v. 28, n. 81, p. 155-174, 2014.

GUIMARÃES, E. N.; LEME, H. J. de C. Caracterização histórica e configuração espacial da estrutura produtiva do Centro-Oeste. In: HOGAN, D. J. et al. (Ed.). **Migração e ambiente no Centro-Oeste**. Campinas: Nepo/Unicamp; Pronex, 2002. p. 17-85.

GUIMARÃES NETO, L. Desigualdade e políticas regionais no Brasil: caminhos e descaminhos. **Planejamento e Políticas Públicas**, n. 15, p. 41-99, June 1997.

JAKOB, A. A. E.; JAKOB, R. de O. S. E. A migração recente na Amazônia brasileira: como analisar? In: CONGRESSO DA ASSOCIAÇÃO LATINO AMERICANA DE SOCIOLOGIA, 30., 2015, San José, Costa Rica. **Proceedings...** San José: Alas, 2015.

JUSTO, W. R. et al. Os determinantes da migração e da migração de retorno intermunicipal no Brasil. In: ENCONTRO NACIONAL DE ECONOMIA, 38., 2010, Salvador, Bahia. **Proceedings...** Niterói: Anepc, 2010.

JUTTEL, L. P. Norte e Centro-Oeste, novos pólos de migração. **Ciência e Cultura**, São Paulo, v. 59, n. 4, p. 10-11, 2007.

LEE, E. S. Uma teoria sobre a migração. In: MOURA, H. A. de. (Ed.). **Migração interna: textos selecionados**. Fortaleza: BNB, 1980. p. 89-114.

LIMA, E. E. C.; BRAGA, F. G. Da rotatividade migratória à baixa migração: uma análise dos padrões da mobilidade populacional no Brasil de 1995-2000. **Revista Brasileira de Estudos da População**, Rio de Janeiro, v. 30, n. 1, p. 57-75, Jan.-June 2013.

LIMA, L. C.; VALE, A. L. F. Migração e mudança social: influência do migrante do Sertão Nordestino no Norte do Brasil. **Scripta Nova**, v. 5, n. 94, 2001.

MAIA, A. G. Estrutura de ocupações e distribuição de rendimentos: uma análise da experiência brasileira nos anos 2000. **Revista de Economia Contemporânea**, Rio de Janeiro, v. 17, n. 2, p. 276-301, May-Aug. 2013.

MAIA, A. G.; QUADROS, W. J. de. Tipologia municipal de classes socioocupacionais: uma nova dimensão para análise das desigualdades territoriais no Brasil. **Revista de Economia e Sociologia Rural**, v. 47, n. 2, p. 389-418, 2009.

MARTINE, G. Migração e metropolização. **São Paulo em Perspectiva**, São Paulo, v. 1, n. 2, p. 28-31, July-Sept. 1987.

MARTINE, G. Brazil. In: NAM, C. B.; SEROW, W. J.; SLY, D. F. (Ed.). **International handbook of internal migration**. Santa Barbara: Greenwood, 1990.

MARTINE, G.; CAMARGO, L. Crescimento e distribuição da população brasileira: tendências recentes. **Revista Brasileira de Estudos de População**, Campinas, v. 1, n. 1/2, p. 99-143, Jan.-Dec. 1984.

MATA, D. da. et al. Quais características das cidades determinam a atração de migrantes qualificados? **Revista Econômica do Nordeste**, Fortaleza, v. 38, n. 3, p. 502-514, 2007.

MENEZES, O. **Migrações: III Recenseamento Geral da População e da Habitação de 2001**. São Tomé e Príncipe: Instituto Nacional de Estatística, 2003.

MENEZES, T. A.; FERREIRA-JÚNIOR, D. **Migração e convergência de renda**. São Paulo: Nereus, 2003. (Texto para Discussão, n. 13).

MYRDAL, G. (Ed.). **Economic theory and the underdeveloped regions**. London: Duckworth Books, 1956.

NETTO JÚNIOR, J. L. da S.; PORTO JUNIOR, S. da S.; FIGUEIREDO, E. A. de. Migração e distribuição de capital humano no Brasil: mobilidade intergeracional educacional e intrageracional de renda. **Revista Econômica do Nordeste**, Fortaleza, v. 39, n. 4, p. 404-427, Oct.-Dec. 2008.

OLIVEIRA, K. F.; JANNUZZI, P. de M. Motivos para migração no Brasil e retorno ao Nordeste: padrões etários, por sexo e origem/destino. **São Paulo em Perspectiva**, v. 19, n. 4, p. 134-143, 2005.

OLIVEIRA, L. A. P.; OLIVEIRA, A. T. R. de. (Ed.). **Reflexões sobre os deslocamentos populacionais no Brasil**. 1. ed. Rio de Janeiro: IBGE, 2011. v. 1, 103 p.

PACHECO, C. A. **Fragmentação da nação**. Campinas: Unicamp/IE, 1998.

PACHECO, C. A.; PATARRA, N. Movimentos migratórios nos anos 80: novos padrões? In: ENCONTRO NACIONAL SOBRE MIGRAÇÃO, 1., 1997, Curitiba, Paraná. **Proceedings...** Curitiba: Ipardes/FNUAP, 1997.

QUEIROZ, S. N.; BAENINGER, R. Migração de retorno: o caso recente das migrações cearenses. **Revista Econômica do Nordeste**, v. 44, n. 4, p. 833-850, 2013.

QUEIROZ, S. N.; SANTOS, J. M. Saldos migratórios: uma análise por estados e regiões do Brasil (1986-2006). **Revista Econômica do Nordeste**, v. 42, n. 2, p. 309-332, 2011.

RAMALHO, H. M. de B. **Migração, seleção e desigualdade de renda**: evidências para o Brasil metropolitano a partir do Censo Demográfico 2000. 2005. Dissertação (Mestrado) – Universidade Federal da Paraíba, João Pessoa, Paraíba, 2005.

SANTOS JUNIOR, E. da R. **Migração e seleção**: o caso do Brasil. 2002. 56 f. Dissertação (Mestrado) – Fundação Getúlio Vargas, Rio de Janeiro, 2002.

SANTOS, L. M.; MOREIRA, I. T. Condições do mercado de trabalho no setor de serviços nordestinos. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 15., 2006, Caxambu, Minas Gerais. **Proceedings...** 2006.

SAWYER, D. R. Fluxo e refluxo da fronteira agrícola no Brasil: ensaio da interpretação estrutural e espacial. **Revista Brasileira de Estudos da População**, v. 1, n. 2, p. 3-34, Dec. 1984.

SILVA FILHO, L. A.; SILVA, F. J. F.; QUEIROZ, S. N. Nordeste industrial: a fragmentação territorial de uma região periférica. **Revista Econômica do Nordeste**, Fortaleza, v. 46, n. 2, p. 9-24, Apr.-June 2015.

SILVA, Y. C. L.; SILVA FILHO, L. A.; CAVALCANTI, D. M. Migração, seleção e diferenciais de renda na região Norte do Brasil em 2010. In: ENCONTRO DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS DA POPULAÇÃO, 20., 2016, Foz do Iguaçu, Paraná. **Proceedings...** 2016.

SINGER, P. Migrações internas: considerações teóricas sobre o seu estudo. In: MOURA, H. A. de. (Ed.). **Migração interna**: textos selecionados. Fortaleza: BNB, 1980. p. 211-244.

SJAASTAD, L. A. Os custos e os retornos da migração. In: MOURA, H. A. de. (Ed.). **Migração interna**: textos selecionados. Fortaleza: BNB, 1980. p. 115-143.

TAYLOR, E. J. The new economics of labour migration and the role of remittances in the migration process. **International Migration**, v. 37, n. 1, p. 63-88, 1999.

TODARO, M. P. A migração de mão de obra e o desemprego urbano em países desenvolvidos. In: MOURA, H. A. de. (Ed.). **Migração interna**: textos selecionados. Fortaleza: BNB, 1980. p. 145-172.

UN – UNITED NATIONS. **Methods of measuring internal migrations**. New York: UN, 1970. (Population Studies, n. 47).

VASCONCELLOS, I. R. P.; RIGOTTI, J. I. R. Migrações entre os municípios brasileiros, a partir das informações dos Censos Demográficos de 1991 e 2000. In: ENCONTRO NACIONAL SOBRE MIGRAÇÕES, 4., 2005, Rio de Janeiro. **Proceedings...** Campinas: Abep, 2005.

WOOD, C. H. Equilibrium and historical-structural perspectives on migration. **The International Migration Review**, v. 16, n. 2, p. 298-319, 1982.

WOOLDRIDGE, J. M. **Introdução à econometria: uma abordagem moderna**. São Paulo: Pioneira Thomson Learning, 2006.

Originais submetidos em: jun. 2020.

Última versão recebida em: jul. 2021.

Aprovada em: jul. 2021.

"OS AMIGOS DOS MEUS AMIGOS SÃO MEUS AMIGOS?": AS CONSEQUÊNCIAS DO CAPITAL SOCIAL PARA AS DIFERENÇAS SALARIAIS ENTRE HOMENS E MULHERES NO BRASIL¹

Fábio Lúcio Rodrigues²

Mércia Santos da Cruz³

Wallace Patrick Santos de Farias de Souza⁴

O objetivo deste artigo constitui-se em identificar variáveis representativas para o capital social (CS) no nível do indivíduo, verificando seus efeitos sobre a composição dos salários e a consequente diferença salarial por sexo e quantis de rendimentos. A Pesquisa Nacional de Saúde (PNS) é utilizada para executar diferentes estratégias econométricas, com a finalidade de identificar o impacto do capital social sobre os salários por sexo. Os resultados apontam para um claro efeito do estoque de capital social sobre os salários em geral, mas com impacto superior nos rendimentos femininos, tanto na média quanto nos salários das mulheres que estão nas faixas salariais mais baixas.

Palavras-chave: capital social; decomposição salarial; diferenças salariais; rendimentos do trabalho.

"THE FRIENDS OF MY FRIENDS ARE THEY MY FRIENDS?": THE CONSEQUENCES OF SOCIAL CAPITAL FOR WAGE DIFFERENCES BETWEEN MEN AND WOMEN IN BRAZIL

The purpose of this paper is to identify representative variables for social capital at the level of the individual, verifying its effects on the composition of wages and the consequent wage difference by sex and incomes quantiles. The National Health Survey is used to execute different econometric strategies in order to identify the impact of social capital on wages by sex. The results point to a clear effect of the stock of social capital on wages in general, but with a higher impact on female incomes, both in the average and in the salaries of women who are in the lowest salary ranges.

Keywords: social capital; wage decomposition; wage differences; labor income.

JEL: J00; J16; J31.

1 INTRODUÇÃO

O período iniciado em meados da década de 1970 trouxe um crescente protagonismo das mulheres no contexto do mercado de trabalho mundial, principalmente nos países de tradição democrática. Características tipicamente femininas, como a capacidade de trabalho em equipe (Capraro, 2018), a persuasão argumentativa (Andrews, 2006) e a maior capacidade de cooperação (Molina *et al.*, 2013), têm levado as mulheres profissionais a ganharem cada vez mais espaços no ambiente

1. DOI: <http://dx.doi.org/10.38116/ppe52n3art4>

2. Professor no Departamento de Economia e no Programa de Pós-Graduação em Economia da Universidade do Estado do Rio Grande do Norte (PPE/UERN). *E-mail:* prof.fabiolucio@gmail.com.

3. Professora no Departamento de Economia e no Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba (PPGE/UFPB). *E-mail:* mercia_sc@hotmail.com.

4. Professor no Departamento de Economia e no PPGE/UFPB. *E-mail:* wpsfarias@gmail.com.

laboral (Madalozzo, 2010; Klasen *et al.*, 2019), trazendo à tona, por consequência, a necessidade de novas discussões das diferenças salariais existentes entre os gêneros ao longo dos anos.

Segundo Araújo e Ribeiro (2002), a população economicamente ativa (PEA) feminina cresceu 260% entre 1970 e 1990 no Brasil, enquanto a participação dos homens se elevou em apenas 73% nesse período e o rendimento médio das mulheres representava apenas 41% do masculino. No entanto, apesar da persistência das diferenças salariais das mulheres no que concerne aos homens, esse *gap* remuneratório vem reduzindo-se consideravelmente desde a década de 1970. Dados do Cadastro Central de Empresas (Cempre) do Instituto Brasileiro de Geografia e Estatística (IBGE), expostos em Caoli (2016), mostram que, a partir de 2009, o rendimento médio feminino já representava, em média, 80% do masculino. Diferença média que se mantém nos anos seguintes, conforme enfatizam Gomes e Souza (2019), com informações obtidas na Relação Anual de Informações Sociais (Rais) do Ministério do Trabalho e Emprego (MTE), concernente a 2016, e Paradella (2019), com dados relativos ao quarto trimestre de 2018 da Pesquisa Nacional por Amostra de Domicílios (PNAD) Contínua do IBGE.

É possível que tais reduções não tenham relações apenas com motivos específicos, como aumento da produtividade e escolaridade, ou aspectos demográficos, tal como a redução da fecundidade. Entretanto, há também relação com um contexto de desenvolvimento social mais amplo, no qual a mulher vem buscando, cada vez mais, seu protagonismo na sociedade na qual está inserida e, principalmente, na procura de igualdade de condições no ambiente das relações de trabalho, conforme pode ser visto em Chevalier (2007), Grove, Hussey e Jetter (2011), Madalozzo, Martins e Lico (2015) e Faustino, Araújo e Maia (2017), apesar da aversão feminina à competitividade com o sexo oposto, como evidenciado, por exemplo, em Niederle e Vesterlund (2007) e Morin (2015).

No ambiente corporativo, ainda ocorrem fortes discrepâncias de remuneração entre os sexos, principalmente na parte superior da distribuição de rendimentos; fenômeno conhecido como “teto de vidro”. É visível a percepção discriminatória sofrida pelas mulheres nas empresas, seja por estruturas remuneratórias tradicionalmente patriarcais, seja por mero sexismo organizacional arraigado em décadas de sub-representação feminina no topo da estrutura administrativa (Bertrand e Duflo, 2017; Bertrand, 2018).

Em paralelo, vários autores enfatizam o papel do capital social (CS) como elo fundamental para o desenvolvimento das instituições em economias de mercado, a partir da segunda metade do século XX. O trabalho seminal de Becker (1974) estabelece os pilares fundamentais de uma teoria que busca compreender os efeitos das interações sociais entre os agentes sobre suas decisões econômicas e,

posteriormente, desenvolve seus argumentos em Becker (1996), em que defende que as decisões e as preferências dos indivíduos não têm relação direta com suas necessidades, mas sim com o estoque de capital social acumulado.

Cremonese (2006) enfatiza que a edificação de um sentimento comunitário depende da ação recíproca entre as pessoas que nesta convivem, constatando que redes de interação social são fundamentais para a ampliação do capital social. Por sua vez, Pollán (2013) entende que laços sociais cada vez mais intensos necessitam de tempo e esforço para serem estabelecidos, e que o relacionamento direto em grupos homogêneos ao longo da vida das pessoas é a característica básica do desenvolvimento do capital social.

Apesar de uma grande gama de estudos que incorporam a teoria do capital social para explicações as mais diversas acerca do comportamento e do desenvolvimento de governos e instituições, ainda são escassas as aplicações empíricas que estabeleçam medidas de capital social no nível dos indivíduos e, principalmente, que estudem as influências que o desenvolvimento dos laços sociais acarreta sobre as pessoas e suas relações de trabalho.

Nesse contexto, o objetivo central deste estudo constitui-se em identificar variáveis representativas para o capital social no nível do indivíduo, ao verificar seus efeitos sobre a composição dos salários e a consequente diferença salarial por sexo e quantis de rendimentos. Para atingir o objetivo proposto, são executadas várias aplicações econométricas, a fim de verificar a persistência dos resultados.

É necessário considerar, também, que as condições de saúde dos trabalhadores são fundamentais na sua oferta de trabalho e, conseqüentemente, na composição dos seus rendimentos, principalmente em relação às mulheres (Cruz e Irfi, 2019). Fatores como morbidades adquiridas (Glaser e Grundy, 1997), violência doméstica (Aizer, 2007), condições de prevenção (Rodrigues, Cruz e Paixão, 2015), saúde na terceira idade (Corrêa, Queiroz e Fazito, 2016), estrutura familiar (Cruz, Araujo e Paixão, 2018) e diferenças entre raça e gênero (Treadwell, 2019) são imprescindíveis para a compreensão da composição salarial e dos hiatos de rendimentos entre os sexos. A pandemia causada pelo novo coronavírus e os efeitos devastadores provocados pela covid-19 nas famílias também afetam diretamente a oferta de mão de obra e os rendimentos dos trabalhadores, principalmente em grupos tradicionalmente marginalizados (Gausman e Langer, 2020; Chandler *et al.*, 2021). Portanto, de forma complementar, e apesar de não ser o foco principal deste trabalho, é realizada uma análise sobre os impactos das condições de saúde dos indivíduos na composição dos rendimentos.

Além desta introdução, este trabalho conta, também, na seção 2, com revisão de literatura, assim como apresenta a evolução do conceito de capital social e aborda as alternativas de mensuração que serão utilizadas nesta pesquisa.

A seção 3 traz um detalhamento da estratégia empírica utilizada neste ensaio e a base de dados relacionada. Na seção 4, são discutidos os resultados empíricos. A seção 5 apresenta as considerações finais.

2 REVISÃO DE LITERATURA

O conceito de capital social abrange um intenso aspecto de interdisciplinaridade, na medida em que integra a sociologia, a psicologia, a antropologia, bem como a economia, e combina várias ideias, incluindo-se as tradições culturais, a civilidade, o envolvimento cívico e a coesão social. Estudos empíricos sugerem que o capital social traz benefícios consideráveis para uma série de resultados econômicos e sociológicos (Claridge, 2004). Nesse mote, existem três tipos de capital que interessam diretamente aos estudos econômicos: capital físico, humano e social.

Para Coleman (1990), o capital social representa os recursos, reais ou potenciais, obtidos a partir dos relacionamentos existentes entre as pessoas e suas redes de engajamento. Nesse contexto, o capital social é um bem público e, portanto, depende da disposição dos membros da comunidade em evitar o *free riding* – ou seja, aquele agente oportunista não disposto a arcar com os custos inerentes à geração do capital social. Dessa forma, normas bem estabelecidas, confiança mútua entre as pessoas envolvidas e a possibilidade de aplicação de sanções àqueles que não respeitarem o *status quo* social no qual estão inseridos tornam-se importantes fatores para sustentar esse ativo coletivo.

Putnam, Leonardi e Nonetti (1993) retomam o argumento central de Coleman (1990) e determinaram características para o capital social inerentes a um bem público, na medida em que o nível do capital social estabelecido em um grupo de indivíduos pode ser expandido em nível agregado, criando externalidades positivas para toda a sociedade.

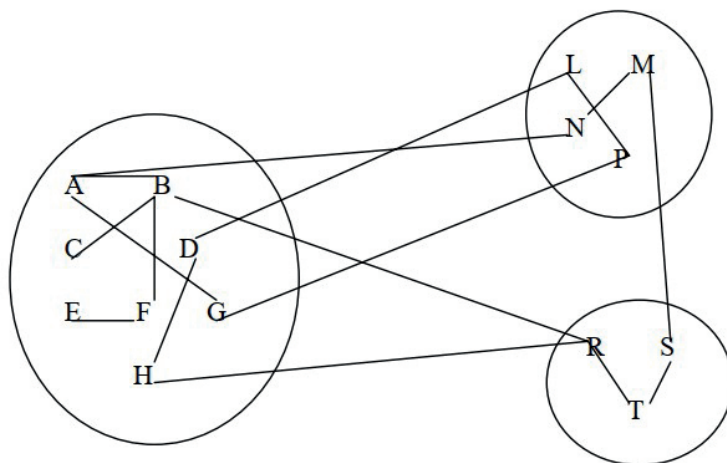
Ademais, conforme complementado por Nunes e Pizzi (2014), o capital social refere-se a uma rede de relações que contribuem para os bens compartilhados em sociedade. As relações recíprocas, a confiança mútua e as redes de parceria entre instituições democráticas e comunidade permitem fortalecer a solidariedade comunitária.

Partindo-se dessa análise, é perceptível que os benefícios do capital social advêm de duas formas distintas, mas integradas. Os laços sociais criados em uma mesma rede de relacionamento são, via de regra, mais intensos e geram benefícios a partir das relações entre seus indivíduos diretamente. Esses benefícios diretos são conhecidos como capital social de vínculo, ou *bonding*, identificando as conexões mais próximas entre indivíduos com características comuns entre si; por exemplo, mesma família, mesmo grupo étnico, culto religioso comum etc. De forma complementar, os laços existentes entre redes de relacionamento diferentes identificam interações mais distantes e mais fracas e levam ao que é chamado de capital social de

ponte, ou *bridging* (Andriani e Karyampas, 2009). A figura 1 descreve três círculos diferentes, que representam, cada um, uma rede de relacionamento distinta com membros diferentes pertencentes a cada círculo.

Na figura 1, as linhas identificam as conexões existentes entre dois ou mais membros, que podem interagir em seu próprio círculo de relacionamentos ou fora deste. As conexões entre dois membros pertencentes ao mesmo círculo – por exemplo, CB, LP ou TS – representam conexões *bonding*, enquanto conexões entre dois ou mais círculos de relacionamento diferentes – a exemplo de RB, RH ou AN – indicam conexões do tipo *bridging*.

FIGURA 1
Conexões *bridging* e *bonding* entre redes de relacionamento



Fonte: Andriani e Karyampas (2009).

Obs.: Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

No escopo da teoria econômica, Becker (1974) estabelece os pressupostos iniciais de uma teoria que permite analisar as interações entre os indivíduos e os comportamentos subjacentes promovidos pelo meio social em que convivem. Em seu artigo seminal, o autor sistematiza o conceito de *renda social* – definida como a soma da renda própria de uma pessoa e o valor monetário para ela das características relevantes dos outros indivíduos de seu meio social – e analisa os efeitos sobre esses gastos de mudanças em variadas fontes de renda e em diferentes preços. Entre suas principais contribuições, está uma abordagem abrangente acerca do comportamento da família como um agente econômico independente, que maximiza sua renda social e promove impactos diretos na mobilidade intergeracional dos indivíduos.

Complementarmente, Becker (1996) argumenta que as decisões e as preferências dos indivíduos têm relação intrínseca com seu capital social e pouco se relacionam com suas necessidades básicas. As decisões de consumo passadas, assim como as demais experiências pessoais adquiridas, são preservadas pelo indivíduo no que o autor chamou de função de utilidade expandida, na qual ocorre a interligação de suas preferências passadas e presentes, evidenciando que o conjunto das interações sociais das pessoas é fator fundamental para as decisões de consumo. Manski (2000) enfatiza que grande parte dos trabalhos da chamada economia social tem estudado a família como entidade maximizadora de utilidade única, com foco nas interações que ocorrem entre os membros dessa entidade. Por sua vez, há desenvolvimentos contemporâneos que lançam mão da teoria dos jogos não cooperativos para modelar famílias e agregados familiares, como grupos cujos membros podem ter objetivos diferentes (Becker, 1996; Manski, 2000).

Akerlof e Kranton (2000) propõem um modelo econômico de comportamento que incorpora o senso de identidade de uma pessoa e analisam como a identidade pode afetar as interações individuais e os resultados econômicos. Os autores concluem que a identidade tem efeito sobre o comportamento econômico por meio de quatro formas diferentes: modificando as recompensas de nossas próprias ações; alterando as recompensas das ações das outras pessoas; escolhendo diferentes identidades; e, por meio de alterações nas categorias sociais e condições comportamentais, afetando as preferências baseadas na identidade.

Apesar de não haver ampla evidência empírica sobre a conexão entre capital social e rendimentos do trabalho, a literatura especializada vem se debruçando cada vez mais na compreensão dessa relação. Borghans, Weel e Weinberg (2005) mostram que as interações interpessoais são importantes determinantes dos resultados do mercado de trabalho e apresentam indícios de que o aumento da importância dessas interações, entre o final dos anos 1970 e o início da década de 1990, ajuda a explicar os motivos de os salários femininos elevarem-se mais rapidamente no período.

Hermann e Kopasz (2011) analisam o efeito do capital social individual sobre a renda salarial de forma agregada em diversos países europeus e apresentam evidências de associação positiva entre capital social e rendimentos, tanto para a amostra conjunta quanto para a maioria dos países individualmente. Por sua vez, Collischon e Eberl (2021) demonstram que a quantidade de indivíduos homens entre amigos mais próximos influencia o estoque de capital social e a manutenção das diferenças salariais entre homens e mulheres no Reino Unido. Harb e Rouhana (2020) também demonstram que o estoque de capital social tem impacto significativo sobre os rendimentos dos trabalhadores libaneses, além de corroborar com evidências empíricas o fenômeno conhecido por “teto de vidro”, descrito por Bertrand (2018).

No Brasil, a literatura empírica acerca dos efeitos do capital social ainda é escassa e predominantemente relacionada a aspectos institucionais. Trabalhos como os de Bilert *et al.* (2011) e Duque (2013) relacionam o estoque de capital social acumulado pela sociedade com aspectos do desenvolvimento sustentável. Por sua vez, Silva (2001) e Arraes (2002) mostram que condições institucionais, como corrupção, nível de liberdades civis e infraestrutura associativa, estão intimamente ligadas ao desempenho econômico. Ribeiro e Araújo (2018) analisam o impacto do capital social na renda e na redução da pobreza no Brasil e encontram evidências de que quanto menor a escala de renda, maior o impacto do capital social na geração de renda e, conseqüentemente, na redução da pobreza da sociedade. Buss e Pellegrini Filho (2007) e Vieira, Ribeiro e Zeller (2012) relacionam o estoque de capital social nacional a aspectos ligados a saúde coletiva e taxas de mortalidade.

Diante do contexto da falta de meios de mensuração direta do capital social e, mais especificamente, da relação deste com a composição salarial, a abordagem tratada neste trabalho visa identificar as conseqüências diretamente relacionadas às questões de composição de rendimentos e das diferenças salariais entre os sexos. A expectativa é que o capital social, por meio de suas redes de laços e confiança comunitária, reduza as diferenças salariais existentes entre homens e mulheres, precipuamente se considerarmos o desenvolvimento e o aumento da influência feminina em todos os tecidos sociais a partir da segunda metade do século XX.

A suposição de preponderância feminina acerca dos efeitos positivos do capital social pode ser construída e embasada em diversos estudos. Em uma visão conservadora, Coleman (1990) e Putnam (2000) defendem que as famílias são exemplos inatos de capital social, relacionando a família dita tradicional e nuclear com altos níveis de confiança social e engajamento cívico. Molyneux (2002) enfatiza que as mulheres são figuras centrais da formação de capital social e que governos estão ansiosos para mobilizar esse estoque de capital social para a construção de programas de redução da pobreza e desenvolvimento comunitário.

Para Bilac (2006), as mulheres são as principais encarregadas do trabalho de produção e acumulação do capital social por intermédio da construção e do fortalecimento das redes de engajamento, sejam familiares, sejam de amizades recíprocas, naturalmente baseadas no trabalho emotivo ou afetivo, próprio da essencialidade feminina. A autora também argumenta que a mobilização das redes de engajamentos e relacionamentos, operadas prioritariamente por mulheres, desempenha importante suporte econômico e social para a sobrevivência das famílias, sobretudo na América Latina.

No âmbito empresarial, Sappleton (2009) mostra que, quando o capital social é medido em termos de confiança, envolvimento comunitário e redes sociais construídas nas corporações, mulheres que administram essas empresas em

setores tradicionalmente femininos têm maior acúmulo de capital social. Isso evidencia contraste com aqueles indivíduos que trabalham em setores tradicionalmente masculinos.

Essencialmente, as mulheres são relacionadas a maiores habilidades de sociabilidade e, por isso, a uma maior capacidade de construir e manter laços sociais mais abrangentes. Nesse sentido, Kaztman (1999) defende que as mulheres são mais presentes em redes de fortes laços sociais, cuja estrutura de socialização requer relacionamentos mais densos e mais homogêneos, como em redes de parentesco e vizinhança. Figueiró *et al.* (2010), em estudo com universitários brasileiros, também encontram evidências de forte capacidade de sociabilidade das mulheres quando comparadas aos homens. Em síntese, a capacidade e habilidade femininas em desenvolver redes de relacionamento estruturalmente mais fortes e abrangentes creditam às mulheres, por hipótese, a preponderância dos efeitos positivos do capital social sobre os rendimentos do trabalho.

3 ESTRATÉGIA EMPÍRICA

Este estudo se propõe a identificar variáveis empíricas representativas para o capital social no nível do indivíduo, ao verificar a influência de tais medidas sobre a composição dos salários e a consequente diferença salarial por sexo em diferentes extratos salariais. A hipótese fundamental defendida reside no fato de que o capital social reduz as diferenças salariais de gênero, além de ser fator relevante para a melhoria na composição dos salários das mulheres em diferentes faixas de renda do trabalho.

3.1 Base de dados e tratamentos das variáveis

O banco de dados utilizado neste estudo é a Pesquisa Nacional de Saúde (PNS), realizada em 2013 em parceria formada pela Fundação Oswaldo Cruz (Fiocruz) e o IBGE. A justificativa para a utilização da PNS reside no fato de que, além de produzir dados em âmbito nacional sobre a situação da saúde, a pesquisa se propôs a levantar informações acerca dos estilos de vida da população brasileira. O questionário aplicado apresenta informações sobre aspectos da vida do indivíduo com a família e os amigos, bem como algumas atividades em grupo, além de identificar a frequência na qual o indivíduo participa das atividades relacionadas.

A PNS é composta por três questionários. Dois são respondidos por um residente do domicílio e observam questões sobre as características particulares daquele domicílio, bem como da situação socioeconômica e de saúde de todos os residentes. O terceiro questionário individual é obrigatoriamente respondido por um residente que tenha idade igual ou superior a 18 anos, selecionado entre os

adultos residentes no domicílio, e contempla informações acerca de morbidade e estilos de vida dos residentes naquele domicílio (Brasil, 2020).

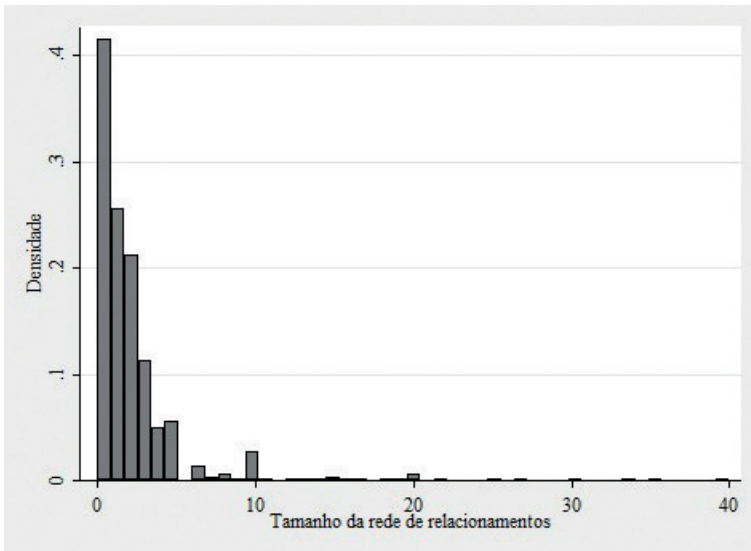
A amostra total da PNS corresponde a 205.546 indivíduos. Contudo, quando utilizado apenas o plano amostral de moradores com 18 anos ou mais de idade, os quais responderam aos módulos extras do questionário, a base de dados é reduzida para 64.653 indivíduos. Por fim, para evitar problemas com viés de seletividade em relação àqueles indivíduos que não possuem rendimento, são excluídos todos os indivíduos que informaram não ter renda monetária do trabalho na pesquisa, além daqueles que não tenham respondido à informação sobre jornada de trabalho semanal, levando a base de dados a 38.398 pessoas.

A variável dependente é o logaritmo do salário ajustado pelas horas trabalhadas ($\log w/h$), construída a partir das informações sobre rendimento mensal e horas efetivamente trabalhadas obtidas da base de dados da PNS. A fim de reduzir significativamente o desvio-padrão da *proxy* de rendimento, o salário por hora trabalhada considera como valor mínimo o equivalente à remuneração por hora relacionada ao salário-mínimo de 2013 (ano de realização da pesquisa) e é limitado a R\$ 150,00, o que corresponde a aproximadamente 99% da amostra. Considerando-se o escopo desse trabalho, o qual parte da suposição de que os laços de relacionamento entre os indivíduos reduzem as diferenças salariais entre homens e mulheres, foi definida como variável de interesse a existência de rede de relacionamento (amigos) por parte do indivíduo, denominada de *capital_social*. Também são definidas duas variáveis para fins de utilização nas estimações via regressões por variáveis instrumentais, *cs_esp* e *cs_relig*, definidas conforme o quadro 1.

A construção da variável *capital_social* foi realizada a partir das repostas da pergunta M15 do questionário da PNS 2013. Os indivíduos responderam à questão “com quantos amigos o(a) sr(a) se sente à vontade e pode falar sobre quase tudo (sem considerar os familiares ou parentes)?”, na qual foram consideradas as repostas com valores superiores a um como positivas para a existência de rede de relacionamentos pelo indivíduo. O gráfico 1 apresenta o histograma de frequências das repostas sobre a quantidade de amigos concernentes à pergunta M15.

No gráfico 1, é possível perceber que as redes de relacionamentos são constituídas, em sua imensa maioria, por até seis amigos, sendo essa frequência correspondente a 95% do tamanho da amostra, aproximadamente. Considerando-se que as repostas ao questionário estruturado não trazem nenhuma indicação sobre a intensidade das amizades, sendo esta, conforme sugerem Souza e Hutz (2008) e Schlösser (2020), de caráter fortemente abstrato e emocional, a incidência de rede de relacionamentos é atribuída admitindo-se que o indivíduo possua laços de amizades declarados, independentemente da quantidade.

GRÁFICO 1
Distribuição de amigos por indivíduo (2013)



Fonte: Questionário-base da PNS. Disponível em: <https://www.pns.icict.fiocruz.br/questionarios>. Acesso em: 14 mar. 2020. Elaboração dos autores.

Obs.: Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

As variáveis explicativas, por sua vez, são divididas em três grupos. O grupo 1 identifica as variáveis relacionadas às características pessoais, civis e familiares do indivíduo, além de *dummies* para as faixas de escolaridade (analfabeto, escolarizado, fundamental, médio e superior), em que os indivíduos analfabetos são o grupo-base de comparação.

QUADRO 1
Descrições das variáveis

Tipo	Variável	Descrição
dependente	log w/h	Logaritmo do salário ajustado pelas horas trabalhadas.
impacto	capital_social	Estoque de capital social. Assume valor um se o indivíduo possui rede de relacionamentos e zero caso contrário.
	cs_mulher	Variável de interação. Assume valor um se o indivíduo for do sexo feminino e possuir estoque de capital social e zero caso contrário.
instrumento	cs_esp	Frequência em atividades esportivas ou artísticas em grupo. Assume valor um se o indivíduo participou mais de uma vez por semana e zero caso contrário.
	cs_relig	Frequência em que compareceu a atividades religiosas. Assume valor um se o indivíduo participou mais de uma vez por semana e zero caso contrário.

(Continua)

(Continuação)

Tipo	Variável	Descrição
grupo 1	idade	Idade em anos.
	idade2	Idade ao quadrado.
	tempo_emp	Experiência profissional.
	tempo_emp2	Experiência profissional ao quadrado.
	mulher	Assume valor um se o indivíduo for do sexo feminino e zero caso contrário.
	branco	Assume valor um se a pessoa se autodeclara da raça branca e zero caso contrário.
	casado	Assume valor um se o indivíduo for casado e zero caso contrário.
	urbano	Assume valor um se o indivíduo reside na zona urbana e zero caso contrário.
	resp_dom	Assume valor um se o indivíduo é o responsável pelo domicílio e zero caso contrário.
	morad_dom	Quantidade de moradores residentes no domicílio.
	escolaridade	Assume valor um se o indivíduo estiver enquadrado na faixa de escolaridade específica (analfabeto, alfabetizado, fundamental, médio e superior) e zero caso contrário.
grupo 2	doen_cancer	Assume valor um se o indivíduo recebeu diagnóstico de câncer e zero caso contrário.
	doen_coracao	Assume valor um se o indivíduo recebeu diagnóstico de doença do coração e zero caso contrário.
	doen_cronica	Assume valor um se o indivíduo recebeu diagnóstico de doença crônica e zero caso contrário.
	viol_conh	Assume valor um se o indivíduo sofreu violência por pessoa conhecida e zero caso contrário.
	viol_desconh	Assume valor um se o indivíduo sofreu violência por pessoa desconhecida e zero caso contrário.
grupo 3	macroregioes	Assume valor um se o indivíduo é residente em uma macrorregião específica (Norte, Nordeste, Sul, Sudeste e Centro-Oeste) e zero caso contrário.

Fonte: Questionário-base da PNS. Disponível em: <https://www.pns.icict.fiocruz.br/questionarios>. Acesso em: 14 mar. 2020. Elaboração dos autores.

O grupo 2 reúne as variáveis relacionadas à ocorrência de diagnóstico de doenças e à violência sofrida pelos indivíduos. Tais variáveis visam identificar o risco à saúde dos indivíduos e sua influência na composição dos rendimentos. O grupo 3 inclui um conjunto de *dummies* que identificam a macrorregião do país na qual o indivíduo reside, sendo a região Sudeste a variável-base. O quadro 1 descreve as variáveis selecionadas.

No grupo 2, a identificação das variáveis que buscam identificar se o indivíduo sofreu violência mostra-se importante, visto que pesquisas sobre as sequelas na saúde mental e física das mulheres que sofreram violência perpetrada por parceiros sugerem a existência de sequelas em mais de 10% dos casos (Campbell, 2002⁵ *apud*

5. Campbell, J. C. Health consequences of intimate partner violence. *The Lancet*, v. 359, n. 9314, p. 1331-1336, 2002.

Cruz e Irffi, 2019). Todas as estimativas para o grupo 2 possuem expectativas de sinal negativo.

A PNS possui um desenho complexo de amostragem, com probabilidades desiguais de seleção dos indivíduos pesquisados. Dessa forma, para a análise dos dados da PNS e a posterior extrapolação dos dados amostrais para a população, são aplicados fatores de expansão ou pesos amostrais dos domicílios e de todos os seus moradores, assim como do morador escolhido para responder à terceira parte do questionário, com o objetivo de calibrar os dados, de forma a buscar a consistência das estimativas com os totais conhecidos (Szwarcwald *et al.*, 2014).

A tabela 1 apresenta as estatísticas descritivas da amostra da PNS e da expansão para a população.

TABELA 1
Análise descritiva das variáveis (2013)

Variáveis	Amostra				Amostra expandida	
	Média	Desvio-padrão	Mínimo	Máximo	Média	Erro-padrão
capital_social	0,6961	0,4599	0	1	0,6861	0,0045
cs_esp	0,1166	0,3209	0	1	0,1187	0,0032
cs_relig	0,1462	0,3533	0	1	0,1407	0,0033
cs_mulher	0,3201	0,4665	0	1	0,2967	0,0040
sal_hora	10,3879	13,8622	0,7179	150	10,2232	0,1405
idade	39,0038	12,2879	18	75	38,5714	0,1100
tempo_emp	8,6103	9,9333	0,0833	50	8,5560	0,0931
mulher	0,4542	0,4979	0	1	0,4249	0,0041
branco	0,4033	0,4906	0	1	0,4862	0,0048
casado	0,3955	0,890	0	1	0,4462	0,0047
urbano	0,8499	0,3572	0	1	0,8882	0,0031
resp_dom	0,6020	0,4895	0	1	0,5190	0,0048
morad_dom	3,2019	1,5672	1	16	3,5321	0,0168
analfabeto	0,0993	0,2991	0	1	0,0828	0,0024
alfabetizado	0,2068	0,4050	0	1	0,2094	0,0040
fundamental	0,1556	0,3625	0	1	0,1594	0,0035
medio	0,3687	0,4825	0	1	0,3818	0,0047
Superior	0,1695	0,3742	0	1	0,1666	0,0039
doen_cancer	0,0089	0,0941	0	1	0,0092	0,0008
doen_coracao	0,0205	0,1419	0	1	0,0261	0,0017
doen_cronica	0,3255	0,4686	0	1	0,3642	0,0045
viol_conh	0,0272	0,1627	0	1	0,0250	0,0013
viol_desconh	0,0379	0,1909	0	1	0,0385	0,0019

(Continua)

(Continuação)

Variáveis	Amostra				Amostra expandida	
	Média	Desvio-padrão	Mínimo	Máximo	Média	Erro-padrão
Norte	0,2165	0,4119	0	1	0,0722	0,0012
nordeste	0,2748	0,4464	0	1	0,2404	0,0034
sul	0,1288	0,3250	0	1	0,1601	0,0027
sudeste	0,2418	0,4282	0	1	0,4463	0,0043
centrooeste	0,1381	0,3450	0	1	0,0810	0,0014

Fonte: PNS 2013. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/saude/29540-2013-pesquisa-nacional-de-saude.html?edicao=9177&t=microdados>. Acesso em: 12 mar. 2020.

Elaboração dos autores.

3.2 Modelos empíricos e métodos de estimação

As diferenças entre os rendimentos auferidos por homens e mulheres é um assunto amplamente discutido desde a revolução feminista, observada a partir do final dos anos 1960. A hipótese neoclássica de que a oferta e a demanda por mão de obra determinavam os salários no mercado de trabalho não mais respondia ao fato de que os trabalhadores, ao possuírem níveis diferenciados de capital humano, eram remunerados de forma desigual.

Assim, Shultz (1961) e Becker (1962), ao constatarem que a mão de obra era heterogênea no mercado de trabalho e que isso se relacionava fortemente com a determinação dos salários, sintetizaram suas premissas acerca dos determinantes dos salários na chamada teoria do capital humano. Elas partem do princípio de que a educação e a experiência profissional dos indivíduos no mercado de trabalho possuem um efeito positivo em sua renda, supondo que, à medida que o volume de capital humano se eleve, a produtividade também crescerá, aumentando os rendimentos do trabalho.

Nesse interim, o trabalho lança mão de estratégia empírica com diferentes aplicações econométricas, não necessariamente concorrentes entre si, com o objetivo único de testar a hipótese básica formulada. Os modelos utilizados levam em consideração a decomposição dos efeitos do capital social por sexo, assim como a desagregação por grupo de sexo. Também admitem possíveis efeitos gerados pela endogeneidade entre as variáveis da especificação.

Mincer (1974) ampliou os estudos de Schultz (1961) e Becker (1962) sobre a influência do capital humano para a composição salarial dos indivíduos e desenvolveu uma abordagem em que é possível estimar os retornos marginais do investimento em capital humano sobre os rendimentos dos trabalhadores. A equação de salários minceriana pode ser representada formalmente da seguinte forma:

$$\ln w_{ki} = X'_{ki} \beta_k + \varepsilon_{ki}, \quad (1)$$

em que $\ln w$ é o logaritmo da renda ajustada pelas horas de trabalho, X' são vetores de características observáveis de cada i indivíduos, relacionadas ao seu capital humano, além de características de controle, e ε é o termo de erro estocástico.

Inicialmente, são obtidas as estimativas com mínimos quadrados ordinários (MQO). Contudo, a utilização de MQO não permite dizer qual parcela da diferença dos rendimentos se deve a características individuais presentes entre homens e mulheres. Assim, faz-se necessária uma metodologia que identifique os componentes dos movimentos dos diferenciais ao longo da curva de rendimento. O método de decomposição de diferenciais por meio de uma análise contrafactual foi proposto por Oaxaca (1973) e desenvolvido por Blinder (1973), ficando conhecido na literatura como método de Oaxaca-Blinder (O-B) de decomposição de rendimentos.

Esse método de decomposição divide o diferencial de salários entre dois grupos, sendo uma parte explicada pelas características de produtividade relacionadas ao capital humano e outra parte residual, não explicada pelas diferenças de produtividade. A parte não explicada é não apenas comumente utilizada como uma medida de discriminação, mas também se refere às diferenças de elementos não observados pela estimação. Assim, a diferença de rendimentos entre sexos pode ser determinada como a diferença entre as equações estimadas por sexo, de forma que:

$$\Delta w = \ln \bar{w}_m - \ln \bar{w}_f = \bar{X}'_m \hat{\beta}_m - \bar{X}'_f \hat{\beta}_f, \quad (2)$$

em que m se refere às características dos homens e f faz referência às características femininas. Multiplicando-se, rearranjando-se e isolando-se os termos de (2), obtém-se:

$$\Delta w = (\bar{X}_m - \bar{X}_f)' \hat{\beta}_f + \bar{X}'_f (\hat{\beta}_m - \hat{\beta}_f) + (\bar{X}_m - \bar{X}_f)' (\hat{\beta}_m - \hat{\beta}_f). \quad (3)$$

A equação (3) é a decomposição salarial de Oaxaca-Blinder. Com a expressão definida por (3), é possível identificar a parcela do *gap* salarial que é explicado pelas diferenças nos atributos produtivos e a outra parcela, a qual não é relacionada à produtividade dos trabalhadores.

Entretanto, a estimação de coeficientes pela média obtida partir da decomposição de Oaxaca-Blinder pode esconder características importantes do comportamento dos salários dos indivíduos. Dessa forma, a regressão por quantis surge como alternativa para a estimação dos retornos educacionais ao longo da curva de distribuição dos salários, com a característica adicional de produzir estimativas mais robustas com relação a valores extremos, se comparadas àquelas obtidas em regressões à média.

O método de decomposição para além da média que mais se aproxima dos objetivos da metodologia original de Oaxaca-Blinder é o desenvolvido por Firpo,

Fortin e Lemieux (2009) e aplicado em Firpo, Fortin e Lemieux (2011), e permite a realização de uma decomposição que computa o efeito individual de cada covariada na distribuição incondicional (Santos e Hermeto, 2019).

O método é baseado na abordagem da função de influência recentrada (RIF – em inglês, *recentered influence function*). A função de influência (IF – em inglês, *influence function*) de determinada estatística ($\nu(F_Y)$) representa a influência de observação individual sobre aquela distribuição e é um método bastante empregado para estimações robustas, utilizado sempre que a IF for definida para a estatística de interesse. Conforme definido por Firpo, Fortin e Lemieux (2009) e melhor formalizada por Silva e França (2016), a RIF é dada por:

$$RIF(y, F_Y) = \nu(F_Y) + IF(y, F_Y). \quad (4)$$

Tal que o valor esperado da $RIF(y, F_Y)$ é:

$$E[RIF(y, F_Y)] = \int RIF(y, F_Y) dF_Y(y) = \nu(F_Y). \quad (5)$$

A equação (5) indica que qualquer estatística de interesse $\nu(F_Y)$ pode ser expressa em termos de um valor esperado. Considerando-se que a variável Y é observada na presença de um conjunto de covariadas arranjadas em um vetor X , pode-se assumir que Y e X são conjuntamente distribuídas de acordo com $F_{Y,X}(y, x)$ e que a função de distribuição de Y condicionada em X é dada por $F_{Y|X}(y | X = x)$. Empregando-se a lei de expectativas iteradas para expressar $\nu(F_Y)$ em termos da expectativa condicional da $RIF(y; \nu, F_Y)$, dado X , sobre a função de distribuição incondicional (marginal) de Y , obtém-se:

$$\nu(F_Y) = \int RIF(y, F_Y) dF_Y(y) = \int E[RIF(y, F_Y) | X = x] dF_X(x). \quad (6)$$

Segundo Silva e França (2016), a expressão (6) apresenta uma propriedade essencial na abordagem de quantis incondicionais, provendo uma forma simples de escrever qualquer funcional $\nu(F_Y)$ em termos do valor esperado da $RIF(y, F_Y)$. Em suma, a regressão RIF propõe a linearização da distribuição marginal das variáveis; isso torna possível atender à lei das expectativas iteradas e realizar a decomposição salarial seguindo a mesma lógica do método de Oaxaca-Blinder.

Cabe destacar que, para obter quaisquer diferenças lineares e não lineares na distribuição conjunta entre esses grupos, Firpo, Fortin e Lemieux (2018) descrevem o uso de regressões RIF em combinação com uma estratégia ponderada, conforme DiNardo, Fortin e Lemieux (1996), como uma metodologia viável para decompor diferenças nas estatísticas de distribuição além da média. Esse método é chamado de *decomposição RIF* e possui três grandes vantagens: a simplicidade de sua implementação; a possibilidade de obter contribuições detalhadas de

covariáveis individuais na decomposição agregada; e a possibilidade de expandir a análise para qualquer estatística para a qual uma RIF possa ser definida.

Conforme Rios-Avila (2020), dado que exista uma função de distribuição conjunta que descreva todos os relacionamentos entre a variável dependente Y , as características exógenas X e a variável categórica T que identifica a associação ao grupo da população ($f_{Y,X,T}$), a distribuição cumulativa de Y dado T pode ser escrita como

$$F_{Y|T=k} = \int F_{Y|X,T=k} dF_{X|T=k}. \quad (7)$$

Dessa forma, para analisar as diferenças entre dois grupos, a distribuição condicional cumulativa de Y pode ser usada para calcular a diferença na estatística distributiva ν :

$$\Delta\nu = \nu\left(\int F_{Y|X,T=1} dF_{X|T=1}\right) - \nu\left(\int F_{Y|X,T=0} dF_{X|T=0}\right), \quad (8)$$

em que $T=0,1$ são os grupos de análise. A equação (8) mostra que as diferenças nas estatísticas de interesse $\Delta\nu$ surgirão devido a diferenças na distribuição de X ($dF_{X|T=1} \neq dF_{X|T=0}$), ou em razão de diferenças nas relações entre Y e X ($dF_{Y|X,T=1} \neq dF_{Y|X,T=0}$).

Para identificar quão importantes são as diferenças nas características (efeito de composição) e as diferenças nos coeficientes (efeito da estrutura salarial) para explicar a diferença geral na estatística distributiva ν , é necessário criar um cenário contrafactual. Como mostra Rios-Avila (2020), a estatística contrafactual ν_c pode ser definida da seguinte maneira:

$$\nu_c = \nu(F_Y^c) = \nu\left(\int F_{Y|X,T=0} dF_{X|T=1}\right). \quad (9)$$

Usando-se esse contrafactual, a diferença na estatística de distribuição ν pode ser desagregada em dois componentes:

$$\begin{aligned} \Delta\nu &= \Delta\nu_S + \Delta\nu_X \\ \Delta\nu &= \nu_1 - \nu_c + \nu_c - \nu_0, \end{aligned} \quad (10)$$

em que $\Delta\nu_X$ reflete a diferença atribuída às diferenças nas características e $\Delta\nu_S$ corresponde às diferenças atribuídas às relações entre Y e X (efeito da estrutura). Dessa forma, regressões RIF distintas podem ser estimadas para cada grupo, com o objetivo de que a estatística contrafactual possa ser identificada da seguinte forma:

$$\nu_1 = E\left[RIF\left\{y, \nu(F_{Y|T=1})\right\}\right] = \bar{X}' \hat{\beta}^1$$

$$\begin{aligned} v_0 &= E\left[RIF\{y, v(F_{y|T=0})\} \right] = \bar{X}' \hat{\beta}^0 \\ v_c &= \bar{X}' \hat{\beta}^0. \end{aligned} \tag{11}$$

A expressão (11) corresponde à decomposição de Oaxaca-Blinder padrão, em que

$$\Delta v_x = (\bar{X}^1 - \bar{X}^0)' \beta^0 \text{ e } \Delta v_s = \bar{X}' (\hat{\beta}^1 - \hat{\beta}^0). \tag{12}$$

É necessário considerar, também, uma possível relação de simultaneidade existente entre o capital social e os salários dos indivíduos. Ou seja, a possibilidade de existência de endogeneidade entre as variáveis escolhidas para representação do capital social com as demais variáveis de análise. Assim, dada a possibilidade de presença de endogeneidade na especificação, as estimações são realizadas utilizando-se o método de regressão por variáveis instrumentais (IV, na sigla em inglês), com vistas ao controle de prováveis efeitos não observáveis.

De acordo com Wooldridge (2016), para obter estimadores consistentes dos β 's em (1) quando for imperativa a correção dos efeitos da endogeneidade, é necessário definir uma nova variável z que seja observável e que satisfaça às seguintes condições: i) z não seja correlacionada com ε – isto é, $Cov(z, \varepsilon) = 0$; e ii) z seja correlacionada com x de interesse, $Cov(z, x) \neq 0$. Atendidas essas condições, a variável z é uma variável instrumental de x , ou simplesmente um instrumento para x , definidos conforme descrito no quadro 1.

As especificações dos modelos empíricos adaptam o disposto em Shultz (1961), Becker (1962), Oaxaca (1973), Blinder (1973), Mincer (1974), Firpo, Fortin e Lemieux (2009), Moraes Filho (2011), Madalozzo, Martins e Lico (2015), Santos e Hermeto (2019) e Cruz e Irffi (2019). A fim de verificar se os resultados obtidos com as estimativas a partir da amostra reverberam para a população, são estimados, a partir do plano amostral expandido, os modelos com os métodos MQO e IV, utilizando-se as mesmas especificações já definidas anteriormente. Os demais métodos econométricos aplicados para a amostra (decomposição de Oaxaca-Blinder, RIF e decomposição RIF) não são estimados para a expansão da população em decorrência da falta de adequação metodológica para a amostra expandida.

4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

O objetivo central deste trabalho é pôr em teste a hipótese de que o capital social reduz as diferenças salariais entre homens e mulheres, na medida em que é fator relevante para a melhoria na composição dos salários em geral, mas com efeito preponderante sobre os rendimentos das mulheres. Os modelos utilizados levam em consideração a decomposição dos efeitos do capital social por sexo, assim

como a desagregação para homens e mulheres separadamente, para os indivíduos maiores de 18 anos.

A tabela 2 traz as estimativas da equação minceriana (resultados robustos com MQO), e a tabela 3 apresenta os resultados obtidos com variáveis instrumentais. A tabela 4 traz os resultados da decomposição de Oaxaca-Blinder com estimativas na média. Por sua vez, a tabela 5 apresenta os resultados para as regressões quantílicas incondicionais, enquanto a tabela 6 traz as estimativas obtidas com a decomposição RIF por intermédio da metodologia de Oaxaca-Blinder.

Os resultados relacionados aos modelos de 1 a 4 na tabela 2⁶ são aqueles que consideram apenas uma equação na análise com especificação distinta, enquanto os modelos 5 e 6 apresentam os resultados para as estimações feitas para cada sexo, sendo o modelo 5 relativo aos homens e o modelo 6, às mulheres. Considerando-se apenas uma única equação, o *gap* salarial calculado sem o efeito do capital social é 21,28% (modelo 1). Quando o capital social é incluído na equação (modelo 2), o *gap* praticamente mantém-se no mesmo percentual e o impacto médio do capital social na variável dependente é 4,61%. Quando o capital social é mensurado pela variável interada *cs_mulher* (modelo 3), que representa o estoque de capital social das mulheres (resultado do produto entre *mulher* e *capital_social*), o impacto de ser mulher com estoque de capital social na variável dependente é de 4,54%; porém, o *gap* “direto” eleva-se para 24,37%.

TABELA 2
Impactos do capital social sobre os salários: estimativas com MQO (2013)

Variáveis	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
mulher	-0,2128*** (0,0075)	-0,2132*** (0,0075)	-0,2437*** (0,0109)	-0,2134*** (0,0128)	-	-
capital_social	-	0,0461*** (0,0078)	-	0,0460*** (0,0104)	0,0455*** (0,0104)	0,0528*** (0,0116)
cs_mulher	-	-	0,0454*** (0,0115)	0,0003 (0,0153)	-	-
R ²	0,4035	0,4041	0,4038	0,4041	0,4140	0,3944
Obs.	35.032	35.032	35.032	35.032	18.751	16.281

Fonte: Com base nos resultados das estimações.

Elaboração dos autores.

Obs.: 1. Erros-padrão robustos das estimativas entre parênteses.

2. Significância: *** = 1%; ** = 5%; * = 10%.

Quando são incluídas as três variáveis de interesse na equação (modelo 4), o *gap* mantém-se praticamente o mesmo que o obtido no modelo 1, e o impacto do capital social continua positivo, na mesma magnitude e significante; porém, a

6. As estimativas completas podem ser obtidas com os autores.

variável interada não se mostra significativa. Na estimação de equações por grupo, o capital social revela-se significativo para ambos os sexos, com impacto superior nos salários das mulheres (5,28%) em relação aos dos homens (4,55%).

Os resultados indicam que, em todas as especificações estimadas com MQO, o capital social provoca um efeito de melhoria na composição salarial, e os modelos 3 e 6 sugerem que as mulheres possuem vantagem relativa nessa melhoria. Assim, o resultado relacionado ao impacto do capital social sobre os rendimentos é significativo e positivo para os salários médios, quando considerados ambos os sexos.

É necessário considerar, também, a dubiedade existente não somente no fato de que as diferenças salariais podem ser menores devido ao capital social, como também de que um salário maior das mulheres (diferença salarial menor) induz o resultado de melhoria salarial obtido com o uso do capital social. Ou seja, existe a possibilidade de o capital social ser correlacionado com o termo de erro da equação de salários. Nessa situação, dada uma possível relação de simultaneidade existente entre o capital social e os salários dos indivíduos, surge a possibilidade de existência de endogeneidade entre as variáveis escolhidas para representação da ligação entre o capital social e a composição dos salários.

Portanto, dada a possibilidade de haver uma relação endógena, e considerando-se a necessidade de tratamento dessa endogeneidade entre as variáveis de interesse, são utilizados dois instrumentos, conforme definido no quadro 1. Os instrumentos são construídos partindo-se do pressuposto de que os amigos descritos na *proxy* de capital social estão inseridos nas redes de relacionamento construídas nas atividades esportivas artísticas e religiosas, mas que essas amizades não possuem relação direta com a composição salarial dos indivíduos da amostra.

A tabela 3 resume os resultados com utilização de variáveis instrumentais. O método generalizado dos momentos em duas etapas (GMM2S, na sigla em inglês) foi utilizado para gerar estimativas eficientes dos coeficientes, bem como estimativas consistentes dos erros-padrão. O modelo 1 considera o total de observações da amostra utilizada. Por sua vez, os modelos 2 e 3 são relativos às estimativas por sexo, homens e mulheres, respectivamente. Os resultados obtidos para a variável de interesse mostram-se estatisticamente significantes em todos os modelos estimados. Os testes de subidentificação e validade dos instrumentos utilizados também se revelam satisfatórios, exceto o teste de Hansen para o modelo 2, o que não invalida o resultado, visto que o importante para essa análise são as comparações por grupo.

A magnitude dos resultados difere bastante das obtidas com MQO, provavelmente devido à utilização do conjunto de instrumentos. Para os homens, a existência de capital social provoca um aumento de 140% nos seus salários. Por sua

vez, para a amostra das mulheres, os resultados indicam que seus salários aumentam, aproximadamente, 196% devido ao estoque de capital social. Nesse sentido, para as mulheres, o impacto do capital social sobre a composição dos seus salários é de aproximadamente 56 pontos percentuais (p.p.) a mais em relação aos homens.

Os resultados da tabela 3 sugerem duas coisas distintas: a primeira é que o capital social realmente promove a melhoria na composição dos salários, independentemente do sexo do indivíduo; e a segunda, que o impacto dessa melhoria é evidentemente superior para as mulheres.

Em relação aos impactos causados pelas condições de saúde dos indivíduos sobre a composição dos seus salários, as estimativas com MQO sugerem que o fato de possuir diagnóstico positivo de doenças crônicas e do coração provoca redução salarial, sobretudo nos salários femininos.

TABELA 3
Impactos do capital social sobre os salários: estimativas com IV (2013)

Especificações	capital_social	Teste de subidentificação	Teste de validade dos instrumentos	Observação
Modelo 1 (amostra total)	2,2314*** (0,3739)	52,496 (0,0000)	3,452 (0,0632)	35.032
Modelo 2 (homens)	1,4063*** (0,3921)	26,543 (0,0000)	4,321 (0,0376)	18.751
Modelo 3 (mulheres)	1,9562*** (0,4412)	32,263 (0,0000)	0,763 (0,3824)	16.281

Fonte: Com base nos resultados das estimações.

Elaboração dos autores.

Obs.: 1. Erro-padrão das estimativas e *p-values* dos testes entre parênteses.

2. Significância: *** = 1%; ** = 5%; * = 10%.

Por sua vez, a violência ocasionada por pessoa conhecida impacta negativamente os salários das mulheres, enquanto os salários masculinos parecem sofrer impactos relativamente maiores quando a violência é provocada por pessoa desconhecida.

Nas estimativas com variáveis instrumentais, percebeu-se que, enquanto os diagnósticos positivos das doenças analisadas passaram a não ser significativos, a violência causada por pessoa conhecida é fator relevante para a redução dos salários para ambos os sexos.

Mensurar diretamente o capital social pode ser algo de difícil execução. Assim sendo, dada a assunção de sua existência e, portanto, a consequente influência que exerce nas relações sociais entre grupos e indivíduos, sua identificação pode ocorrer a partir da constatação dos efeitos que sua ocorrência causa nos agentes envolvidos.

Dessa forma, os efeitos do capital social na decomposição de Oaxaca-Blinder são mensurados de duas formas: direta e indiretamente. A decomposição medida pelo impacto indireto compara dois grupos, homens e mulheres, e consiste em estimar as diferenças salariais tendo *mulher* como variável de separação de grupos e obter os resultados para três grupos específicos: especificação padrão; especificação que considera somente aquelas pessoas que não têm capital social (*capital_social* = 0); e especificação para aquelas pessoas que possuem estoque de capital social (*capital_social* = 1).

Por sua vez, o impacto direto do capital social sobre os salários estima a decomposição comparando mulheres com estoque de capital social e os demais indivíduos da amostra, independentemente do sexo. Consiste em substituir a variável de separação de grupos pela variável interada *cs_mulher*, o que significa que haverá um grupo em que estarão todas as mulheres que possuem estoque de capital social e outro grupo com os demais indivíduos. A tabela 4 apresenta os resultados obtidos para essas condições de especificação, com utilização da decomposição de Oaxaca-Blinder.

TABELA 4
Decomposição de Oaxaca-Blinder sobre os salários na média (2013)

Método de estimação (impacto indireto)	Homens	Mulheres	Diferença salarial	Observação
O-B padrão	1,9550*** (0,0063)	1,8493*** (0,0067)	0,1057*** (0,0092)	35.032
O-B para grupo sem capital social (<i>capital_social</i> = 0)	1,8515*** (0,0102)	1,6677*** (0,0111)	0,1838*** (0,0151)	11.455
O-B para grupo com capital social (<i>capital_social</i> = 1)	2,0085*** (0,0080)	1,9314*** (0,0082)	0,0771*** (0,0114)	23.577
Método de estimação (impacto direto)	Outros	<i>cs_mulher</i>	Diferença salarial	Observação
O-B padrão (<i>cs_mulher</i>)	1,8939*** (0,0056)	1,9314*** (0,0082)	- 0,0375*** (0,0099)	35.032

Fonte: Com base nos resultados das estimações.

Elaboração dos autores.

Obs.: 1. Erro-padrão das estimativas entre parênteses.

2. Significância: *** = 1%; ** = 5%; * = 10%.

3. A descrição *outros* faz referência a todos os indivíduos da amostra, exceto mulheres com estoque de capital social positivo.

O primeiro ponto a ser observado na tabela 4 é que as estimações que consideram a influência do capital social geram estimativas salariais maiores para ambos os sexos. Considerando-se os resultados a partir do impacto indireto do capital social, é perceptível que a existência de estoque de capital social por parte das mulheres produz um efeito de melhoria na composição salarial feminina. Dessa forma, reduz a diferença salarial em mais de 10 p.p. em relação àquelas mulheres que não possuem tal condição. A variável interada produz uma diferença salarial em favor das mulheres de 0,0375.

Os resultados apontam que há claro impacto positivo do estoque de capital social para as mulheres que o possuem, visto que seu acúmulo melhora a composição salarial das mulheres via impacto indireto e inverte o sentido do hiato salarial em desfavor dos homens, quando medido via impacto direto. Esses resultados sugerem que o acúmulo de estoque de capital social por parte das mulheres possibilita melhor composição salarial, em relação tanto aos homens em geral quanto às mulheres que não o possuem.

A tabela 5 resume os resultados das estimativas no método baseado na abordagem da função de influência recentrada. Os resultados relacionados aos modelos de 1 a 3 são aqueles que consideram especificações distintas e apenas uma equação na análise, enquanto os modelos 4 e 5 apresentam os resultados de mensuração do capital social para as estimações feitas por grupo de sexo. As estimativas foram realizadas para cada quantil pontual selecionado.

O modelo 1, exposto na tabela 5, apresenta os resultados para o hiato salarial entre homens e mulheres para os quantis selecionados ao longo da distribuição de salários, mostrando que menores diferenças salariais ocorrem entre o 20º e o 30º quantil. À medida que os salários crescem acima do 40º quantil, as diferenças de rendimentos elevam-se e ficam sempre superiores a 20%, em desfavor das mulheres.

As estimativas das variáveis de interesse para o modelo 2 da tabela 5 mantêm aproximadamente os mesmos valores para os hiatos salariais vistos no modelo 1 e incluem a medida do impacto do capital social na distribuição dos salários. Os resultados mostram que o capital social provoca um efeito positivo na composição dos salários da amostra entre 2,36% e 7,89%, ao longo da distribuição dos rendimentos, e que a predominância dos maiores efeitos ocorre acima do 40º quantil.

TABELA 5
Estimativas por quantis incondicionais dos salários para diferentes especificações (2013)

quantil	Modelo 1		Modelo 2				Modelo 3			Modelo 4		Modelo 5	
	mulher	R ²	mulher	cap_social	R ²	mulher	cap_social	cs_mulher	R ²	cap_social (homem)	R ²	cap_social (mulher)	R ²
0,10	-0,2460*** (0,0165)	0,1284	-0,2463*** (0,0165)	0,0323* (0,0181)	0,1291	-0,3244*** (0,0312)	-0,0197 (0,0229)	0,1158*** (0,0360)	0,1294	0,0207 (0,0183)	0,1454	0,0608** (0,0294)	0,1265
0,30	-0,1543*** (0,0068)	0,2175	-0,1545*** (0,0068)	0,0236*** (0,0072)	0,2177	-0,1900*** (0,0122)	-0,0001 (0,0093)	0,0526*** (0,0143)	0,2181	0,0219* (0,0114)	0,2314	0,0326*** (0,0095)	0,1804
0,50	-0,2336*** (0,0090)	0,2624	-0,2340*** (0,0090)	0,0442*** (0,0096)	0,2629	-0,2513*** (0,0159)	0,0326*** (0,0127)	0,0257 (0,0188)	0,2629	0,0512*** (0,0125)	0,2722	0,0553*** (0,0137)	0,2613
0,80	-0,2720*** (0,0171)	0,3000	-0,2726*** (0,0171)	0,0789*** (0,0172)	0,3003	-0,1871*** (0,0275)	0,1358*** (0,0236)	-0,1267*** (0,0339)	0,3006	0,0991*** (0,0221)	0,3137	0,0468** (0,0218)	0,3037
0,90	-0,2716*** (0,0211)	0,2207	-0,2721*** (0,0210)	0,0586*** (0,0201)	0,2209	-0,1877*** (0,0313)	0,1147*** (0,0281)	-0,1250*** (0,0400)	0,2211	0,0747*** (0,0281)	0,2595	0,0669** (0,0269)	0,2014
Obs.	35.032		35.032	35.032		35.032	35.032	35.032		18.751		16.281	

Fonte: Com base nos resultados das estimações.

Elaboração dos autores.

Obs.: 1. Erro-padrão das estimativas entre parênteses.

2. Significância: *** = 1%; ** = 5%; * = 10%.

Nos resultados do modelo 3, pode-se perceber que o *gap* é consideravelmente maior que nos modelos anteriores para a distribuição de salários abaixo da mediana, chegando a ser superior a 30% em favor dos homens no décimo quantil. Todavia, o resultado mais interessante para essa especificação diz respeito à comparação dos resultados para as variáveis *capital_social* e *cs_mulher*. Até o 40º quantil, não há efeito significativo do capital social se considerados ambos os sexos, entretanto esse impacto é significativo e positivo para a composição salarial das mulheres, chegando a ter impacto de até 11,58% sobre os acréscimos de rendimentos na faixa dos menores salários, perdendo significância após o 40º quantil da distribuição de salários.

Quando considerada a distribuição acima do 70º quantil, há um claro efeito do capital social na composição de rendimentos dos homens em detrimento dos salários femininos, visto que, enquanto o *capital_social* aumenta, o *cs_mulher* mostra reversão dos efeitos do capital social sobre os rendimentos femininos. Tais resultados sugerem que a influência do capital social para a melhoria da composição salarial das mulheres ocorre na parte inferior da distribuição de rendimentos e vai perdendo significância quando a distribuição salarial se aproxima de sua mediana.

Ainda da tabela 5, os modelos 4 e 5 apresentam um resumo das estimativas considerando as equações por sexo. No modelo 4, os resultados sugerem a inexistência de efeito do capital social sobre os salários dos homens para as faixas abaixo do 40º quantil e um impacto positivo e crescente a partir desse ponto da distribuição de rendimentos. Por sua vez, para as mulheres, todas as estimativas se mostram significativas a 5%, com predominância de efeitos superiores aos dos homens até a mediana da distribuição de rendimentos. Assim, corrobora-se a hipótese de que o capital social promove melhorias salariais médias sobre os salários femininos com impacto superior àquelas provocadas nos salários masculinos, predominantemente até o quantil relativo à mediana da distribuição de salários.

Partindo-se das definições dos impactos direto e indireto anteriores, os resultados da implementação da decomposição RIF por intermédio do método de Oaxaca-Blinder são expostos na tabela 6. As estimativas também foram realizadas para cada quantil pontual selecionado.

A primeira observação interessante que pode ser feita a partir dos resultados para os efeitos indiretos na tabela 6 é que os salários do grupo que possui estoque positivo de capital social são maiores em relação àqueles que não o possuem, independentemente do sexo. Resultados que se repetem quando esse mesmo grupo é comparado com a amostra que não diferencia os indivíduos pelo estoque de capital social, o que evidencia, *per se*, o efeito positivo do capital social na composição salarial dos indivíduos.

Quando comparados os grupos com e sem estoque de capital social, percebe-se claramente que as diferenças salariais, apesar de persistirem, ocorrem em valores

significativamente menores. Enquanto o hiato salarial para o grupo que não possui estoque de capital social ocorre no intervalo entre 0,0699 e 0,2356, para o grupo que possui estoque de capital social esse intervalo ocorre entre 0,0389 e 0,1132 ao longo da distribuição de rendimentos, não existindo diferença significativa para o décimo quantil e a mediana.

O último grupo de resultados da tabela 6 apresenta as estimativas para o impacto direto do capital social sobre os salários, em que a variável de separação de grupos passa a representar, como anteriormente, todas as mulheres que possuem estoque de capital social, enquanto no grupo adverso estão os demais indivíduos da amostra.

A obtenção dos salários com utilização da variável interada ao longo da distribuição de rendimentos resulta em estimativas maiores para as mulheres em quase todos os quantis sugeridos para análise. Nos 30º, 40º e 50º quantis, os hiatos salariais não se mostraram significativos estatisticamente, o que indica que não há diferenças salariais significativas nesses pontos da distribuição dos salários. A influência do capital social na composição salarial feminina e a consequente redução do *gap* salarial por sexo parecem ocorrer, em grande parte, próximo das caudas da distribuição de probabilidade dos rendimentos.

Os impactos das condições de saúde dos trabalhadores nas estimativas por *quantis* sugerem que o diagnóstico positivo de doenças provoca efeitos de reduções salariais naquelas mulheres que possuem menores remunerações, enquanto a violência sofrida parece reduzir mais diretamente os salários dos homens. Em síntese, as condições de saúde dos indivíduos impactam diretamente sua composição de rendimentos, apesar de sua inclusão nas estimativas realizadas não provocarem efeitos significativos na relação entre o capital social e os diferenciais salariais por sexo tratados neste estudo.

Foram realizadas também estimativas com a extrapolação dos dados amostrais para a população, seguindo Szwarcwald *et al.* (2014), com o objetivo de definir os fatores de expansão dos domicílios e do morador selecionado para responder à terceira parte do questionário e no intuito de verificar se os resultados obtidos com as estimativas a partir da amostra reverberam para a população. Os modelos com os métodos MQO e IV foram executados utilizando-se as mesmas especificações já definidas anteriormente.

TABELA 6
Comparação dos efeitos do capital social sobre os diferenciais de salários com decomposição quantílica incondicional (2013)

quantil	Efeitos indiretos										Efeitos diretos				
	OB-RIF padrão ¹					OB-RIF para grupo de indivíduos sem CS (capital_social = 0)					OB-RIF para grupo de indivíduos com CS (capital_social = 1)		OB-RIF para grupo de mulheres com CS		
	Homem	Mulher	Diferença salarial	Homem	Mulher	Diferença salarial	Homem	Mulher	Diferença salarial	Homem	Mulher	Diferença salarial	Outros	cs_mulher	Diferença salarial
0,10	1,0600*** (0,0086)	0,9601*** (0,0122)	0,0999*** (0,0149)	0,9661*** (0,0171)	0,7768*** (0,0224)	0,1894*** (0,0282)	1,0796*** (0,0108)	1,0677*** (0,0111)	0,0119 (0,0154)	0,9673*** (0,0101)	1,0684*** (0,0121)	-0,1011*** (0,0157)			
0,30	1,4730*** (0,0059)	1,4260*** (0,0048)	0,0470*** (0,0076)	1,4058*** (0,0094)	1,3359*** (0,0081)	0,0699*** (0,0124)	1,5181*** (0,0078)	1,4200*** (0,0065)	0,0981*** (0,0102)	1,4082*** (0,0046)	1,4200*** (0,0062)	-0,0118 (0,0077)			
0,50	1,8481*** (0,0068)	1,6874*** (0,0070)	0,1607*** (0,0098)	1,7783*** (0,0107)	1,5727*** (0,0099)	0,2056*** (0,0146)	1,8932*** (0,0088)	1,7889*** (0,0091)	0,1043 (0,0127)	1,7958*** (0,0060)	1,7888*** (0,0091)	0,0070 (0,0109)			
0,80	2,6387*** (0,0125)	2,4583*** (0,0127)	0,1804*** (0,0178)	2,4858*** (0,0184)	2,2502*** (0,0212)	0,2356*** (0,0280)	2,7145*** (0,0154)	2,6649*** (0,0174)	0,0496** (0,0233)	2,5416*** (0,0108)	2,6661*** (0,0180)	-0,1244*** (0,0210)			
0,90	3,2118*** (0,0161)	3,0394*** (0,0152)	0,1724*** (0,0221)	2,9650*** (0,0289)	2,7616*** (0,0288)	0,2034*** (0,0408)	3,2694*** (0,0188)	3,1924*** (0,0189)	0,0770*** (0,0267)	3,0939*** (0,0141)	3,1924*** (0,0189)	-0,0985*** (0,0236)			
Obs.	18,751	16,281	35,032	6,389	5,066	11,455	12,362	11,215	23,577	23,817	11,215	35,032			

Fonte: Com base nos resultados das estimações.

Elaboração dos autores.

Nota: ¹ OB-RIF: Decomposição de Oaxaca-Blinder com função de influência recentrada.

Obs.: 1. Erro-padrão das estimativas entre parênteses.

2. Significância: *** = 1%; ** = 5%; * = 10%.

Os resultados para os hiatos salariais relacionados aos modelos 1, 2 e 3 seguem o mesmo sentido das estimativas para a amostra obtidas anteriormente, das quais as estimativas para a amostra expandida produzem diferenças salariais 2,5 p.p. superiores, em média. No modelo 2, o estoque de capital social é significativo e positivo, mas pouco inferior ao obtido nas estimativas para a amostra. Por sua vez, o estoque de capital social aplicado somente às mulheres no modelo 3 é significativo e 1 p.p. superior no que concerne ao plano amostral. Na especificação para o modelo 4, as evidências apontam que os impactos do capital social sobre os salários para a amostra expandida somente são significativos a 10%.

Considerando-se as estimativas realizadas por sexo, os resultados do modelo 5 apontam para a falta de impacto significativo do capital social sobre os salários masculinos para a população, resultado que diverge do obtido para a amostra, em que o capital social produz efeito positivo sobre os salários dos homens. Por sua vez, para as estimativas do modelo 6, o qual considera somente as mulheres no cálculo, os resultados para o impacto do capital social sobre os salários femininos se mostraram significativos e 0,7 p.p. superior em relação à amostra.

As estimativas com variáveis instrumentais para a amostra expandida, ao se utilizarem os instrumentos já definidos conforme o quadro 1, mostraram que o impacto médio do estoque de capital social sobre os salários é significativo e da ordem de quase 143%, independentemente do sexo do indivíduo. Até mesmo sendo relativamente alto, o resultado é inferior àquele obtido para a amostra em quase 80 p.p. Quando considerado apenas o grupo de homens, o impacto do capital social somente é significativo a 90% de confiança. Por sua vez, para as mulheres, o efeito do capital social sobre os salários é positivo e fortemente significativo, apesar de um pouco inferior em relação às estimativas da amostra.

Diante dos resultados obtidos, é possível corroborar os resultados que este trabalho vem construindo ao longo das estimações realizadas, embasado no fato de que o estoque de capital social melhora a composição dos salários de homens e mulheres, mas tem impacto superior sobre a remuneração feminina. Dessa forma, sugere-se que o hiato salarial existente entre os sexos tende a diminuir quando as mulheres constroem um círculo de amigos, não apenas com amigos diretos, mas também ampliando essa interação social com “os amigos dos amigos” – isto é, com a acumulação de estoque de capital social *bonding* e *bridging*.

5 CONSIDERAÇÕES FINAIS

O capital social, tanto quanto as demais formas de capital (físico e humano), promove a geração de valor, exercendo influência direta na composição dos rendimentos da força de trabalho. Contudo, sua mensuração não ocorre de maneira fácil, e não há variáveis representativas que possam expressá-lo de forma objetiva, principalmente no nível do indivíduo. Nesse sentido, esse trabalho se propõe a

identificar variáveis representativas para o capital social no nível do indivíduo, ao verificar seus efeitos sobre a composição dos salários e a consequente diferença salarial por sexo ao longo da distribuição de rendimentos.

Utilizando-se os dados amostrais da PNS, foram executados diferentes modelos econométricos com estimativas na média e por quantis selecionados, com a intenção de pôr à prova a hipótese de que o capital social reduz as diferenças salariais entre homens e mulheres, na medida em que é fator relevante para a melhoria na composição dos salários em geral, mas com efeito preponderante sobre os rendimentos das mulheres em diferentes faixas de remuneração do trabalho. Os métodos aplicados não se configuraram como concorrentes entre si, tendo-se em vista que a intenção principal foi criar um arcabouço de resultados, sobre diversos aspectos econométricos, para identificar a persistência dos resultados e, assim, corroborar a hipótese em teste.

Os resultados relacionados às condições de saúde dos indivíduos sugerem que a cobertura privada de saúde parece dar algum tipo de garantia à melhoria da remuneração, e esse efeito é tão maior quanto maiores forem os salários auferidos, enquanto os rendimentos femininos parecem absorver em maior magnitude as restrições impostas pela saúde pessoal. De forma geral, os resultados obtidos para os impactos das condições de saúde parecem seguir aqueles já descritos por Fernandes, Bertoldi e Barros (2009), Rodrigues, Cruz e Paixão (2015), Amendola *et al.* (2017) e Cruz e Irffi (2019).

Em linhas gerais, os resultados obtidos apontam que a construção das redes de amizades e relacionamentos tem significativo impacto na melhoria da composição dos salários. Desse modo, sugere-se que o conjunto de interações sociais (amigos e amigos dos amigos) produzem efeitos do estoque de capital social sobre os salários em geral, mas com impacto superior nos rendimentos femininos quando na média. Impacto este que se percebe predominante nos salários mais baixos, principalmente naqueles inferiores à mediana da distribuição de rendimentos, nas faixas salariais femininas mais baixas.

REFERÊNCIAS

- AIZER, A. **Wages, violence and health in the household**. Cambridge, Estados Unidos: NBER, 2007. (Working Paper, n. 13494).
- AKERLOF, G. A.; KRANTON, R. E. Economics and identity. **The Quarterly Journal of Economics**, v. 115, n. 3, p. 715-753, 2000.
- AMENDOLA, F. *et al.* Índice de vulnerabilidade a incapacidades e dependência (IVF-ID), segundo condições sociais e de saúde. **Ciência e Saúde Coletiva**, v. 22, n. 6, p. 2063-2071, 2017.

ANDREWS, P. Gender differences in persuasive communication and attribution of success and failure. **Human Communication Research**, v. 13, n. 3, p. 372-385, 2006.

ANDRIANI, L.; KARYAMPAS, D. **A new proxy of social capital and the economic performance across the Italian regions**. Londres: University of London, 2009. (Birkbeck Working Paper in Economics and Finance, n. 903).

ARAÚJO, V. F.; RIBEIRO, E. P. Diferenciais de salários por gênero no Brasil: uma análise regional. **Revista Econômica do Nordeste**, v. 33, n. 2, p. 1-22, 2002.

ARRAES, R. A. Capital social, fatores institucionais e custos de transação no crescimento econômico: cenários para países e estados do Brasil. *In*: CONGRESSO BRASILEIRO DE CUSTOS, 9., 2002, São Paulo. **Anais ...** São Paulo: ABC, 2002.

BECKER, G. S. Investment in human capital: a theoretical analysis. **Journal of Political Economy**, v. 5, n. 70, p. 9-49, 1962.

BECKER, G. S. A theory of social interactions. **Journal of Political Economy**, v. 82, n. 6, p. 1063-1093, 1974.

BECKER, G. S. **Accounting for tastes**. Cambridge, Estados Unidos: Harvard University Press, 1996.

BERTRAND, M. Coase lecture: the glass ceiling. **Economica**, v. 85, n. 338, p. 205-231, 2018.

BERTRAND, M.; DUFLO, E. Field experiments on discrimination. **Handbook of Economic Field Experiments**, v. 1, p. 309-393, 2017.

BILAC, E. D. Gênero, vulnerabilidade das famílias e capital social: algumas reflexões. *In*: CUNHA, J. M. P. (Org.). **Novas metrópoles paulistas: população, vulnerabilidade e segregação**. Campinas: Nepo/Unicamp, 2006.

BILERT, V. S. S. *et al.* A contribuição do capital social para o desenvolvimento local sustentável. **Ciências Sociais Aplicadas em Revista**, v. 11, n. 21, p. 29-42, 2011.

BLINDER, A. S. Wage discrimination: reduced form and structural estimates. **The Journal of Human Resources**, v. 8, n. 4, p. 436-455, 1973.

BORGHANS, L.; WEEL, B. ter; WEINBERG, B. A. **People people: social capital and the labor-market outcomes of underrepresented groups**. Bonn: IZA, 2005. (IZA Discussion Paper, n. 1494).

BRASIL. Ministério da Saúde. **Pesquisa Nacional de Saúde: referencial teórico**. Brasília: MS, 2020. Disponível em: <https://svs.aids.gov.br/daent/aceso-a-informacao/inqueritos-de-saude/pns/2013/apresentacao/referencial-teorico>. Acesso em: 19 mar. 2020.

BUSS, P. M.; PELLEGRINI FILHO, A. A saúde e seus determinantes sociais. **Physis**, v. 17, n. 1, p. 77-93, 2007.

CAOLI, C. Mulheres receberam 80% do salário dos homens em 2014, mostra IBGE. **Notícias G1**, 17 jun. 2016. Disponível em: <http://g1.globo.com/economia/noticia/2016/06/mulheres-receberam-80-do-salario-dos-homens-em-2014-mostra-ibge.html>. Acesso em: 8 ago. 2019.

CAPRARO, V. **Women are slightly more cooperative than men (in one-shot prisoner's dilemma games played online)**. Londres: Middlesex University London, 21 maio 2018. Disponível em: <https://ssrn.com/abstract=3182429>. Acesso em: 14 mar. 2020.

CHANDLER, R. *et al.* The impact of covid-19 among black women: evaluating perspectives and sources of information. **Ethnicity & Health**, v. 26, n. 1, p. 80-93, 2021.

CHEVALIER, A. Education, occupation and career expectations: determinants of the gender pay gap for UK graduates. **Oxford Bulletin of Economics and Statistics**, v. 69, n. 6, p. 819-842, 2007.

CLARIDGE, T. **Social capital and natural resource management: an important role for social capital?** Brisbane: University of Queensland, jun. 2004.

COLEMAN, J. S. **Foundations of social theory**. Cambridge, Estados Unidos: Harvard University Press, 1990.

COLLISCHON, M.; EBERL, A. Social capital as a partial explanation for gender wage gaps. **The British Journal of Sociology**, v. 72, n. 3, p. 757-773, 2021.

CORRÊA, C. S.; QUEIROZ, B. L.; FAZITO, D. Relação entre tamanho e estrutura da rede de apoio e o tempo individual dedicado à atenção ao idoso na cidade de São Paulo, 2000. **Revista Brasileira de Estudos de População**, v. 33, n. 1, p. 75-97, 2016.

CREMONESE, D. Insolidarismo e cordialidade: uma análise das mazelas políticas do Brasil. *In*: BAQUERO, M.; CREMONESE, D. (Org.). **Capital social: teoria e prática**. Ijuí: Unijuí, 2006.

CRUZ, M. S.; ARAUJO, J.; PAIXAO, A. N. Family structure and its impacts on the restrictions of self-perception of elderly health levels in Brazil. **Ciência & Saúde Coletiva**, v. 23, n. 8, p. 2751-2762, 2018.

CRUZ, M. S.; IRFFI, G. D. Qual o efeito da violência contra a mulher brasileira na autopercepção da saúde? **Ciência & Saúde Coletiva**, v. 24, n. 7, p. 2531-2542, 2019.

DINARDO, J.; FORTIN, N.; LEMIEUX, T. Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. **Econometrica**, v. 64, n. 5, p. 1001-1044, 1996.

DUQUE, E. Capital social como instrumento de desenvolvimento sustentável. **Configurações**, v. 11, p. 189-201, 2013.

FAUSTINO, I. A.; ARAÚJO, E.; MAIA, K. Mercado de trabalho e discriminação: uma análise das diferenças salariais e discriminação por gênero no Brasil e macrorregiões (2004-2014). *In*: ENCONTRO NACIONAL DE ECONOMIA, 45., 2017, Natal. **Anais...** Natal: Anpec, 2017.

FERNANDES, L. C. L.; BERTOLDI, A. D.; BARROS, A. J. D. Utilização dos serviços de saúde pela população coberta pela Estratégia de Saúde da Família. **Revista de Saúde Pública**, v. 43, n. 4, p. 595-603, 2009.

FIGUEIRÓ, M. T. *et al.* Traços de personalidade de estudantes de psicologia. **Psicólogo Informação**, v. 14, n. 14, p. 13-28, 2010.

FIRPO, S.; FORTIN, N.; LEMIEUX, T. Unconditional quantile regressions. **Econometrica**, v. 77, n. 3, p. 953-973, 2009.

FIRPO, S.; FORTIN, N.; LEMIEUX, T. Decomposition methods in economic capital social. *In*: ACEMOGLU, D.; AUTOR, D. (Ed.). **Handbook of labor economics**. Amsterdam: Elsevier, 2011.

FIRPO, S.; FORTIN, N.; LEMIEUX, T. Decomposing wage distributions using recentered influence function regressions. **Econometrics**, v. 6, n. 2, p. 28-68, 2018.

GAUSMAN, J.; LANGER, A. Sex and gender disparities in the covid-19 pandemic. **Journal of Women's Health**, v. 29, n. 4, p. 465-467, 2020.

GLASER, K.; GRUNDY, E. Marital status and long-term illness in Great Britain. **Journal of Marriage and the Family**, v. 59, n. 1, p. 156-164, 1997.

GOMES, M. R.; SOUZA, S. C. I. Assimetrias salariais de gênero e a abordagem regional no Brasil: uma análise segundo a admissão no emprego e setores de atividade. **Revista de Economia Contemporânea**, v. 22, n. 3, p. 1-31, 2019.

GROVE, W. A.; HUSSEY, A.; JETTER, M. The gender pay gap beyond human capital: heterogeneity in noncognitive skills and in labor market tastes. **The Journal of Human Resources**, v. 46, n. 4, p. 827-874, 2011.

HARB, N.; ROUHANA, T. Earnings and gender wage gap in Lebanon: the role of the human and social capital. **Applied Economics**, v. 52, n. 44, p. 4834-4849, 2020.

HERMANN, Z.; KOPASZ, M. The effects of social capital on wage income: a multi-country analysis. **Corvinus Journal of Sociology and Social Policy**, v. 2, n. 2, p. 3-26, 2011.

KAZTMAN, R. **Activos y estructuras de oportunidades**: estudios sobre las raíces de la vulnerabilidad social en Uruguay. [s.l.]: Repositório CEPAL, 1999. Disponível em: <https://repositorio.cepal.org/server/api/core/bitstreams/7cb04f5a-220f-40de-89ee-114ccb1d4195/content>. Acesso em: 19 mar. 2020.

KLASEN, S. *et al.* **What drives female labor force participation?** Comparable micro-level evidence from eight developing and emerging economies. Bonn: IZA, 2019. (IZA Discussion Paper, n. 12067).

MADALOZZO, R. Occupational segregation and the gender wage gap in Brazil: an empirical analysis. **Economia Aplicada**, v. 14, n. 2, p. 147-168, 2010.

MADALOZZO, R.; MARTINS, S.; LICO, M. **Segregação ocupacional e hiato salarial entre os gêneros**. São Paulo: Insper, 2015. (Insper Working Paper, n. 357).

MANSKI, C. F. Economic analysis of social interactions. **Journal of Economic Perspectives**, v. 14, n. 3, p. 115-136, 2000.

MINCER, J. **Schooling, experience and earnings**. Nova York: Columbia University Press, 1974.

MOLINA, J. A. *et al.* Gender differences in cooperation: experimental evidence on high school students. **Plos One**, v. 8, n. 12, p. 1-10 (p. e83700), 2013.

MOLYNEUX, M. Gender and the silences of social capital: lessons from Latin America. **Development and Change**, v. 33, n. 2, p. 167-188, 2002.

MORAIS FILHO, G. B. **Discriminação salarial e segregação ocupacional**: uma análise de gênero no mercado de trabalho paraibano (1995-2009). 2011. Dissertação (Mestrado) – Programa de Pós-Graduação em Economia, Universidade Federal da Paraíba, João Pessoa, 2011.

MORIN, L. Do men and women respond differently to competition? Evidence from a major education reform. **Journal of Labor Economics**, v. 33, n. 2, p. 443-491, 2015.

NIEDERLE, M.; VESTERLUND, L. Do women shy away from competition? Do men compete too much? **The Quarterly Journal of Economics**, v. 122, n. 3, p. 1067-1101, 2007.

NUNES, V. F.; PIZZI, J. The social capital category in Putnam: for a socially responsible. **International Journal of Advances in Social Science and Humanities**, v. 2, n. 10, p. 38-48, 2014.

OAXACA, R. Male-female wage differentials in urban labor markets. **International Economic Review**, v. 14, n. 3, p. 693-709, 1973.

PARADELLA, R. Diferença cai em sete anos, mas mulheres ainda ganham 20,5% menos que homens. **Agência de Notícias IBGE**, Rio de Janeiro, 11 mar. 2019. Disponível em: [https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/23924-diferenca-cai-em-sete-anos-mas-mulheres-ainda-ganham-20-5-menos-que-homens#:~:text=Mesmo%20com%20uma%20leve%20queda,Domic%C3%ADlios%20Cont%C3%ADnua%20\(PNAD%20Cont%C3%ADnua\)](https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/23924-diferenca-cai-em-sete-anos-mas-mulheres-ainda-ganham-20-5-menos-que-homens#:~:text=Mesmo%20com%20uma%20leve%20queda,Domic%C3%ADlios%20Cont%C3%ADnua%20(PNAD%20Cont%C3%ADnua)). Acesso em: 20 fev. 2021.

POLLÁN, M. M. **Capital social**: glosario. Santiago de Compostela: Camiño do Faro, 2013.

PUTNAM, R. D. **Bowling alone**: the collapse and revival of American community. Nova York: Simon & Schuster, 2000.

PUTNAM, R. D.; LEONARDI, R.; NONETTI, R. Y. **Making democracy work**: civic traditions in modern Italy. Princeton: Princeton University Press, 1993.

RIBEIRO, L. L.; ARAÚJO, J. A. Capital social e pobreza no Brasil. **Revista de Economia Política**, v. 38, n. 4, p. 749-765, 2018.

RIOS-AVILA, F. Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition. **The Stata Journal**, v. 20, n. 1, p. 51-94, 2020.

RODRIGUES, J. D.; CRUZ, M. S.; PAIXÃO, A. N. Uma análise da prevenção do câncer de mama no Brasil. **Ciência & Saúde Coletiva (online)**, v. 20, n. 10, p. 3163-3176, 2015. Disponível em: <https://www.scielo.br/j/csc/a/FhNNWR8rXswhXgnL7QYzk7F/?lang=pt>.

SANTOS, M. M.; HERMETO, A. M. Uma abordagem de decomposição quantílica para diferenciais regionais de salários no Brasil metropolitano. **Pesquisa e Planejamento Econômico**, v. 49, n. 1, p. 99-132, 2019.

SAPPLETON, N. Women non-traditional entrepreneurs and social capital. **International Journal of Gender and Entrepreneurship**, v. 1, n. 3, p. 192-218, 2009.

SCHLÖSSER, A. Elementos caracterizadores das representações sociais da amizade para universitários. **Revista de Psicologia**, Fortaleza, v. 11, n. 1, p. 12-19, 2020.

SHULTZ, T. Investment in human capital. **The American Economic Review**, v. 51, n. 1, p. 1-17, 1961.

SILVA, M. F. G. **Cooperação, capital social e desempenho econômico**: um estudo analítico. Relatório de pesquisa. São Paulo: [s.n.], 2001.

SILVA, V. H. M. C.; FRANÇA, J. M. S. Decompondo o diferencial regional de salários entre Sudeste e Nordeste: uma aplicação da abordagem quantílica incondicional. **Revista Econômica do Nordeste**, v. 47, n. 3, p. 109-129, 2016.

SOUZA, L. K.; HUTZ, C. S. Relacionamentos pessoais e sociais: amizade em adultos. **Psicologia em Estudo**, v. 13, n. 2, p. 257-265, 2008.

SZWARCWALD, C. L. *et al.* Pesquisa Nacional de Saúde no Brasil: concepção e metodologia de aplicação. **Ciência & Saúde Coletiva**, v. 19, n. 2, p. 333-342, 2014.

TREADWELL, H. M. Wages and women in health care: the race and gender gap. **American Journal of Public Health**, v. 109, n. 2, p. 208-209, 2019.

VIEIRA, L. A.; RIBEIRO, L. C.; ZELLER, C. B. Capital social e saúde: associação entre *proxys* de capital social e a taxa de mortalidade por causas externas por agressão e por lesões autoprovocadas em municípios brasileiros com mais 100.000 habitantes. *In*: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 18., 2012, Águas de Lindóia. **Anais... Águas de Lindóia**: Abep, 2012.

WOOLDRIDGE, J. M. **Introdução à econometria**: uma abordagem moderna. 3. ed. São Paulo: Cengage Learning, 2016.

Originais submetidos em: jul. 2020.

Última versão recebida em: jun. 2021.

Aprovada em: jun. 2021.

RETORNO SALARIAL E DOENÇAS CARDIOVASCULARES: EVIDÊNCIAS PARA O CASO BRASILEIRO¹

Victor Rodrigues de Oliveira²

Giácomo Balbinotto Neto³

Paulo de Andrade Jacinto⁴

Eduardo da Silveira⁵

O objetivo deste artigo é analisar a relação entre doenças cardiovasculares e salários no Brasil, com base nos dados da Pesquisa Nacional de Saúde 2013. Esta relação é analisada adotando-se uma estratégia empírica em dois estágios. Em primeiro lugar, estima-se um *propensity score matching* para avaliar a probabilidade de um trabalhador reportar ter alguma doença cardiovascular. No segundo passo, é estimado um modelo de regressão quantílica incondicional, em que o escore de propensão é utilizado como peso para garantir as condições de balanceamento nos grupos de suporte comum. Os resultados indicaram que indivíduos que reportaram ter alguma doença cardiovascular têm perdas salariais em torno de 11%. As evidências encontradas indicam que a prevenção primária de doenças cardiovasculares melhorará os resultados no mercado de trabalho. Além disso, a prevenção e o tratamento secundários, se puderem reduzir as complicações da doença, também poderão ter um impacto significativo sobre o salário e o mercado de trabalho.

Palavras-chave: economia da saúde; doenças cardiovasculares; salários; regressão quantílica incondicional.

SALARY RETURN AND CARDIOVASCULAR DISEASES: EVIDENCE FOR THE BRAZILIAN CASE

The purpose of this article is to analyze the relationship between cardiovascular diseases and wages in Brazil based on data from the National Health Survey 2013. This relationship is analyzed using a two-stage empirical strategy. First, a propensity score matching is estimated to assess the probability of a worker reporting having a cardiovascular disease. In the second step, an unconditional quantile regression model is estimated, in which the propensity score is used as a weight to ensure the conditions of balance in the common support groups. The results indicated that individuals who reported having a cardiovascular disease have wage losses around 11%. The evidence found indicates that primary prevention of cardiovascular disease will improve outcomes in the labor market. Furthermore, secondary prevention and treatment, if they can reduce the complications of the disease, can also have a significant impact on wages and the labor market.

Keywords: health economics; heart problems; wage; unconditional quantile regression.

JEL: C31; I10; J31.

1. DOI: <http://dx.doi.org/10.38116/ppe52n3art5>

2. Professor do Programa de Pós-Graduação em Desenvolvimento Econômico da Universidade Federal do Paraná (PPGDE/UFPR). *E-mail:* victoroliv.rod@ufpr.br.

3. Professor do Programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande do Sul (PPGE/UFRGS). *E-mail:* giacomo.balbinotto@ufrgs.br.

4. Professor da UFPR. *E-mail:* paulo.jacinto@ufpr.br.

5. Mestre em economia pela UFRGS. *E-mail:* edusilveira74@gmail.br.

1 INTRODUÇÃO

Informações do relatório da Organização Mundial da Saúde (OMS) mostram que, em 2019, as doenças crônicas não transmissíveis⁶ (DCNTs) foram responsáveis por cerca de 32 milhões de mortes no mundo (WHO, 2016b). Em termos relativos, esse número representa 56,60% do total de óbitos. Em 2000, o número de óbitos era de aproximadamente 26 milhões e correspondia a 49,70% das mortes totais. Assim, em quinze anos, houve um aumento expressivo de óbitos decorrentes de DCNTs.

Entre esse conjunto de doenças, as doenças cardiovasculares⁷ (DCVs) tiveram a maior participação nas mortes, passando de 27,70%, em 2000, para 31,30%, em 2015. Na Europa, as DCVs representaram 49% de todos os óbitos em 2015. No Brasil, como pode ser visualizado no gráfico 1, os números não são diferentes. Cerca de 44,83% dos óbitos registrados no país tiveram como causa as DCNTs contra 46,08% em 2000. Por grupo de doenças, as DCVs foram responsáveis por 27,94% das mortes em 2000 *vis-à-vis* 31,53% em 2015.⁸

As DCVs são a principal causa de morte na população brasileira, sendo responsáveis por, pelo menos, 20% das mortes na população com mais de 30 anos de idade (Mansur *et al.*, 2001; Mansur e Favarato, 2012; 2016; Brant e Ribeiro, 2017). Esse conjunto de doenças tem respondido por grande parte das mortes antes dos 70 anos de idade e da perda de qualidade de vida, gerando incapacidades e alto grau de limitação das pessoas doentes em suas atividades de trabalho e lazer, de modo a exercer significativa pressão sobre os serviços de saúde (Schmidt *et al.*, 2011; Rtveldze *et al.*, 2013; Ribeiro *et al.*, 2016).

Os custos econômicos estimados das DCVs no Brasil, de acordo com Siqueira, Siqueira-Filho e Land (2017), foram de R\$ 37,1 bilhões em 2015, um aumento de 17% no período 2010-2015. Os custos da morte prematura por DCVs representaram 61% do custo total da DCV, os custos diretos com hospitalizações e consultas foram 22%, e os custos associados à perda de produtividade relacionada à doença foram 15% do total. Em termos do produto interno bruto (PIB), o custo

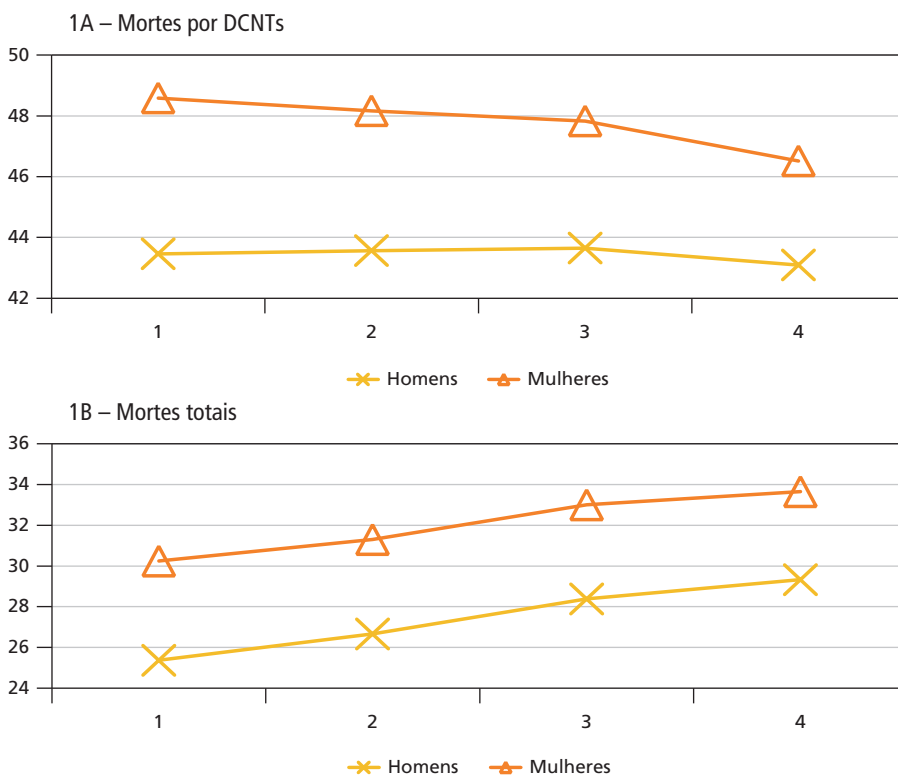
6. As DCNTs são doenças multifatoriais que se desenvolvem no decorrer da vida e são de longa duração. As quatro DCNTs de maior impacto mundial são: doenças cardiovasculares, diabetes, câncer e doenças respiratórias crônicas.

7. As DCVs são um grupo de doenças do coração e dos vasos sanguíneos, que incluem: doença cardíaca coronária; doença cerebrovascular; doença arterial periférica; doença cardíaca reumática; doença cardíaca congênita; e trombose venosa profunda e embolia pulmonar. Para detalhes, ver WHO (2014).

8. A preponderância de DCVs nas causas de mortalidade globalmente e o compartilhamento de seus fatores de risco com os de outras doenças crônicas nortearam a formulação de estratégias preventivas pela World Health Organization (WHO) para o enfrentamento das DCVs (WHO, 2005). O impacto socioeconômico das DCVs está ameaçando o progresso das Metas de Desenvolvimento do Milênio, incluindo a redução da pobreza, a equidade, a estabilidade econômica e a segurança humana, podendo atuar como um freio no próprio desenvolvimento econômico das nações (WHO, 2011). O debate internacional sobre esse assunto tem sido intenso, culminando em Reunião de Alto Nível na Assembleia da Organização das Nações Unidas (ONU), em setembro de 2011 (United Nations, 2012). A declaração desse encontro, ao afirmar que a carga das DCVs e seu impacto constituem um dos grandes desafios para o desenvolvimento no século XXI, reconheceu o papel e a responsabilidade primordial dos governos na resposta ao desafio. Contudo, é necessário um engajamento de todos os setores da sociedade para gerar respostas acertadas na prevenção e controle das DCVs (United Nations, 2012).

médio das DCVs foi estimado em 0,7% do PIB. Comparativamente, o estudo de Balbinotto Neto e Silva (2008) avaliou os custos das DCVs no Brasil para o ano de 2003 e concluiu que a perda chegou a 1,74% do PIB.

GRÁFICO 1
Óbitos por doenças cardiovasculares – Brasil (2000-2015)
(Em %)



Fontes: WHO (2016a) e United Nations (2016).
Elaboração dos autores.

As perdas relacionadas ao mercado de trabalho podem ser analisadas sob diversas óticas. A literatura indica que há relação entre o tipo de trabalho realizado tanto no que diz respeito à prevalência quanto à incidência das DCVs (Gaziano, Galea e Reddy, 2007; Eller *et al.*, 2009; Gaziano *et al.*, 2010; Song *et al.*, 2015; Dugani e Gaziano, 2016). Indivíduos que sofrem destes tipos de problemas de saúde apresentam, frequentemente, uma redução em sua produtividade e uma consequente diminuição em seus rendimentos do trabalho. Esse problema pode ser exacerbado quando o trabalhador precisa se afastar do trabalho por longos períodos, o que pode se traduzir na perda do emprego e de oportunidades de promoção, em

defasagem técnica e de treinamento em relação ao restante da força de trabalho, em absenteísmo ou, ainda, em aposentadoria precoce.

Os trabalhadores afetados não só podem abandonar a força de trabalho, como também impõem uma carga de cuidados de saúde aos outros, reduzindo a renda e o crescimento, por exemplo, por meio de tributos distorcidos. Os trabalhadores com menos saúde ainda realizam suas funções, mas com um nível de esforço reduzido, uma vez que a doença pode afetar a ambição e a durabilidade do trabalho, o que, por sua vez, reduz a produtividade média dos trabalhadores. A produtividade do trabalhador não só tem impacto no nível de renda, mas também na taxa de crescimento, uma vez que a produtividade do trabalhador afeta imediatamente o retorno do investimento, a acumulação de capital e, eventualmente, o crescimento da renda. Além da morbidade, a expectativa de vida em si pode ter um impacto nos níveis de renda e no crescimento. Visto que a educação é um custo fixo no estágio inicial da vida de um indivíduo que precisa ser coberto por prêmios salariais em todas as fases posteriores de seu ciclo de vida profissional, o retorno à acumulação de capital humano é negativamente afetado por uma redução na expectativa de vida útil. Uma menor taxa de acumulação de capital humano, em seguida, leva a um menor crescimento da renda.

As evidências empíricas sugerem que o efeito da depreciação da saúde é a queda na oferta de mão de obra. A redução de incentivos para a acumulação de capital humano ou físico ou o aumento dos gastos com saúde do indivíduo também podem decorrer da presença de doenças. De fato, uma literatura empírica considerável indica os vários impactos negativos das DCVs no nível microeconômico, como a redução dos salários, da probabilidade de ser empregado, das horas ofertadas, entre outras (Chirikos e Nestel, 1985; Mitchell e Butler, 1986; Mitchell e Burkhauser, 2006; Currie e Madrian, 1999; Kidd, Sloane e Ferko, 2000; Wilson, 2001; Pelkowski e Berger, 2004; Cai e Kalb, 2006; García e Nicolás, 2006; Adeyi, Smith e Robles, 2007).

Para a Austrália, Harris (2009) encontrou evidências de que um evento cardiovascular reduz a participação no mercado de trabalho para homens em 13% e em 9% para mulheres. Com as estimativas corrigidas para a endogeneidade das DCVs, esses percentuais mudam para 13,60% e 9,60%, respectivamente.

Evidências da redução do crescimento econômico em decorrência da prevalência e da mortalidade oriundas das DCVs são apresentadas por Suhrcke e Urban (2010) e mostraram que há impacto para os países de renda alta, mas sem impacto nos países de baixa e média rendas. As evidências empíricas da relação entre as DCVs e o mercado de trabalho estão concentradas nos países desenvolvidos, com pouca ou quase nenhuma literatura para os países de renda média.

Nesse contexto, as DCVs impõem limitações à qualidade de vida relacionadas a aspectos físicos, sociais, financeiros e de saúde dos indivíduos. Consistente com os critérios empregados pela literatura sobre economia da saúde, adotou-se a abordagem do capital humano para estimar as perdas salariais decorrentes do ônus imposto por esse conjunto de doenças. O objetivo deste estudo é contribuir com a escassa literatura sobre o tema, mensurando a perda de rendimento dos trabalhadores brasileiros oriunda de DCVs a partir dos microdados da Pesquisa Nacional de Saúde (PNS) 2013. Computa-se essa relação ao se utilizar uma estratégia empírica em dois passos. Em primeiro lugar, estima-se um *propensity score matching* (PSM) para avaliar a probabilidade de um trabalhador reportar ter alguma DCV. Em segundo, estima-se um modelo de regressão quantílica incondicional, em que o escore de propensão é utilizado como peso.

Os principais resultados indicam que, ao longo da distribuição de rendimentos, a prevalência de DCVs está negativamente correlacionada aos salários de homens, com reduções de até 16%, e de mulheres, com reduções de até 20%. Além desta introdução, este artigo é composto por mais quatro seções. A seção 2 descreve a base de dados. A seção 3 apresenta a metodologia empírica utilizada. A seção 4 expõe e discute os resultados obtidos, e a seção 5 traz as considerações finais.

2 BASE DE DADOS

Neste estudo, foram utilizados os microdados da PNS 2013, realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Com desenho próprio, elaborado, especificamente, para coletar informações sobre saúde da população brasileira em 2013, a PNS foi planejada para a estimação de vários indicadores com a precisão desejada e para assegurar a continuidade no monitoramento da grande maioria dos indicadores do Suplemento Saúde da Pesquisa Nacional por Amostra de Domicílios (PNAD). A elaboração da PNS foi fundamentada em três eixos principais: o desempenho do sistema nacional de saúde; as condições de saúde da população brasileira; e a vigilância das DCNTs e dos fatores de risco associados. A pesquisa foi representativa para: Brasil, áreas urbanas e rurais; Grandes Regiões; Unidades da Federação (UFs); e capitais dos estados.

A utilização da PNS mostrou-se relevante devido à sua representatividade nacional e pela disponibilidade de variáveis de saúde, de dados antropométricos e nutricionais e de informações do mercado de trabalho. A principal vantagem no uso da PNS consiste em observar dados atualizados para os problemas cardiovasculares no Brasil. Até o momento, a PNS abrange os microdados mais recentes das características de saúde e rendimentos.

Diferentemente de pesquisas nas quais o entrevistado reporta muitas de suas características, a PNS se preocupou em levar os instrumentos para a aferição de

peso, de altura, de circunferência da cintura e de pressão arterial,⁹ seguindo os mesmos procedimentos da Pesquisa de Orçamentos Familiares (POF) 2008-2009.¹⁰ Este cuidado na obtenção de dados é fundamental, uma vez que a notificação incorreta de dados autorreportados, quando sistemática e não aleatória, produz estimativas inconsistentes e compromete os resultados e a robustez das pesquisas.

Para tornar a amostra homogênea e estimar corretamente a taxa de salários, restringiu-se a análise aos indivíduos com idade entre 25 e 64 anos, que haviam trabalhado na última semana (com rendimentos diferentes de zero), que não estavam ocupados na agricultura nem na administração pública,¹¹ que não estavam gestantes e para as quais existem informações relativas a número de anos de estudo.

O estudo proposto promoveu o ajuste da amostra visando respeitar o plano amostral que lhe deu origem. A pesquisa é domiciliar, e o plano amostral empregado foi a amostragem conglomerada em três estágios, com estratificação das unidades primárias de amostragem. Os setores censitários ou o conjunto de setores formam as unidades primárias de amostragem, os domicílios são as unidades de segundo estágio e os moradores com 18 anos ou mais de idade definem as unidades de terceiro estágio.¹²

A PNS estratificou as DCVs em quatro grupos: infarto, que é a obstrução ou entupimento das artérias coronárias que irrigam o coração; angina, que é a dor ou o desconforto no peito; insuficiência cardíaca, que corresponde à situação em que o coração deixa de bombear o sangue na quantidade necessária à manutenção do corpo; e outras doenças do coração, por exemplo, por incapacidade das artérias coronárias, quando obstruídas, de conduzirem adequadamente o oxigênio indispensável para o trabalho do músculo cardíaco (cardiopatia coronariana). Em virtude do pequeno número de observações para cada uma das DCVs em separado, a análise será realizada para o conjunto das DCVs.

Em decorrência das características biológicas distintas entre homens e mulheres, as quais influenciam a predisposição e a resistência a certos tipos de doenças, as estimativas serão produzidas separadamente por gênero. As doenças ocupacionais afetam distintamente homens e mulheres devido aos tipos de trabalho tradicionalmente efetuados por ambos. A exposição ao risco de adoecimento, sobretudo decorrente do tipo de trabalho, dos hábitos de vida e de consumo, a percepção da saúde e a decisão de busca pelos serviços ambulatoriais, médicos e hospitalares

9. São eles, respectivamente: balança eletrônica portátil, estadiômetro portátil, fita de inserção e aparelho de pressão digital.

10. Para mais detalhes, ver IBGE (2010).

11. Foram desconsideradas mulheres desses segmentos devido à dinâmica própria que rege a administração pública e o setor agrícola da economia brasileira.

12. Para detalhes, ver Souza-Júnior *et al.* (2015).

também têm sido destacadas como causas das diferenças por gênero na utilização de serviços de saúde (Ven e Ellis, 2000).

Os estudos em economia da saúde também destacam que os homens tendem, em média, a reportar menos problemas de saúde – prevalência de doenças crônicas e número médio de dificuldades motoras. A demanda das mulheres por serviços médicos geralmente é mais elevada do que a dos homens na idade adulta, fenômeno associado ao período fértil e à maior percepção feminina de sua saúde. Na velhice, a demanda masculina ultrapassa a feminina, fato relacionado ao desgaste dos homens no trabalho, o que leva a maior prevalência de doenças nessa fase da vida e, geralmente, à morte com idade menos avançada do que as mulheres. De forma geral, as mulheres utilizam mais os serviços ambulatoriais e preventivos, enquanto os homens tendem a usar mais os serviços curativos. Outrossim, a população masculina possui taxas de internação maiores nas idades adultas avançadas e maior tempo de permanência por internação (Cai e Kalb, 2006; Gaziano, Galea e Reddy, 2007; Zweifel, Breyer e Kifmann, 2009; Jäckle e Himmler, 2010; Johar e Katayama, 2012).

3 MÉTODO

Explorar e identificar a correlação entre a presença de DCVs e os resultados no mercado de trabalho exigem considerar os efeitos de causalidade e também os possíveis efeitos de causalidade reversa. Além dos casos em que há viés de simultaneidade, a endogeneidade ainda pode ser decorrente do viés de variável relevante omitida ou por meio de erros de mensuração. Em suma, a prevalência de DCV pode não ser exógena em modelos salariais.

A literatura de economia da saúde já propôs diferentes formas para corrigir o problema da endogeneidade.¹³ O método mais empregado é o uso de variáveis instrumentais. O objetivo é encontrar um instrumento altamente correlacionado com a variável indicadora de DCV (relevância do instrumento) e que seja ortogonal ao erro na equação de salários (exogeneidade do instrumento).

Contudo, as dificuldades na obtenção de instrumentos válidos dão origem a críticas que colocam em dúvida a hipótese de exogeneidade do instrumento. Por exemplo, Atella, Pace e Vuri (2008) argumentam que a melhor estratégia seria utilizar as informações de irmãos (gêmeos) de mesmo sexo. Contudo, isto nem sempre é possível.

13. Gortmaker *et al.* (1993), Cawley (2004) e Larose *et al.* (2016) utilizaram o peso anterior à entrada do indivíduo no mercado de trabalho. Se o salário tem impacto sobre o peso e, conseqüentemente, o aparecimento de DCV, a estratégia consiste em utilizar uma medida pretérita. A literatura também utilizou informações de irmãos para determinar se os resultados refletiriam o ambiente familiar sobre o peso, e o posterior desenvolvimento de um problema cardíaco, ou se, de fato, deve-se à relação entre peso e *status* econômico. Nesse caso, o propósito seria reduzir muitos dos fatores não observados genéticos e do ambiente por meio de efeitos fixos (Averett e Korenman, 1996; Conley e Glauber, 2007).

Em muitas pesquisas domiciliares, como é a PNS, não é possível identificar irmãos que não moram no mesmo domicílio. E muitos que moram, estudam ou não estão trabalhando, o que configura um instrumento fraco em muitos casos. É interessante registrar que em diversos estudos que aplicaram variáveis instrumentais para lidar com a endogeneidade da variável de saúde em modelos de salário, como Cawley (2004), Morris (2006), Shimokawa (2008), entre outros, os testes de Hausman não rejeitaram a hipótese de que os coeficientes de mínimos quadrados ordinários (MQO) e de variáveis instrumentais (IV) são iguais.¹⁴

Com intuito de contornar os possíveis problemas de endogeneidade, foi utilizada uma estratégia empírica em duas etapas. Na primeira etapa, utilizou-se um PSM para estimar a probabilidade de se ter DCV. Na segunda, fez-se uso do escore de propensão como peso nas regressões quantílicas.

3.1 Propensity score matching

Utilizando-se do caráter genético da DCV (predisposição e obesidade hereditária), usou-se o método PSM para estimar a probabilidade de se ter DCV, considerando como controle as características observáveis, como variáveis de perfil antropométrico, nutricional e hábitos de saúde do próprio indivíduo.

O quadro 1 faz uma descrição das variáveis de controle utilizadas na estimação do PSM.

QUADRO 1
Descrição das variáveis

Variável	Descrição
Perfil antropométrico, nutricional e hábitos de saúde	
Altura	Altura média (metros).
Peso	Peso médio (kg).
Índice de Massa Corporal (IMC)	i) baixo peso (IMC menor que 18,5); ii) peso ideal (IMC de 18,51 a 24,99); iii) sobrepeso (IMC de 25 a 29,99); iv) obesidade 1 (IMC de 30 a 34,9); v) obesidade 2 (IMC de 35 a 39,9); e vi) obesidade 3 (IMC maior que 40).
Cintura	Homens: i) cintura normal (até 94 cm); ii) risco médio (de 94 cm até 98 cm); iii) risco alto (de 98 cm até 102 cm); e iv) risco altíssimo (maior que 102 cm). Mulheres: i) cintura normal (até 80 cm); ii) risco médio (de 81 cm até 83 cm); iii) risco alto (de 84 cm até 87 cm); e iv) risco altíssimo (maior que 87 cm).
Pressão arterial (sistólica versus diastólica)	i) normal: 120 x 80 ou menor; ii) pré-hipertensão: de 121 a 139 x 81 a 89; iii) hipertensão estágio 1: de 140 a 159 x 90 a 99; iv) hipertensão estágio 2: de 160 a 179 x 100 a 109; e v) hipertensão estágio 3: acima de 180 x 110.
Índice de hábitos de vida	Consumo de saladas, legumes, sucos naturais, frutas, leite e sal; prática de exercícios físicos; faz dieta; tabagismo presente e pretérito.
Plano de saúde	1 se tem plano de saúde; 0, caso contrário.

(Continua)

14. Isso pode indicar que não há problemas de endogeneidade ou sugerir a falta de instrumentos adequados.

(Continuação)

Variável	Descrição
Fatores demográficos, educacionais e geográficos	
In(salário)	Logaritmo natural do rendimento bruto do último mês proveniente do trabalho principal.
Idade	Idade (anos).
Escolaridade	Faixas de estudo: i) sem instrução; ii) fundamental incompleto ou equivalente; iii) fundamental completo ou equivalente; iv) médio incompleto ou equivalente; v) médio completo ou equivalente; vi) superior incompleto ou equivalente; e vii) superior completo.
Raça/cor	1 se branco; 0, caso contrário. 1 se preto; 0, caso contrário. 1 se amarelo; 0, caso contrário.
Cônjuge	1 se casado; 0, caso contrário.
Filhos	1 se tem filhos; 0, caso contrário.
Unidade da Federação	Variáveis <i>dummy</i> para UFs.

Fonte: Microdados da PNS 2013.

Elaboração dos autores.

Obs.: Este quadro apresenta as variáveis utilizadas no cômputo do PSM.

O escore de propensão é definido como a probabilidade de um indivíduo ter uma DCV dadas as características observáveis, $[P(X)_i = Pr[DCV_i = 1|X_i]$, e será utilizado para computar os pesos para balancear os indivíduos, tornando o grupo de controle (não ter a DCV), em média, similar ao de tratado (ter a DCV).

A utilização do escore de propensão toma como base duas hipóteses-chave. A primeira, a seleção nos observáveis, requer que a participação no programa seja independente dos resultados, condicional nas covariáveis. A segunda hipótese refere-se à existência de um suporte comum. Esta condição requer que existam unidades de ambos os grupos, tratamento e controle, para cada característica X para a qual se deseja comparar. Estas hipóteses são testadas verificando a área de sobreposição da distribuição de probabilidade de tratamento de ambos os grupos e por meio de um teste t para igualdade de médias entre os grupos de controle e de tratamento, antes e após o pareamento.

A estimação do escore de propensão permitirá atribuir pesos diferenciados para os trabalhadores do contrafactual de acordo com as características e probabilidades de se ter um diagnóstico de DCV. A metodologia é robusta por ter a vantagem de conduzir a um ganho adicional de robustez devido à remoção do efeito direto das variáveis omitidas (regressão) e ao reduzir a correlação entre as variáveis omitidas e incluídas (reponderação). Esses pesos serão empregados como ponderadores no modelo de regressão quantílica incondicional. O intuito é manter a parcimônia dos modelos, a fim de atingir as condições de balanceamento nos grupos de suporte comum do PSM.

3.2 Regressão quantílica incondicional

O objetivo é avaliar a existência de uma relação entre a prevalência de DCVs e os salários ao longo da distribuição da variável dependente por meio de um modelo de regressão quantílica incondicional, seguindo a proposta de Firpo, Fortin e Lemieux (2009). Este método representa um avanço na literatura sobre o tema, pois permite a obtenção do efeito marginal de uma covariada sobre a variável dependente de forma direta, principalmente quando as variáveis independentes são discretas. Desse modo, seja o τ -ésimo, $0 < \tau < 1$, quantil de uma variável aleatória Y com função de distribuição $F_Y(y) = P(Y \leq y)$, isto é,

$$q_\tau := F^{-1}(\tau) = \inf\{y: F_Y(y) \geq \tau\}. \quad (1)$$

Observe que q_τ é uma estatística de ordem e pode ser obtida pela minimização de uma função perda (linear) assimétrica:

$$\frac{1}{N} [\tau \sum_{i: y_i \geq q} |y_i - q| + (1 - \tau) \sum_{i: y_i < q} |y_i - q|] = \frac{1}{N} \sum \rho_\tau(y_i - q), \quad (2)$$

em que $\rho_\tau(\varepsilon) = (\tau - 1\{\varepsilon < 0\})\varepsilon$ é a função *check* e ε é o termo de erro.

O método de Firpo, Fortin e Lemieux (2009) consiste da regressão de uma transformação da variável de resultado sobre as variáveis explicativas e é construído sob o conceito de função de influência recentrada (RIF), como segue:

$$RIF(Y; v, F_Y(y)) = \left[q_\tau + \frac{\tau - 1}{f_Y(q_\tau)} \right] + \left[\frac{1}{f_Y(q_\tau)} \right] 1\{Y \geq q_\tau\}. \quad (3)$$

Para estimar a nova variável, como apresentada em (3), $f_Y(q_\tau)$ deve ser estimado em um passo anterior. Nesse caso, obtém-se a densidade de Y por meio de um estimador não paramétrico, como segue:

$$\hat{f}_Y(\hat{q}_\tau) = \frac{1}{N} \sum \frac{1}{h} K_Y\left(\frac{Y_i - \hat{q}_\tau}{h}\right), \quad (4)$$

em que K_Y é a função Kernel e h é a *bandwidth*.

Assim, $\hat{f}_Y(\hat{q}_\tau)$ é estimada por (4) e o resultado de interesse é estimado por uma regressão de $RIF(Y; \hat{q}_\tau)$ sobre X por MQO.

4 RESULTADOS

Antes de verificar a relação entre as DCVs e os salários, as duas hipóteses básicas do PSM¹⁵ devem ser checadas. Para verificar se os indivíduos tratados (ter a DCV) e os indivíduos pareados são comparáveis, é necessário realizar um teste de médias para as diferentes variáveis. As tabelas 1 e 2 mostram que, pelo teste t , as covariadas estão balanceadas adequadamente, com redução substancial do viés. Para o teste conjunto de variáveis, os resultados também indicaram redução do valor médio do viés, com diminuição significativa da estatística de teste qui-quadrado.

15. As estimativas do PSM estão dispostas na tabela A.1 no apêndice A.

O *p*-valor indica que a hipótese nula de igualdade de médias entre os grupos não pode ser rejeitada.

TABELA 1
Teste para igualdade de médias para grupo de tratados e de controle antes e depois do matching: homens (2013)

Variáveis	Antes do matching			Depois do matching		
	Controle	Tratado	$p > t $	Controle	Tratado	$p > t $
Altura	170,02	170,46	0,35	170,02	170,39	0,57
Peso	81,11	77,15	0,00	81,11	81,80	0,64
IMC: peso ideal	0,30	0,38	0,02	0,30	0,28	0,51
IMC: sobrepeso	0,37	0,43	0,05	0,37	0,37	1,00
IMC: obesidade 1	0,20	0,14	0,01	0,20	0,24	0,21
IMC: obesidade 2	0,11	0,03	0,00	0,11	0,10	0,67
IMC: obesidade 3	0,01	0,01	0,27	0,01	0,02	0,52
Cintura: risco médio	0,12	0,12	0,97	0,12	0,11	0,69
Cintura: risco alto	0,07	0,11	0,10	0,07	0,09	0,63
Cintura: risco altíssimo	0,39	0,21	0,00	0,39	0,44	0,22
Pré-hipertensão	0,16	0,21	0,03	0,16	0,18	0,56
Hipertensão estágio 1	0,06	0,05	0,51	0,06	0,04	0,31
Hipertensão estágio 2	0,02	0,01	0,06	0,02	0,03	0,43
Hipertensão estágio 3	0,01	0,01	0,69	0,01	0,00	0,56
Índice de hábitos de vida	0,30	0,29	0,05	0,30	0,31	0,39
Plano de saúde	0,24	0,25	0,60	0,24	0,26	0,49
Idade	48,29	40,78	0,00	48,29	48,45	0,86
Fundamental incompleto	0,34	0,27	0,01	0,34	0,33	0,65
Fundamental completo	0,12	0,11	0,52	0,12	0,15	0,26
Médio incompleto	0,03	0,05	0,06	0,03	0,01	0,36
Médio completo	0,21	0,27	0,02	0,21	0,16	0,15
Superior incompleto	0,02	0,04	0,08	0,02	0,03	0,43
Superior completo	0,14	0,11	0,09	0,14	0,21	0,05
Branco	0,48	0,40	0,01	0,48	0,55	0,12
Preto	0,51	0,60	0,05	0,51	0,45	0,14
Chefe	0,75	0,73	0,56	0,75	0,79	0,22
Casado	0,52	0,46	0,04	0,52	0,52	0,93
Filhos	0,51	0,67	0,01	0,51	0,48	0,76
Amostra	R ²	LR	p>chi2	Viés médio	Viés mediano	-
Unmatched	0,09	237,68	0,00	12,40	9,50	-
Matched	0,05	35,45	0,99	4,80	5,10	-

Fonte: PNS 2013. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/saude/9160-pesquisa-nacional-de-saude.html?=&t=-microdados>. Acesso em: 29 nov. 2016.

Elaboração dos autores.

Obs.: Esta tabela apresenta um teste de diferença de médias para verificar a aderência do PSM. Controles adicionais incluem altura ao quadrado, peso corporal ao quadrado, idade ao quadrado e variáveis *dummy* de estado.

TABELA 2
Teste para igualdade de médias para grupo de tratados e controle antes e depois do matching: mulheres (2013)

Variáveis	Antes do matching			Depois do matching		
	Controle	Tratado	$p > t $	Controle	Tratado	$p > t $
Altura	157,82	158,77	0,03	157,82	158,13	0,60
Peso	68,95	67,78	0,18	68,95	69,38	0,74
IMC: peso ideal	0,36	0,39	0,34	0,36	0,35	0,93
IMC: sobrepeso	0,33	0,36	0,22	0,33	0,33	1,00
IMC: obesidade 1	0,18	0,16	0,23	0,18	0,18	0,82
IMC: obesidade 2	0,09	0,06	0,02	0,09	0,11	0,46
IMC: obesidade 3	0,02	0,02	0,68	0,02	0,02	1,00
Cintura: risco médio	0,11	0,12	0,40	0,11	0,11	0,89
Cintura: risco alto	0,13	0,12	0,66	0,13	0,13	1,00
Cintura: risco altíssimo	0,56	0,50	0,05	0,56	0,58	0,72
Pré-hipertensão	0,15	0,14	0,42	0,15	0,14	0,71
Hipertensão estágio 1	0,04	0,03	0,15	0,04	0,07	0,24
Hipertensão estágio 2	0,01	0,01	0,19	0,01	0,01	1,00
Hipertensão estágio 3	0,01	0,00	0,07	0,01	0,00	0,32
Índice de hábitos de vida	0,33	0,32	0,16	0,33	0,34	0,85
Plano de saúde	0,27	0,29	0,56	0,27	0,29	0,70
Idade	45,35	40,08	0,00	45,35	45,40	0,96
Fundamental incompleto	0,27	0,22	0,04	0,27	0,29	0,77
Fundamental completo	0,14	0,11	0,13	0,14	0,14	1,00
Médio incompleto	0,07	0,06	0,21	0,07	0,06	0,48
Médio completo	0,25	0,33	0,01	0,25	0,23	0,61
Superior incompleto	0,02	0,04	0,21	0,02	0,01	0,48
Superior completo	0,10	0,14	0,06	0,10	0,13	0,40
Branco	0,44	0,41	0,36	0,44	0,41	0,42
Preto	0,54	0,58	0,23	0,54	0,57	0,48
Chefe	0,55	0,56	0,69	0,55	0,48	0,13
Casado	0,40	0,38	0,35	0,40	0,44	0,42
Filhos	0,48	0,71	0,00	0,48	0,43	0,51
Amostra	R ²	LR	p>chi2	Viés médio	Viés mediano	-
Unmatched	0,06	134,58	0,00	9,30	8,20	-
Matched	0,04	27,19	1,00	4,70	3,90	-

Fonte: PNS 2013. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/saude/9160-pesquisa-nacional-de-saude.html?=&t=microdados>. Acesso em: 29 nov. 2016.

Elaboração dos autores.

Obs.: Esta tabela apresenta um teste de diferença de médias para verificar a aderência do PSM. Controles adicionais incluem altura ao quadrado, peso corporal ao quadrado, idade ao quadrado e variáveis *dummy* de estado.

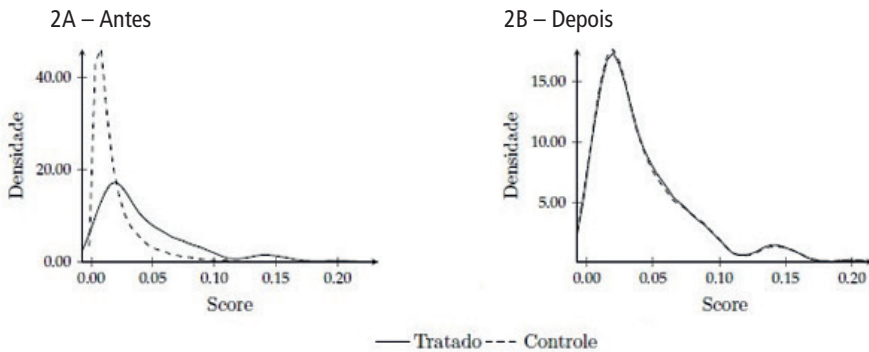
A hipótese de existência de uma área de suporte comum pode ser verificada por meio dos gráficos 2 e 3, os quais apresentam as distribuições de probabilidade do tratamento antes e depois do pareamento. Observa-se que, antes do pareamento, a distribuição de probabilidades para os indivíduos que têm DCV e os que não têm são distintas e, após o pareamento, passam a apresentar um comportamento

muito similar. O p -valor do teste de Kolmogorov-Smirnov corrobora a hipótese de igualdade das distribuições após o balanceamento pelo escore de propensão. Após verificar que as duas hipóteses básicas do PSM são atendidas, se procederá à regressão quantílica, levando em consideração o uso do escore de propensão como peso.

Uma questão relevante são os efeitos heterogêneos dos problemas cardiovasculares na distribuição de salários. A correlação entre a prevalência de DCVs e os rendimentos dos trabalhos pode ser diferente ao longo da distribuição salarial.

GRÁFICO 2

Distribuição da probabilidade do tratamento para tratados e controle: homens (2013)



Fonte: PNS 2013. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/saude/9160-pesquisa-nacional-de-saude.html?=&t=microdados>. Acesso em: 29 nov. 2016.

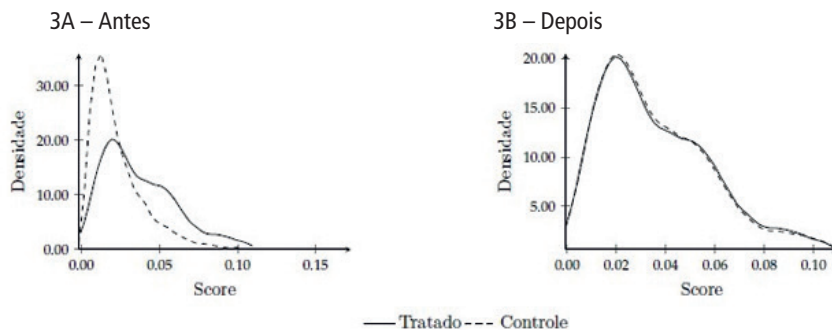
Elaboração dos autores.

Obs.: 1. Este gráfico apresenta a distribuição de probabilidade de ter DCV sobre os salários (em logaritmo).

2. Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

GRÁFICO 3

Distribuição da probabilidade do tratamento para tratados e controle: mulheres (2013)



Fonte: PNS 2013. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/saude/9160-pesquisa-nacional-de-saude.html?=&t=microdados>. Acesso em: 29 nov. 2016.

Elaboração dos autores.

Obs.: 1. Este gráfico apresenta a distribuição de probabilidade de ter DCV sobre os salários (em logaritmo).

2. Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

Por um lado, a oferta de mão de obra pode reagir de forma ambígua à redução da saúde. Enquanto o efeito substituição prevê uma maior demanda por lazer, quando o retorno do trabalho diminui, o efeito renda prevê que o indivíduo busca compensar a redução dos ganhos ao longo da vida, aumentando sua oferta de trabalho. Por outro lado, a produtividade pode diminuir diretamente, pela menor atividade física e mental, ou indiretamente, porque essas pessoas podem fazer um uso menos eficiente da tecnologia, do maquinário e dos equipamentos disponíveis. É nesse sentido que a adoção de uma abordagem tradicional, como a regressão na média, pode levar à perda de informações pertinentes que seriam obtidas por meio de uma estimação quantílica.

Os gráficos 4 e 5 reportam os resultados das estimativas econométricas para homens e mulheres, respectivamente. A variável dependente é o logaritmo natural do rendimento bruto do último mês proveniente do trabalho principal. O parâmetro central de interesse mensura a correlação entre a prevalência de DCVs e o rendimento do trabalho. Foram estimadas duas especificações: a primeira sem controles (denotada por modelo reduzida) e a segunda com controles (denotada por modelo completo). O propósito dessa estratégia é verificar se há mudanças consideráveis nos resultados que possam indicar a presença de variáveis omitidas relevantes. Para o cálculo dos erros-padrão, foram considerados *clusters* de estados. Indivíduos que vivem no mesmo estado estão sujeitos às mesmas condições do mercado de trabalho local e sofrem choques econômicos comuns e, portanto, seus rendimentos são correlacionados.

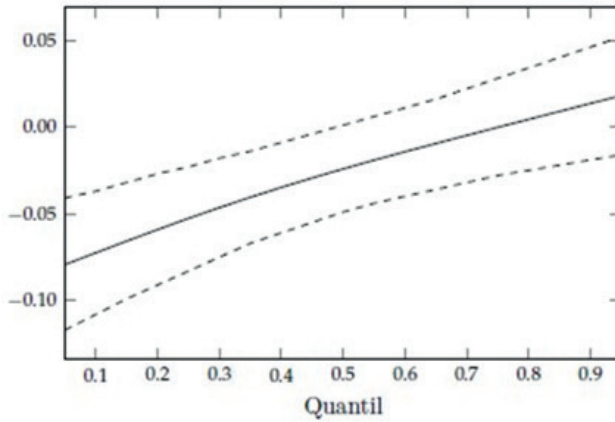
Em termos de significância estatística, observa-se que, no modelo reduzido, apenas para homens com rendimentos acima da mediana, o acometimento por alguma DCV não apresenta correlação estatisticamente diferente de zero com os rendimentos oriundos do trabalho. Na especificação do modelo completo, essa correlação não é significativa para os homens com salários superiores aos observados para o nono decil de renda. Para o grupo das mulheres, as estimativas da correlação entre salários e DCVs são estatisticamente significativas ao longo de toda distribuição de rendimentos.

Os resultados indicam que os homens com presença de DCV, quando comparados àqueles que não reportaram tal problema, têm um salário que é menor em até 12% no modelo reduzido (sem controles) e que chega a 16% no modelo completo (com controles). Para as mulheres, essa redução é de até 21% em ambos os modelos. Com base nas estimativas, observa-se que a presença de DCV pode estar impondo uma penalidade salarial, principalmente entre os trabalhadores cujo salário está abaixo da mediana da distribuição de rendimentos.

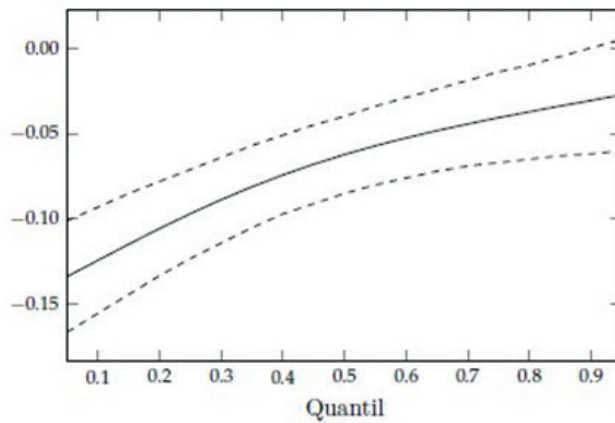
GRÁFICO 4

Relação entre doenças cardiovasculares e rendimento do trabalho: homens

4A – Modelo reduzido



4B – Modelo completo



Elaboração dos autores.

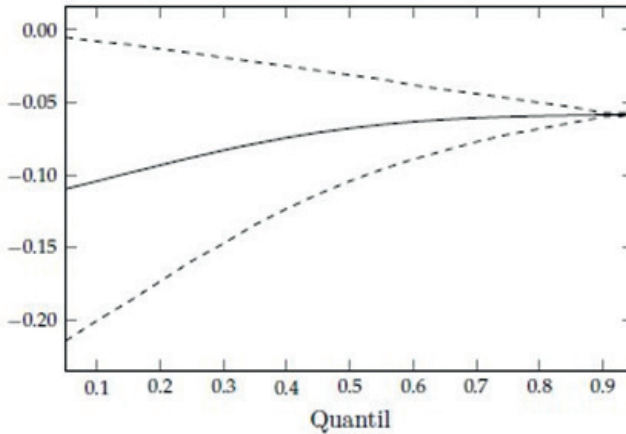
Obs.: 1. A linha pontilhada representa o intervalo de confiança de 95% obtido a partir de 1.000 replicações *bootstrap* corrigidas por *clusters* de estados. A variável dependente é o logaritmo natural do rendimento bruto do último mês proveniente do trabalho. O modelo reduzido inclui apenas a variável indicadora de DCV. O modelo completo inclui como controles os fatores demográficos, educacionais e geográficos da tabela 1. Todos os modelos são ponderados pelo escore de propensão para balancear os indivíduos e torná-los homogêneos.

2. Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

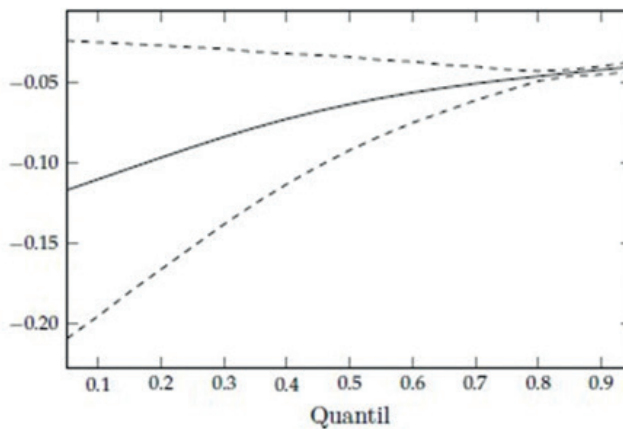
GRÁFICO 5

Relação entre doenças cardiovasculares e rendimento do trabalho: mulheres

5A – Modelo reduzido



5B – Modelo completo



Elaboração dos autores.

Obs.: 1. A linha pontilhada representa o intervalo de confiança de 95% obtido a partir de 1.000 replicações *bootstrap* corrigidas por *clusters* de estados. A variável dependente é o logaritmo natural do rendimento bruto do último mês proveniente do trabalho. O modelo reduzido inclui apenas a variável indicadora de DCV. O modelo completo inclui como controles os fatores demográficos, educacionais e geográficos da tabela 1. Todos os modelos são ponderados pelo escore de propensão para balancear os indivíduos e torná-los homogêneos.

2. Ilustração cujos leiaute e textos não puderam ser padronizados e revisados em virtude das condições técnicas dos originais (nota do Editorial).

À medida que se desloca à direita a distribuição de salários, a perda de rendimentos associada à presença de DCVs torna-se menor, em decorrência do forte impacto do grau de instrução sobre os salários. A educação tem um duplo papel, ao reduzir a taxa de depreciação do estoque de capital humano (a capacidade física e mental e as características adquiridas afetam tanto a saúde quanto a escolaridade

na mesma direção) e ao aumentar o acesso à informação. As pessoas menos impulsivas são melhores no planejamento do seu futuro e, portanto, buscam adquirir mais educação, se envolvem em comportamentos preventivos e com menor risco, e adquirem melhores hábitos de saúde e de alimentação. Esse processo está associado à preferência pelo tempo, que é responsável tanto pela quantidade de educação que uma pessoa recebe quanto por sua saúde (Grossman, 1972).

Por sua vez, a diferença nas estimativas para homens e mulheres pode estar associada ao perfil distinto de cuidados médicos desses grupos. Esse processo é resultado da interação entre a necessidade do indivíduo e a sua decisão de procurar cuidados médicos. Os determinantes dessa utilização estão associados às características dos indivíduos, ao sistema de saúde e ao padrão de prática médica. Assim como a saúde influencia a quantidade de tempo disponível para o trabalho saudável, ela também influencia a qualidade do tempo disponível para os cuidados com a saúde. O fato de uma pessoa estar saudável o suficiente para ir para o trabalho não significa necessariamente que ela está trabalhando em todo seu potencial. A perda de produtividade que ocorre quando os funcionários vêm para trabalhar, mas, como consequência de doença ou outras condições de saúde, não são totalmente funcionais é uma fonte da perda de produtividade relacionada com a saúde. Como se viu, a relação negativa encontrada aqui pode levar à redução da produção e da produtividade dos trabalhadores afetados (e também, potencialmente, à redução da produtividade dos colegas de trabalho e familiares).

Os resultados encontrados aqui estão de acordo com os dados mais recentes sobre os custos das DCVs no Brasil. Seguindo a abordagem do capital humano, Stevens *et al.* (2018) apontam que a renda perdida por indivíduo em 2015, em média, foi de 18%. A implicação deste resultado é que a prevenção primária de DCVs melhorará os resultados no mercado de trabalho, mas a prevenção e o tratamento secundários, se puderem reduzir as complicações da doença, também poderão ter um impacto significativo sobre os salários (Song *et al.*, 2015). Disso decorre que as DCVs determinam substanciais custos financeiros e perda de bem-estar no Brasil e deveriam ser uma prioridade de saúde pública e receber mais atenção.

5 CONSIDERAÇÕES FINAIS

O objetivo deste artigo foi examinar a relação entre a presença de DCVs e o rendimento salarial no mercado de trabalho brasileiro. Foram utilizados os microdados da PNS 2013. Para alcançar o objetivo proposto, foi empregada uma estratégia empírica em duas etapas. A primeira consistiu na estimação de um PSM para avaliar a probabilidade de um indivíduo reportar ter tido uma DCV. Na segunda etapa,

foi estimado um modelo de regressão quantílica em que o escore de propensão serviu como peso nas regressões.

Os resultados indicaram que indivíduos que reportaram ter alguma DCV têm perdas salariais em torno de 11%. Quando se incluem controles sociodemográficos, antropométricos e nutricionais, há um aumento da correlação entre problemas cardiovasculares e salários para homens. Também se observou que a perda salarial é decrescente para os trabalhadores com maiores rendimentos.

Os resultados obtidos são consistentes com outros trabalhos da literatura e são robustos do ponto de vista econométrico, tendo em vista o tratamento dado ao problema estatístico e ao uso de técnicas adequadas. Ademais, as evidências indicam que os custos indiretos das DCVs em termos de redução salarial são relevantes.

As implicações em termos de políticas públicas desses resultados são de que, mesmo que não seja viável eliminar essa doença crônica, especialmente a curto prazo, a redução de alguns dos fatores de risco, como exercício físico insuficiente, tabagismo e obesidade, aumentará a participação da força de trabalho e os ganhos salariais. A diminuição desses fatores de risco pode não apenas reduzir a prevalência de DCVs, mas também ter algum impacto em outras doenças crônicas graves – como câncer, artrite, asma e doença mental – que afetam os resultados do mercado de trabalho. Embora não seja explicitamente abordado aqui, também é provável que um melhor tratamento de DCVs reduza as complicações e melhore a participação, a força de trabalho e os retornos monetários.

REFERÊNCIAS

- ADEYI, O.; SMITH, O.; ROBLES, S. (Ed.). **Public policy and the challenge of chronic noncommunicable diseases**. Washington: World Bank, 2007.
- ATELLA, V.; PACE, N.; VURI, D. Are employers discriminating with respect to weight? European evidence using quantile regression. **Economics & Human Biology**, v. 6, n. 3, p. 305-329, dez. 2008.
- AVERETT, S.; KORENMAN, S. The economic reality of the beauty myth. **Journal of Human Resources**, v. 31, n. 2, p. 304-330, 1996.
- BALBINOTTO NETO, G.; SILVA, E. N. da. Os custos da doença cardiovascular no Brasil: um breve comentário econômico. **Arquivos Brasileiros de Cardiologia**, v. 91, n. 4, p. 217-218, 2008.
- BRANT, L. C. C.; RIBEIRO, A. L. P. Variations and particularities in cardiovascular disease mortality in Brazil and Brazilian states in 1990 and 2015: estimates from the Global Burden of Disease. **Revista Brasileira de Epidemiologia**, v. 20, n. 1, p. 116-128, maio 2017.

CAI, L.; KALB, G. Health status and labour force participation: evidence from Australia. **Health Economics**, v. 15, n. 3, p. 241-261, 2006.

CAWLEY, J. The impact of obesity on wages. **Journal of Human Resources**, v. 39, n. 2, p. 451-474, 2004.

CHIRIKOS, T. N.; NESTEL, G. Further evidence on the economic effects of poor health. **The Review of Economics and Statistics**, v. 67, n. 1, p. 61-69, fev. 1985.

CONLEY, D.; GLAUBER, R. Gender, body mass, and socioeconomic status: new evidence from the PSID. **Advances in Health Economics and Health Services Research**, v. 17, p. 253-275, 2007.

CURRIE, J.; MADRIAN, B. C. Health, health insurance and the labor market. *In*: ASHENFELTER, O.; CARD, D. (Ed.). **Handbook of labor economics**. Springer: Amsterdam, 1999. v. 3, p. 3309-3416.

DUGANI, S.; GAZIANO, T. A. 25 by 25: achieving global reduction in cardiovascular mortality. **Current Cardiology Reports**, v. 18, n. 10, p. 1-6, jan. 2016.

ELLER, N. H. *et al.* Work-related psychosocial factors and the development of ischemic heart disease: a systematic review. **Cardiology in Review**, v. 17, n. 2, p. 83-97, mar.-abr. 2009.

FIRPO, S.; FORTIN, N. M.; LEMIEUX, T. Unconditional quantile regressions. **Econometrica**, v. 77, n. 3, p. 953-973, maio 2009.

GARCÍA, P. G.; NICOLÁS, A. L. Health shocks, employment and income in the Spanish labour market. **Health Economics**, v. 15, n. 9, p. 997-1009, set. 2006.

GAZIANO, T. A. *et al.* Growing epidemic of coronary heart disease in low- and middle-income countries. **Current Problems in Cardiology**, v. 35, n. 2, p. 72-115, fev. 2010.

GAZIANO, T. A.; GALEA, G.; REDDY, K. S. Scaling up interventions for chronic disease prevention: the evidence. **The Lancet**, v. 370, n. 9603, p. 1939-1946, dez. 2007.

GORTMAKER, S. L. *et al.* Social and economic consequences of overweight in adolescence and young adulthood. **New England Journal of Medicine**, v. 329, n. 14, p. 1008-1012, set. 1993.

GROSSMAN, M. On the concept of health capital and the demand for health. **Journal of Political Economy**, v. 80, n. 2, p. 223-255, mar.-abr. 1972.

HARRIS, A. Diabetes, cardiovascular disease and labour force participation in australia: an endogenous multivariate probit analysis of clinical prevalence data. **Economic Record**, v. 85, n. 271, p. 472-484, dez. 2009.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa de Orçamentos Familiares 2008-2009**: antropometria e estado nutricional de crianças, adolescentes e adultos no Brasil. Rio de Janeiro: IBGE, 2010.

JÄCKLE, R.; HIMMLER, O. Health and wages panel data estimates considering selection and endogeneity. **Journal of Human Resources**, v. 45, n. 2, p. 364-406, 2010.

JOHAR, M.; KATAYAMA, H. Quantile regression analysis of body mass and wages. **Health Economics**, v. 21, n. 5, p. 597-611, maio 2012.

KIDD, M. P.; SLOANE, P. J.; FERKO, I. Disability and the labour market: an analysis of British males. **Journal of Health Economics**, v. 19, n. 6, p. 961-981, nov. 2000.

LAROSE, S. L. *et al.* Does obesity influence labour market outcomes among working-age adults? Evidence from Canadian longitudinal data. **Economics & Human Biology**, v. 20, p. 26-41, mar. 2016.

MANSUR, A. de P. *et al.* Trends in death from circulatory diseases in Brazil between 1979 and 1996. **Arquivos Brasileiros de Cardiologia**, v. 76, n. 6, p. 497-510, 2001.

MANSUR, A. de P.; FAVARATO, D. Mortality due to cardiovascular diseases in Brazil and in the metropolitan region of São Paulo: a 2011 update. **Arquivos Brasileiros de Cardiologia**, v. 99, v. 2, p. 755-761, 2012.

MANSUR, A. de P.; FAVARATO, D. Trends in mortality rate from cardiovascular disease in Brazil, 1980-2012. **Arquivos Brasileiros de Cardiologia**, v. 107, n. 1, p. 755-761, 2016.

MITCHELL, J. M.; BURKHAUSER, R. V. Disentangling the effect of arthritis on earnings: a simultaneous estimate of wage rates and hours worked. **Applied Economics**, v. 22, n. 10, p. 1291-1309, 2006.

MITCHELL, J. M.; BUTLER, J. S. Arthritis and the earnings of men: an analysis incorporating selection bias. **Journal of Health Economics**, v. 5, n. 1, p. 81-98, mar. 1986.

MORRIS, S. Body mass index and occupational attainment. **Journal of Health Economics**, v. 25, n. 2, p. 347-364, mar. 2006.

PELKOWSKI, J. M.; BERGER, M. C. The impact of health on employment, wages, and hours worked over the life cycle. **The Quarterly Review of Economics and Finance**, v. 44, n. 1, p. 102-121, fev. 2004.

RTVELADZE, K. *et al.* Health and economic burden of obesity in Brazil. **PLoS ONE**, v. 8, n. 7, e68785, jul. 2013.

RIBEIRO, A. L. P. *et al.* Cardiovascular health in Brazil: trends and perspectives. **Circulation**, v. 133, n. 4, p. 422-433, jan. 2016.

SCHMIDT, M. I. *et al.* Chronic non-communicable diseases in Brazil: burden and current challenges. **The Lancet**, v. 377, n. 9781, p. 1949-1961, jun. 2011.

SHIMOKAWA, S. The labour market impact of body weight in China: a semi-parametric analysis. **Applied Economics**, v. 40, n. 8, p. 949-968, 2008.

SIQUEIRA, A. D. S. E.; SIQUEIRA-FILHO, A. G. D.; LAND, M. G. P. Analysis of the economic impact of cardiovascular diseases in the last five years in Brazil. **Arquivos Brasileiros de Cardiologia**, v. 109, n. 1, p. 39-46, 2017.

SONG, X. *et al.* Productivity loss and indirect costs associated with cardiovascular events and related clinical procedures. **BMC Health Services Research**, v. 15, n. 1, p. 245-259, jun. 2015.

SOUZA-JÚNIOR, P. R. B. de. *et al.* Desenho da Amostra da Pesquisa Nacional de Saúde 2013. **Epidemiologia e Serviços de Saúde**, Brasília, v. 24, n. 2, p. 207-216, abr.-jun. 2015.

STEVENS, B. *et al.* Os custos das doenças cardíacas no Brasil. **Arquivos Brasileiros de Cardiologia**, v. 111, n. 1, p. 29-36, 2018.

SUHRCKE, M.; URBAN, D. Are cardiovascular diseases bad for economic growth? **Health Economics**, v. 19, n. 12, p. 1478-1496, dez. 2010.

UNITED NATIONS. **Political declaration of the high-level meeting of the General Assembly on the prevention and control of non-communicable diseases**. Washington: UN, 2012.

UNITED NATIONS. **World population prospects: the 2015 revision**. Nova York: UN, 2016.

VEN, W. P. M. M. van de; ELLIS, R. P. Risk adjustment in competitive health plan markets. *In*: CULYER, A. J.; NEWHOUSE, J. P. (Ed.). **Handbook of health economics**. Amsterdam: Elsevier, 2000. v. 1, p. 755-845.

WHO – WORLD HEALTH ORGANIZATION. **Preventing chronic diseases: a vital investment**. Geneva: WHO, 2005.

WHO – WORLD HEALTH ORGANIZATION. **Global status report on noncommunicable diseases 2010**. Geneva: WHO, 2011.

WHO – WORLD HEALTH ORGANIZATION. **Global status report on noncommunicable diseases**. Geneva: WHO, 2014.

WHO – WORLD HEALTH ORGANIZATION. **Global health estimates 2015**: deaths by cause, age, sex, by country and by region, 2000-2015. Geneva: WHO, 2016a.

WHO – WORLD HEALTH ORGANIZATION. **World health statistics 2016**: monitoring health for the SDGs sustainable development goals. Geneva: WHO, 2016b.

WILSON, S. E. Work and the accommodation of chronic illness: a re-examination of the health-labour supply relationship. **Applied Economics**, v. 33, n. 9, p. 1139-1156, 2001.

ZWEIFEL, P.; BREYER, F.; KIFMANN, M. (Ed.). **Health economics**. Amsterdam: Springer Science & Business Media, 2009.

BIBLIOGRAFIA COMPLEMENTAR

AMERICAN HEART ASSOCIATION. **Cardiovascular disease**: a costly burden for America – projections through 2035. Dallas: American Heart Association, 2017.

BOWRY, A. D. K. *et al.* The burden of cardiovascular disease in low-and middle-income countries: epidemiology and management. **Canadian Journal of Cardiology**, v. 31, n. 9, p. 1151-1159, set. 2015.

FIRPO, S.; FORTIN, N. M.; LEMIEUX, T. Supplement to “unconditional quantile regressions”. **Econometrica**, v. 77, n. 3, p. 953-973, maio 2009.

GALLO, P.; GENÉ-BADIA, J. Cuts drive health system reforms in Spain. **Health Policy**, v. 113, n. 1-2, p. 1-7, nov. 2013.

GAULD, R. New Zealand’s post-2008 health system reforms: toward re-centralization of organizational arrangements. **Health Policy**, v. 106, n. 2, p. 110-113, jul. 2012.

HADDAD, S.; NOUGTARA, A.; FOURNIER, P. Learning from health system reforms: lessons from Burkina Faso. **Tropical Medicine & International Health**, v. 11, n. 12, p. 1889-1897, dez. 2006.

KRUK, M. E.; FREEDMAN, L. P. Assessing health system performance in developing countries: a review of the literature. **Health Policy**, v. 85, n. 3, p. 263-276, mar. 2008.

NIGENDA, G. *et al.* Evaluating the implementation of Mexico’s health reform: the case of Seguro Popular. **Health Systems & Reform**, v. 1, n. 3, p. 217-228, 2015.

PAIM, J. *et al.* The Brazilian health system: history, advances, and challenges. **The Lancet**, v. 377, n. 9779, p. 1778-1797, maio 2011.

TURK-ADAWI, K.; SARRAFZADEGAN, N.; GRACE, S. L. Global availability of cardiac rehabilitation. **Nature Reviews Cardiology**, v. 11, n. 10, p. 586-596, out. 2014.

VUORENKOSKY, L.; MLADOVSKY, P.; MOSSIALOS, E. Finland: health system review. **Health Systems in Transition**, v. 10, n. 4, p. 1-168, 2008.

WAGSTAFF, A. *et al.* China's health system and its reform: a review of recent studies. **Health Economics**, v. 18, n. S2, p. S7-S23, jul. 2009.

APÊNDICE A

TABELA A.1
Resultados do *propensity score matching: logit (odds ratio)*

	Homens	Mulheres
Altura	0,379*** (0,0000242)	0,448*** (0,0000289)
Altura ao quadrado	-0,00109*** (0,00001)	-0,00148*** (0,00001)
Peso corporal	-0,0184*** (0,00000346)	0,0194*** (0,00000383)
Peso corporal ao quadrado	0,0000606*** (0,00000001)	-0,0000961*** (0,00000001)
Cintura: risco médio	0,131*** (0,0000182)	-0,0286*** (0,000021)
Cintura: risco alto	-0,311*** (0,0000231)	0,171*** (0,0000207)
Cintura: risco altíssimo	0,392*** (0,0000191)	-0,0424*** (0,00002)
Pressão: pré-hipertensão	-0,343*** (0,0000143)	0,0222*** (0,0000151)
Pressão: hipertensão estágio 1	-0,205*** (0,0000228)	0,188*** (0,0000266)
Pressão: hipertensão estágio 2	0,134*** (0,000036)	0,397*** (0,0000501)
Pressão: hipertensão estágio 3	-0,410*** (0,0000607)	0,764*** (0,0000508)
IMC: peso ideal	0,561*** (0,0000623)	-0,376*** (0,000043)
IMC: sobrepeso	0,665*** (0,0000683)	-0,733*** (0,0000516)
IMC: obesidade 1	0,909*** (0,0000758)	-0,850*** (0,0000622)
IMC: obesidade 2	1,631*** (0,0000853)	-0,787*** (0,0000737)
IMC: obesidade 3	1,317*** (0,000106)	-1,291*** (0,0000936)

(Continua)

(Continuação)

	Homens	Mulheres
Índice de hábitos de vida	0,290*** (0,0000404)	0,376*** (0,0000421)
Plano de saúde	-0,281*** (0,0000139)	0,0321*** (0,0000134)
Idade	0,0347*** (0,00000457)	0,0484*** (0,0000046)
Idade ao quadrado	0,000350*** (0,00000001)	-0,0000574*** (0,00000001)
Fundamental incompleto	0,240*** (0,0000171)	-0,232*** (0,0000181)
Fundamental completo	0,299*** (0,0000212)	-0,0326*** (0,0000207)
Médio incompleto	-0,269*** (0,0000352)	0,166*** (0,000025)
Médio completo	0,120*** (0,0000194)	-0,413*** (0,0000191)
Superior incompleto	-0,287*** (0,000041)	-0,825*** (0,0000411)
Superior completo	0,552*** (0,0000224)	-0,611*** (0,0000246)
Branco	0,572*** (0,0000842)	0,256*** (0,0000603)
Preto	0,569*** (0,0000842)	0,244*** (0,0000602)
Chefe de família	-0,176*** (0,0000122)	-0,0739*** (0,0000127)
Casado	0,0311*** (0,0000125)	0,226*** (0,0000129)
Filhos	0,0551*** (0,00000661)	-0,139*** (0,00000763)
Tamanho amostral	13.885	10.879

Elaboração dos autores.

Obs.: Significância: *** 1%; ** 5%; * 10%.

Originais submetidos em: jun. 2018.

Última versão recebida em: ago. 2021.

Aprovada em: ago. 2021.

DESIGUALDADE DE OPORTUNIDADES EDUCACIONAIS: INFLUÊNCIA DAS CIRCUNSTÂNCIAS DA INFÂNCIA NO DESEMPENHO ESCOLAR NO BRASIL¹

Sabrina Martins de Araújo²

Cláudia Malbouisson³

Wallace Patrick Santos de Farias Souza⁴

Este artigo analisa como as circunstâncias da infância explicam o resultado educacional para alunos do 5º ano do ensino fundamental no Brasil. Utilizaram-se dados do Sistema de Avaliação da Educação Básica (Saeb) 2017 e, como estratégia empírica, o *generalized additive model* (GAM) de Hastie e Tibshirani (1986; 1987; 1990) e o índice de desigualdade de oportunidades (IOP) de Ferreira e Gignoux (2014). Verificou-se que, quanto melhor o conjunto de oportunidades, maior o desempenho escolar. A decomposição do IOP revela a importância de novas variáveis de circunstâncias relativas aos hábitos culturais, à relação entre pais e filhos, e à escola, para a desigualdade de oportunidade educacional.

Palavras-chave: ensino fundamental; circunstâncias da infância; desigualdade de oportunidades.

INEQUALITY OF EDUCATIONAL OPPORTUNITIES: INFLUENCE OF CHILDREN'S CIRCUMSTANCES ON SCHOOL PERFORMANCE IN BRAZIL

This article analyzes how childhood circumstances explain the educational outcome in Brazil, for students in the 5th year of elementary school. System of Assessment for Basic Education (Saeb) 2017 data were used and, as empirical strategy, the generalized additive model (GAM) of Hastie and Tibshirani (1986; 1987; 1990) and the inequality of opportunities index (IOP) of Ferreira and Gignoux (2014). It was found that the better the set of opportunities, the greater the school performance. The decomposition of the IOP reveals the importance of new variables of circumstances related to cultural habits, the relationship between parents and children, and the school, for the inequality of educational opportunity.

Keywords: elementary school; childhood circumstances; inequality of opportunities.

JEL: I20; I21; I24.

1 INTRODUÇÃO

A desigualdade de renda é uma questão antiga sobre a qual a literatura econômica tem se debruçado no sentido de compreender suas causas, consequências, definições e formas de mensuração, a fim de subsidiar a formulação e avaliação de

1. DOI: <http://dx.doi.org/10.38116/ppe52n3art6>

2. Professora substituta do Departamento de Administração e Economia da Universidade Estadual da Paraíba (Daec/UEPB). *E-mail:* sabrinaeconomia@gmail.com.

3. Professora do Programa de Pós-Graduação em Economia da Universidade Federal da Bahia (PPGE/UFBA). *E-mail:* claudiamalbo@gmail.com.

4. Professor do Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba (PPGE/UEPB). *E-mail:* wpsfarias@gmail.com.

políticas públicas voltadas ao seu combate. Trata-se de um problema que assola de forma mais intensa as economias subdesenvolvidas, com destaque para os países da América Latina em geral, mostrando-se como um desafio para estas (Barros e Mendonça, 1995; Ferreira, 2000; Hoffmann, 2001; Mendéz e Waltenberg, 2018).

O Brasil é um país historicamente desigual e ainda ocupa a sétima posição no *ranking* da desigualdade de renda medida pelo índice de Gini, segundo o *Relatório de Desenvolvimento Humano 2019* do Programa das Nações Unidas para o Desenvolvimento – PNUD (UNDP, 2019). A partir de Langoni (1973), estudos sobre os determinantes da desigualdade de renda no país apontam a desigualdade educacional e os altos retornos para níveis elevados de escolaridade como alguns dos seus principais fatores explicativos (Barros e Mendonça, 1995; Ferreira, 2000; Procópio, Freguglia e Chein, 2015).

Com relação a esse último aspecto, dados do Instituto Brasileiro de Geografia e Estatística (IBGE) indicam que a desigualdade educacional vem apresentando redução ao longo dos últimos anos. Entre 2007 e 2015, observa-se aumento do percentual de indivíduos com 25 anos ou mais de idade nas faixas de escolaridade mais elevada; elevação do percentual de pessoas com onze a catorze anos de estudo, de 24,7% para 30,5%; e aumento da população com quinze anos ou mais de estudo, de 8,9% para 13,0%.⁵ Outro dado importante refere-se à taxa de escolarização das crianças de 6 a 14 anos, que era 98,6% em 2015, o que representa a quase universalização do ensino fundamental.⁶

Contudo, considerando-se a importância da qualidade da educação para a redução da desigualdade educacional, a melhoria vem sendo observada apenas para os anos iniciais do ensino fundamental. O Índice de Desenvolvimento da Educação Básica (Ideb) revela que o país vem superando, para esta etapa, as metas estabelecidas – 5,2 em 2013 (meta 4,9); 5,5 em 2015 (meta 5,2); 5,8 em 2017 (meta 5,5); e 5,9 em 2019 (meta 5,7) –, mas os resultados ainda são modestos. Nos anos finais do ensino fundamental e no ensino médio, as metas estabelecidas não foram atingidas (Inep, 2020).

No comparativo internacional, a situação é ainda mais crítica. Conforme o *Relatório Brasil no Pisa*⁷ 2018, das 79 economias participantes da avaliação (Brasil, 2019), o Brasil ocupa o intervalo 55-59 no *ranking* de leitura, 69-72 no de matemática, e 64-67 no de ciências. Isso significa que as melhorias observadas do ponto de vista quantitativo, por exemplo, em termos de anos de estudo, não necessariamente garantem oportunidades iguais, já que não refletem qualidade

5. Disponível em: <https://brasilemsintese.ibge.gov.br/educacao/anos-de-estudo.html>. Acesso em: 5 set. 2018.

6. Disponível em: <https://brasilemsintese.ibge.gov.br/educacao/taxa-de-escolarizacao-das-pessoas-de-6-a-14-anos.html>. Acesso em: 5 set. 2018.

7. Programme for International Student Assessment.

educacional e, portanto, igual obtenção de conhecimento e habilidade individual (Nogueira, 2015).

No que tange à discussão teórica sobre desigualdade de resultados, de modo geral, se considera a perfeita igualdade como parâmetro de avaliação para definir se uma sociedade é justa. Isto pode ser observado por meio da utilização de medidas tradicionais de desigualdade de renda como os índices de Gini e Theil. Contudo, essa visão tem sido questionada. Entre os anos 1970 e 1980, um debate no campo da filosofia política contribuiu para mover a demanda por igualdade do espaço dos resultados individuais para o espaço das oportunidades. Sob esse ponto de vista, a desigualdade deveria ser decomposta em termos justos e injustos, indicando que a justiça social não exige que os resultados individuais sejam todos iguais, mas que as oportunidades das pessoas sejam iguais (Aguirreche, 2012; Ferreira e Peragine, 2015; Souza, Oliveira e Annegues, 2017).

Em economia, a principal contribuição foi dada por Roemer (1993; 1998), para o qual os resultados dos indivíduos seriam decompostos em circunstâncias e esforços. Deste modo, tem-se uma teoria da (des)igualdade de oportunidades baseada em três tipos de informação: resultados, circunstâncias e esforços. Resultados são definidos como realizações educacionais, renda, acesso a serviços básicos etc. Circunstâncias são variáveis que não estão sob o controle do indivíduo (encontram-se além da sua responsabilidade) e que afetam sua capacidade de atingir o resultado, tais como sexo, raça, *background* familiar. Já os esforços são variáveis sob controle do indivíduo (sob sua responsabilidade), a exemplo do nível educacional, horas trabalhadas, decisão de migrar, entre outros⁸ (Carvalho e Waltenberg, 2015).

Assim, define-se desigualdade de oportunidades educacionais como a desigualdade decorrente de fatores de circunstâncias dos estudantes, dado o esforço individual (Aguirreche, 2012). A título de ilustração, se dois estudantes X e Y exercem o mesmo esforço, mas possuem conjuntos de circunstâncias distintos que os impedem de obter os mesmos resultados, tem-se uma situação de desigualdade de oportunidades, que deve ser compensada por políticas públicas (Nogueira, 2015).

Um aspecto fundamental na análise de desigualdade de oportunidades na educação básica refere-se à atribuição de responsabilidade a indivíduos em idade escolar por suas escolhas/esforços, já que se trata de crianças e adolescentes pelos quais os pais e a escola podem ser os responsáveis. Alguns estudos discutem esse assunto sem se aprofundarem ou chegarem a um consenso, apontando ao menos uma responsabilidade parcial a alunos com idade em torno de 15 anos (Waltenberg e Vandenberghe, 2007; Gamboa e Waltenberg, 2012). Procópio, Freguglia

8. Quanto ao esforço, a literatura empírica difere no tratamento dado a ele, com alguns trabalhos considerando-o observável, a exemplo de Bourguignon, Ferreira e Menéndez (2007), e outros, não observável, tais como Fleurbaey (2008) e Pistolesi (2009).

e Chein (2015) discutem o tema no âmbito do ensino fundamental e optam por tratar a parcela da desigualdade não explicada pelas variáveis de circunstâncias como desigualdade residual, e não como desigualdade justa ou decorrente do esforço.

Para Hufe *et al.* (2015), todos os comportamentos e realizações das crianças devem ser compreendidos como consequências das circunstâncias. Conforme estes autores, a omissão de variáveis importantes de circunstâncias da infância levou a medidas de desigualdade de oportunidades inadequadas ou tendenciosas para baixo, pois um indivíduo não deve ser considerado responsável por suas escolhas antes do alcance de uma idade de consentimento (para uma escolha responsável). Deste modo, expandem o conjunto de circunstâncias analisado, incluindo os atributos do indivíduo e seu ambiente quando criança na análise da desigualdade de renda.

Contudo, os trabalhos empíricos sobre desigualdade de oportunidades educacionais se concentram em um conjunto limitado de circunstâncias, muito provavelmente por considerarem parcela significativa dos determinantes do desempenho educacional como variáveis sob responsabilidade do estudante. Ao deixarem de incluir variáveis importantes que não estão sob o controle do aluno, a parcela injusta da desigualdade é subestimada. Este trabalho contribui à literatura no sentido de incorporar um maior número de variáveis circunstanciais – inclusive novas características relativas ao *background* familiar, além de características da escola – por entender que as crianças não realizam de forma independente suas escolhas ou esforços. Logo, assim como Hufe *et al.* (2015), assume-se que as crianças não devem ser responsabilizadas por suas realizações.

Diante disto, o objetivo deste artigo é analisar como as circunstâncias da infância explicam o resultado educacional no Brasil, para alunos do 5º ano do ensino fundamental. De outro modo, busca-se estudar a desigualdade de oportunidade educacional considerando todos os determinantes dos resultados escolares como variáveis de circunstâncias, ao invés de considerar uma parcela como de escolha do próprio estudante, já que normalmente se trata de alunos na faixa etária de 10 a 11 anos. Adicionalmente, faz-se a inclusão de novas variáveis de circunstâncias da infância na análise da desigualdade de oportunidades, como variáveis relativas à relação entre pais e filhos, aos hábitos culturais e à escola.

Para atender ao objetivo delineado, foram utilizados microdados do Sistema de Avaliação da Educação Básica (Saeb), do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), para o ano 2017. São empregadas duas metodologias. A primeira é o *generalized additive model* (GAM), de Hastie e Tibshirani (1986; 1987; 1990), método não paramétrico utilizado para analisar como as circunstâncias da infância determinam o resultado educacional. Esse método também foi aplicado por Hunger *et al.* (2011) para avaliar o impacto da multimorbidade na qualidade de vida dos idosos na Alemanha, por Colón-González *et al.* (2013) e

Prata, Rodrigues e Bermejo (2020) para estudar, respectivamente, o efeito de fatores climáticos na incidência de dengue no México e covid-19 no Brasil, e por Souza, Oliveira e Annegues (2018) para investigar a relação entre características familiares e desempenho escolar.

Já a segunda metodologia busca mensurar a desigualdade de oportunidade educacional, bem como a contribuição de cada conjunto de circunstâncias para ela, através do modelo proposto por Ferreira e Gignoux (2014). Os resultados apontam que quanto melhor o conjunto de oportunidades, maior o desempenho educacional do estudante, sendo o contrário observado para o grupo menos favorecido. A decomposição dos índices de desigualdade de oportunidades (IOPs) ratifica o conjunto base – que contém as variáveis de circunstâncias mais utilizadas na literatura, como sexo, escolaridade dos pais, índice socioeconômico, tipo de escola – como o mais importante, seguido pelos grupos de hábitos culturais e de variáveis relacionais para o desempenho em português, e pelo de hábitos culturais e características da escola para matemática. De modo geral, acrescentar novos conjuntos de circunstâncias aumenta em 15% e 13%, respectivamente, o IOP para português e matemática.

Este artigo está organizado em cinco seções, incluindo esta introdução. A segunda seção apresenta uma breve revisão da literatura empírica. A terceira trata de aspectos da metodologia utilizada para o desenvolvimento da pesquisa como dados, variáveis, estatísticas descritivas e estratégia empírica. Na quarta seção, tem-se os resultados do estudo. A quinta apresenta as considerações finais.

2 REVISÃO DA LITERATURA

A teoria da igualdade de oportunidades foi aplicada a vários trabalhos voltados para a análise da desigualdade nos resultados educacionais. Esses trabalhos utilizam metodologias diversas e analisam a desigualdade em diferentes países e níveis de ensino.

Foguel e Veloso (2012) aferem o grau de desigualdade de oportunidades nos serviços de creche e pré-escola no Brasil. Para isto, constroem um índice de oportunidades a partir do índice de oportunidades humanas usado em Barros *et al.* (2009), com o acesso à educação incluindo, além da frequência real, a escolha dos pais de matricular as crianças. Aplicando dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) 2006, os resultados indicam grandes diferenças entre os índices para as crianças de 0 a 3 anos de idade, e diferenças menores para crianças de 4 a 6 anos. Como explicação tem-se que, à medida que as crianças crescem, o número daquelas que não frequentam a escola devido à escolha dos pais diminui, o que indica que a pré-escola pode estar mais próxima de uma oportunidade básica do que a creche. Constatou-se ainda que uma criança típica do grupo mais desfavorecido é não branca e vive em famílias com renda *per capita* abaixo da média,

cujo chefe não completou o grau primário. O contrário é observado para o grupo menos desfavorecido.

Procópio, Freguglia e Chein (2015), por sua vez, avaliam o comportamento da desigualdade educacional e de oportunidades ao longo dos quatro primeiros anos do ensino fundamental. Os autores utilizam as proficiências em português e matemática de crianças participantes do Projeto Geração Escolar (Geres), no período entre 2005 e 2008. O método aplicado para a mensuração da desigualdade de oportunidades seguiu Ferreira e Gignoux (2011), que indicam uma abordagem paramétrica a partir de uma estimação por mínimos quadrados ordinários (MQO). Os principais resultados revelaram: um aumento na desigualdade educacional em matemática e um comportamento estável em português; uma desigualdade de oportunidades no final do período de 8% para matemática e 11% para português (com esses valores aumentando ao longo do tempo em 80% e 109%, respectivamente); e o nível socioeconômico tendo se apresentado como o fator de maior influência na desigualdade de oportunidade educacional.

Souza, Oliveira e Annegues (2017), também em pesquisa para a educação fundamental, mensuram a desigualdade de oportunidade nos estados do Brasil a partir de dados do Saeb 2005 e 2011, para alunos da 4ª série. Com base no índice proposto por Salehi-Isfahani, Hassine e Assaad (2014), os autores constroem IOPs por meio de regressões paramétricas e não paramétricas. Conforme o ano e o estado analisados, verificam que a desigualdade de oportunidade explica entre 8% e 41% da desigualdade total.

Por sua vez, Tavares, Camelo e Paciência (2018) estudam a desigualdade de oportunidades no ensino fundamental (5º e 9º ano), considerando fatores sobre os quais as redes de ensino e/ou escolas podem exercer influência (fatores de esforço) e aqueles que estão fora do seu controle (fatores de circunstâncias). Usando uma estratégia empírica semelhante à proposta por Bourguignon *et al.* (2007), os resultados apontam que o efeito-escola sobre a desigualdade de notas é de 40% a 45%, e o efeito-rede, de 50% a 60%. Isso mostra que grande parcela da desigualdade de notas está sob o controle de políticas educacionais estabelecidas pelos órgãos centrais e diretores escolares, estando o restante relacionado às circunstâncias da escola.⁹

Diaz (2012), em análise para o ensino médio, examina a evolução do nível de desigualdade de oportunidades utilizando dados do Saeb de 1995 a 2005, das avaliações de português e matemática. Seguindo Ferreira e Ginoux (2008), as medidas de desigualdade foram obtidas de forma não paramétrica a partir do índice Theil L e do índice que corresponde à metade do quadrado do coeficiente

9. Essas circunstâncias são: leis e normas, perfil dos gestores, dos docentes, tempo de dedicação do professor, dotação de recursos financeiros e físicos da escola, tamanho da escola, número de etapas de ensino, localização, tamanho das turmas, além do nível socioeconômico dos alunos e da composição do corpo discente.

de variação.¹⁰ Observou-se que as desigualdades de oportunidades foram menores em português (média na faixa de 16% da desigualdade total) do que em matemática (média em torno de 24%), com esse padrão se repetindo nas escolas pública e privada. Verificou-se também que a desigualdade de oportunidades nas escolas privadas caiu em ambas as áreas de conhecimento (atingindo, em 2005, 8% e 10% em português e matemática, respectivamente), enquanto permaneceu relativamente estável nas escolas públicas.

Nogueira (2015), assim como Diaz (2012), pesquisa a desigualdade de oportunidades no ensino médio no Brasil. O trabalho analisa o impacto da origem individual no desempenho dos alunos no Exame Nacional do Ensino Médio (Enem). Os resultados mostram forte influência de fatores sociais e econômicos como renda familiar, escolaridade dos pais e tipo de escola na determinação da desigualdade de oportunidade educacional.

Na literatura internacional, o estudo de Ferreira e Gignoux (2011) avalia a desigualdade educacional em 57 países, a partir de dados do Pisa 2006. Os autores propõem duas medidas de desigualdade, uma para o desempenho educacional e outra para oportunidades educacionais. A primeira medida é a variância (ou desvio-padrão) dos resultados dos testes, e a segunda, que é a medida de desigualdade de oportunidades educacionais, é dada pela parcela da variância nos resultados explicada pelas circunstâncias. Verificou-se que a desigualdade de oportunidades representa até 35% de todas as disparidades no desempenho educacional, sendo maior na Europa continental e América Latina do que na Ásia, Escandinávia e América do Norte.

Gamboa e Waltenberg (2011) também empregam dados do Pisa (2006 e 2009) e analisam a desigualdade de oportunidades educacionais em seis países da América Latina. Utilizando uma abordagem não paramétrica e variáveis de circunstâncias como tipo de escola (pública ou privada), gênero e educação parental, além de combinações dessas variáveis, constatam que a desigualdade de oportunidades varia de menos de 1% a 25%, com heterogeneidade considerável entre países, indivíduos e anos. Os autores chamam a atenção para as variáveis *educação dos pais e tipo de escola* como importantes fontes de desigualdade injusta, principalmente na Argentina e no Brasil.

Na mesma direção, Aguirreche (2012) mensura a desigualdade de oportunidades educacionais em 65 países com base em dados do Pisa 2009. Empregando um modelo estrutural baseado em Fleurbaey e Schokkaert (2009) e a variância como medida absoluta de desigualdade, verificou que os países com maior grau de desigualdade de oportunidade, superior a 30% da desigualdade total, eram

10. Índice de entropia generalizada com parâmetro de sensibilidade igual a dois.

principalmente países da América do Sul, Europa Oriental e Ásia. Já os países com menor desigualdade injusta eram da América do Norte, Europa Ocidental e Oceania. Observou ainda uma correlação negativa de -0,69 entre a desigualdade de oportunidades e as realizações dos alunos.

Por seu turno, Carvalho e Waltenberg (2015) estudam a desigualdade de oportunidades no acesso ao ensino superior com base em dados da PNAD 2003 e 2013, a partir do índice de oportunidade humana (IOH), desenvolvido por Barros *et al.* (2009). Ademais, utilizam o método de decomposição de Shapley para análise dos fatores mais associados à desigualdade estudada. Os resultados indicam que houve realocação de oportunidades dos grupos não vulneráveis para os vulneráveis e um leve aumento no acesso educacional. Porém, o IOH para o presente caso (IOH = 0,281) ainda se mostrou aquém do ideal (IOH = 1), quando o acesso é universal. Por último, verificou-se que as circunstâncias de maior relevância foram a instrução do chefe do domicílio e a renda domiciliar *per capita*, com uma contribuição moderada do fator cor/raça.

De modo geral, a literatura empírica mostra a importância da desigualdade de oportunidades para a desigualdade educacional. Contudo, como destacado na introdução deste artigo, as análises se concentram em um conjunto limitado de circunstâncias, o que justifica a contribuição desta pesquisa.

3 PROCEDIMENTOS METODOLÓGICOS

3.1 Dados, variáveis e estatísticas descritivas

Para estudar a desigualdade de oportunidades educacionais no Brasil foram utilizados microdados do Saeb, pesquisa realizada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). O ano analisado é 2017 e o público-alvo da pesquisa são os alunos do 5º ano do ensino fundamental, considerando escolas públicas e privadas conjuntamente. São utilizados dados das proficiências nas avaliações em língua portuguesa e matemática e dos questionários aplicados ao aluno, professor, diretor e escola.

O Saeb foi criado em 1990, como uma avaliação única, e desde então passou por algumas mudanças. Em 2005, passou a contar com duas avaliações, a Avaliação Nacional da Educação Básica (Aneb), que manteve as características que já vinham sendo adotadas até aquele momento, e a Avaliação Nacional do Rendimento Escolar (Anresc), também conhecida como Prova Brasil, desenvolvida com o objetivo de avaliar a qualidade do ensino nas escolas públicas. Em 2013, incorpora a Avaliação Nacional de Alfabetização (ANA). O Saeb é atualmente composto por essas três avaliações externas em larga escala (Brasil, 2018).

A Aneb é realizada bianualmente, de forma amostral e censitária, abrangendo escolas e alunos das redes públicas e privadas do país, nas áreas rurais e urbanas, matriculados no 5º e 9º anos do ensino fundamental e 3º ano do ensino médio regulares. A Prova Brasil é bianual e censitária, englobando alunos do 5º e 9º anos do ensino fundamental das escolas públicas que possuem no mínimo vinte alunos matriculados na série avaliada. Por sua vez, a ANA afere os níveis de alfabetização e letramento em língua portuguesa (leitura e escrita) e matemática no 3º ano do ensino fundamental das escolas públicas, sendo realizada de forma censitária (Brasil, 2018).

Com relação às variáveis utilizadas na estratégia empírica, deve-se destacar que estas compõem sete cenários, os quais compreendem diferentes conjuntos de circunstâncias que subsidiam a avaliação da desigualdade de oportunidades no 5º ano do ensino fundamental. Cada cenário incorpora novas variáveis de não responsabilidade que fazem parte de um conjunto específico, o que possibilita a observação de como as estimativas variam conforme cada conjunto de circunstância é sequencialmente introduzido. No quadro 1, tem-se uma visão geral dos cenários e conjuntos de circunstâncias estudados.¹¹

QUADRO 1
Cenários e conjuntos de circunstâncias da infância

Variáveis de não responsabilidade	Conjuntos de circunstâncias	Cenários						
Sexo	Base	1º						
Raça								
Escolaridade da mãe								
Escolaridade do pai								
Índice socioeconômico								
Frequentou creche e/ou pré-escola								
Tamanho da família								
Já foi reprovado								
Já abandonou a escola								
Escola pública ou privada								
Escola urbana ou rural								
Trabalho fora de casa								
Região								

(Continua)

11. O quadro A.1 do apêndice A apresenta a descrição detalhada da natureza das variáveis.

(Continuação)

Variáveis de não responsabilidade	Conjuntos de circunstâncias	Cenários														
Mora com mãe e pai	Relação entre pais e filhos	2ª														
Vê mãe lendo																
Vê pai lendo																
Frequência às reuniões de pais																
Pais incentivam a estudar																
Pais incentivam a fazer dever de casa																
Pais incentivam a ler																
Pais incentivam a ir à escola																
Pais conversam sobre o que acontece na escola																
Lê jornais	Hábitos culturais e uso do tempo			3ª												
Lê livros																
Lê revistas em geral																
Lê revistas em quadrinhos																
Lê sites na internet																
Costuma ir à biblioteca																
Costuma ir ao cinema																
Costuma ir a espetáculos																
Tempo gasto em dias de aula assistindo TV, navegando na internet ou jogando jogos eletrônicos																
Tempo gasto em dias de aula fazendo trabalhos domésticos	Características do professor															
Sexo																
Raça																
Se tem curso superior																
Experiência																
Formação continuada																
Pós-graduação																
Sexo									Características do diretor							
Raça																
Se tem curso superior																
Experiência																
Formação continuada																
Pós-graduação	Características da escola															
Qualidade da escola																
Violência																
Homogênea quanto à idade	Características da turma															
Homogênea quanto ao rendimento escolar																
Heterogênea quanto à idade																
Heterogênea quanto ao rendimento escolar																

Fonte: Hufe *et al.* (2015).
Elaboração dos autores.

O primeiro cenário, chamado cenário base, contém as variáveis de circunstâncias mais utilizadas na literatura sobre desigualdade de oportunidade educacional e sobre os determinantes do desempenho escolar. O segundo incorpora variáveis características do relacionamento entre pais e filhos. Já o terceiro cenário compreende variáveis ligadas aos hábitos culturais e de uso do tempo dos estudantes, enquanto os cenários quatro, cinco, seis e sete adicionam variáveis relacionadas ao professor, ao diretor, à escola e à turma, respectivamente.

As estatísticas descritivas da amostra utilizada são apresentadas na tabela 1. Verifica-se que as médias dos alunos em português e matemática foram 230,9 e 240,4, respectivamente. As estatísticas do conjunto base mostram que 49% dos indivíduos são meninos, 37% brancos e cerca de 25% possuem pai ou mãe com curso superior. Destaca-se ainda que 78% dos estudantes fizeram creche ou pré-escola, suas famílias têm em média quatro pessoas, 15% já foram reprovados e 5% já abandonaram a escola.

No que tange ao tipo de escola e à localização, verifica-se que 98% dos alunos estudam em escolas públicas e 90% em escolas urbanas, estando a maioria concentrados nas regiões Sudeste (43%) e Nordeste (21%) do país. Ademais, 9% dos alunos afirmaram trabalhar fora de casa, recebendo ou não um salário.

O conjunto de variáveis que representam a relação entre pais e filhos é fundamental para esse trabalho pois, segundo Coleman (1988), compõe o capital social da família que é importante para o desempenho escolar das crianças. Observa-se que 68% dos alunos moram com mãe e pai. Esta variável representa a estrutura familiar e é bastante relevante na literatura sobre resultados infantis. Em média, 89% dos discentes veem mãe ou pai lendo, 87% afirmam que os pais frequentam as reuniões de pais, 98%, que os pais os incentivam a realizarem atividades ligadas à educação, evidenciando a importância atribuída por eles ao aprendizado dos filhos, e 86%, que os pais conversam sobre o que acontece na escola.

Quanto aos aspectos culturais, tem-se que a maior parte dos alunos lê livros (95%) e revistas em quadrinhos (89%), 66% dos estudantes costumam frequentar bibliotecas, 69%, cinema, e 57%, espetáculos. As variáveis de uso do tempo mostram que o tempo gasto em dias de aula assistindo TV, navegando na internet ou jogando jogos eletrônicos é de aproximadamente 2,4 horas, e fazendo trabalhos domésticos, algo próximo a 1,6 horas. Essas variáveis formam o capital cultural das famílias e crianças, o que, para Bourdieu (2007), determina o resultado escolar e a desigualdade educacional, tema mais bem abordado nos resultados desse estudo.

Com relação aos grupos que representam a escola, a grande maioria dos professores (87%) e diretores (84%) são mulheres e entre 50% e 54% são de cor branca. Cerca de 95% dos professores têm formação superior, 70% têm dez anos ou mais de experiência, 88% participaram de cursos de formação continuada e 67%

possuem pós-graduação. Quanto aos diretores, 98% possuem curso superior, 39% têm mais de cinco anos de experiência, 85% participaram de cursos de formação continuada e 83% têm pós-graduação.

Já quanto às características das turmas, parte considerável é homogênea quanto à idade (31%) e heterogênea quanto ao rendimento escolar (21%). Deve-se destacar que as variáveis de escola vêm ganhando cada vez mais importância em economia da educação com a disponibilidade de dados de melhor qualidade e o aperfeiçoamento dos métodos de análise.

TABELA 1
Estatísticas descritivas: 5º do ensino fundamental (2017)

Variáveis	Média	Desvio-padrão	Mínimo	Máximo
Variáveis dependentes				
Proficiência em português	230,9	46,81	125,0	334,2
Proficiência em matemática	240,4	45,31	125,0	357,0
Conjuntos de circunstâncias				
Base				
Aluno homem	0,49	0,50	0,00	1,00
Aluno branco	0,37	0,48	0,00	1,00
Mãe com curso superior	0,26	0,44	0,00	1,00
Pai com curso superior	0,24	0,43	0,00	1,00
Índice socioeconômico	-0,27	0,92	-4,18	1,19
Frequentou creche/pré-escola	0,78	0,42	0,00	1,00
Tamanho da família	4,24	1,15	1,00	6,00
Já foi reprovado	0,15	0,36	0,00	1,00
Já abandonou a escola	0,05	0,21	0,00	1,00
Escola pública	0,98	0,13	0,00	1,00
Escola urbana	0,90	0,30	0,00	1,00
Trabalho	0,09	0,29	0,00	1,00
Norte	0,11	0,31	0,00	1,00
Nordeste	0,21	0,41	0,00	1,00
Sudeste	0,43	0,49	0,00	1,00
Sul	0,17	0,37	0,00	1,00
Centro-Oeste	0,08	0,28	0,00	1,00
Relação entre pais e filhos				
Mora com mãe e pai	0,68	0,47	0,00	1,00
Vê mãe lendo	0,93	0,26	0,00	1,00
Vê pai lendo	0,85	0,36	0,00	1,00
Frequência às reuniões de pais	0,87	0,33	0,00	1,00
Incentivo a estudar	0,99	0,12	0,00	1,00

(Continua)

(Continuação)

Variáveis	Média	Desvio-padrão	Mínimo	Máximo
Incentivo a fazer dever de casa	0,98	0,16	0,00	1,00
Incentivo a ler	0,96	0,20	0,00	1,00
Incentivo a ir à escola	0,98	0,16	0,00	1,00
Conversa	0,86	0,35	0,00	1,00
Hábitos culturais e uso do tempo				
Lê jornais	0,39	0,49	0,00	1,00
Lê livros	0,95	0,21	0,00	1,00
Lê revistas em geral	0,61	0,49	0,00	1,00
Lê revistas em quadrinhos	0,89	0,32	0,00	1,00
Lê sites na internet	0,76	0,43	0,00	1,00
Biblioteca	0,66	0,47	0,00	1,00
Cinema	0,69	0,46	0,00	1,00
Espectáculos (museu, teatro, dança)	0,57	0,50	0,00	1,00
Tempo vendo TV	2,37	1,31	0,00	4,00
Tempo com trabalho doméstico	1,57	1,18	0,00	4,00
Características do professor				
Professor homem	0,13	0,23	0,00	1,00
Professor branco	0,50	0,39	0,00	1,00
Curso superior	0,95	0,16	0,00	1,00
Experiência	0,70	0,32	0,00	1,00
Formação continuada	0,88	0,23	0,00	1,00
Pós-graduação	0,67	0,34	0,00	1,00
Características do diretor				
Diretor homem	0,16	0,37	0,00	1,00
Diretor branco	0,54	0,50	0,00	1,00
Curso superior	0,98	0,14	0,00	1,00
Experiência	0,39	0,49	0,00	1,00
Formação continuada	0,85	0,36	0,00	1,00
Pós-graduação	0,83	0,38	0,00	1,00
Características da escola				
Qualidade da escola	0,02	1,16	-3,99	1,55
Violência	-0,21	0,89	-1,06	4,91
Características da turma				
Homogênea: idade	0,31	0,46	0,00	1,00
Homogênea: rendimento escolar	0,05	0,21	0,00	1,00
Heterogênea: idade	0,05	0,21	0,00	1,00
Heterogênea: rendimento escolar	0,21	0,41	0,00	1,00
Número de observações	301.873			

Fonte: Microdados do Saeb 2017. Disponível em: <https://www.gov.br/inep/pt-br/acao-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.

Elaboração dos autores.

No tocante ao índice socioeconômico, às variáveis de qualidade da escola e ao indicador de violência, pode-se dizer que quanto maiores, melhor a condição socioeconômica, melhor a estrutura da escola e maior a violência escolar.

3.2 Estratégia empírica

3.2.1 GAM

Para analisar como as circunstâncias da infância determinam o desempenho do aluno em português e matemática no 5º ano do ensino fundamental das escolas brasileiras, empregou-se uma abordagem não paramétrica. O modelo econométrico utilizado é o GAM ou modelo aditivo generalizado.¹²

Este método foi utilizado por Hunger *et al.* (2011) para avaliar o impacto da multimorbidade na qualidade de vida dos idosos na Alemanha. Os resultados apontaram efeitos negativos de condições crônicas de saúde, tanto isoladamente quanto em combinação com outras condições, destacando-se os últimos. Também por meio do emprego de um modelo aditivo generalizado, Colón-González *et al.* (2013) verificam que o clima influencia significativamente a incidência de dengue no México, enquanto Prata, Rodrigues e Bermejo (2020) utilizam o GAM para estimar a relação entre temperatura e infecção por covid-19 nas capitais brasileiras, com os resultados indicando uma relação negativa, porém sem evidências de redução para temperaturas acima de 25,8 °C. Já Souza, Oliveira e Annegues (2018) mostram, através da estimação do GAM, que variáveis de *background* familiar têm grande influência no desempenho escolar dos alunos em comparação com variáveis relativas à escola em geral.

O GAM foi introduzido por Hastie e Tibshirani (1986; 1987; 1990), revelando-se útil na descoberta de efeitos de covariáveis não lineares. Este modelo substitui a forma linear $\sum \beta_j X_j$ por um soma de funções suaves $\sum f_j(X_j)$, sendo considerado uma extensão dos modelos lineares generalizados, em que o preditor linear $\eta(X) = \alpha + \sum \beta_j X_j$ é substituído pelo preditor aditivo $\eta(X) = \alpha + \sum f_j(X_j)$, utilizando-se uma função de ligação para relacionar a variável dependente aos preditores aditivos, sendo $\eta(X) = g(\mu(X))$. Assim, um modelo aditivo generalizado tem a forma (Hastie e Tibshirani, 1987):

$$g(\mu(X)) = \alpha + \sum_{j=1}^p f_j(X_j), E[f_j(X_j)] = 0. \quad (1)$$

Existem basicamente três tipos de procedimentos para estimar os modelos GAM. O método de integração marginal (Linton e Nielsen, 1995), o método de

12. O modelo aditivo generalizado permite estimativas suaves para todas as covariáveis e ajustes lineares para algumas delas, no último caso podendo se configurar num modelo semiparamétrico (Hastie e Tibshirani, 1986).

dois estágios (Horowitz e Mammen, 2004) e o algoritmo *backfitting* (Buja, Hastie e Tibshirani, 1989; Hastie e Tibshirani, 1990; Souza, Oliveira e Annegues, 2018; Silveira, 2015). Neste trabalho foi utilizado o algoritmo *backfitting*.¹³

O algoritmo *backfitting* é um método iterativo que estima as funções suaves f_j no modelo, uma de cada vez, com base nos resíduos:

$$r_{ij} = y_i - \hat{\alpha} - \sum_{\substack{k=1 \\ k \neq j}}^p \hat{f}_k(X_{ik}). \quad (2)$$

Em cada estágio, o algoritmo suaviza os resíduos contra a próxima covariável. Esses resíduos são obtidos removendo-se as funções estimadas ou os efeitos de todas as outras variáveis. O processo se repete até convergir (Hastie e Tibshirani, 1987; Buja, Hastie e Tibshirani, 1989; Venables e Ripley, 2002). Qualquer método de suavização pode ser utilizado para estimar $f(\cdot)$. Neste trabalho foram utilizadas funções suaves *splines*. Segundo Hastie e Tibshirani (1987), quando funções suaves *splines* cúbicas são usadas no algoritmo *backfitting*, elas convergem para a solução do problema de mínimos quadrados penalizados:¹⁴

$$\min \sum_{j=1}^n (y_i - \sum_{j=1}^p f_j(X_{ji}))^2 + \sum_{j=1}^p \lambda_j \int [f_j''(X_j)]^2 dX_j. \quad (3)$$

3.2.2 IOP educacional

Para mensurar a desigualdade de oportunidade educacional e a contribuição de cada circunstância para ela, este estudo segue o método adotado por Ferreira e Gignoux (2014). Esses autores utilizam uma medida paramétrica para o IOP educacional, com base em uma regressão por MQO do desempenho do estudante (y) em função de um conjunto de circunstâncias (C):

$$IOP = \frac{I(C'_i \hat{\beta})}{I(y)}, \quad (4)$$

em que $\hat{\beta}$ é a estimativa MQO dos coeficientes de regressão em uma estimação simples de y em C :

$$y_i = C'_i \beta + \eta_i. \quad (5)$$

Em (4), $C'_i \hat{\beta}$ denota o vetor das pontuações previstas dos estudantes nos testes a partir da regressão (5). Sob a suposição de uma relação linear entre desempenho e circunstâncias, esse vetor representa uma distribuição suavizada, uma vez que

13. No caso gaussiano, o algoritmo *local scoring* é exatamente o algoritmo *backfitting* do modelo aditivo (Hastie e Tibshirani, 1987).

14. Para mais detalhes sobre o uso do estimador *spline* cúbico, ver Hastie e Tibshirani (1987; 1990) e Hardle e Linton (1994).

todos os indivíduos com circunstâncias idênticas recebem o mesmo valor previsto do desempenho.

Ferreira e Gignoux (2014) então elegem a variância como índice de desigualdade $I(\cdot)$. Dentre outras propriedades importantes, tal escolha ocorre porque essa medida é ordinalmente equivalente quando aplicada a distribuições pré e pós-padronizadas, já que a variância de uma distribuição pós-padronização é uma transformação monotônica da variância de uma distribuição pré-padronização (Procópio, Freguglia e Chein, 2015).

Essa característica é de grande importância para os estudos sobre desigualdade no contexto educacional. Pesquisas nessa área geralmente utilizam notas padronizadas como medidas de desempenho, calculadas por meio da teoria da resposta ao item. O processo de padronização utilizado requer que a média e a variância das notas sejam ajustadas a partir de uma escala arbitrária. Essa transformação impede a aplicação de medidas tradicionais de desigualdade para a desigualdade educacional, como os índices de Gini e Theil, pois o uso dessas medidas sobre notas padronizadas pode não refletir ordinalmente a desigualdade original do fenômeno (Ferreira e Gignoux, 2014; Procópio, Freguglia e Chein, 2015; Tavares, Camelo e Paciência, 2018).

Diante disso, a escolha da variância como índice de desigualdade produz a proposta de medida de desigualdade de oportunidade educacional, como um caso especial de (4):

$$IOP = \frac{Var(C_i' \hat{\beta})}{Var(y_i)}. \quad (6)$$

Este índice é simplesmente o R quadrado de uma regressão MQO do resultado educacional da criança em um vetor C de circunstâncias individuais. Neste artigo, C inclui os grupos de circunstâncias que compõem os sete cenários avaliados.

Deve-se destacar que, para os autores, esse índice é calculado a partir de (5) como uma forma reduzida de um modelo como:

$$y_i = f(C, E, u), \quad (7)$$

$$E = f(C, v), \quad (8)$$

em que y corresponde à nota do aluno no teste. C denota o vetor de circunstâncias. E é o vetor de esforços, que representa todas as variáveis que afetam o resultado educacional, sobre as quais os indivíduos têm alguma medida de controle, e que também pode ser influenciado pelas circunstâncias. u e v são choques aleatórios. Deste modo, na equação (5) as variáveis de esforço são deliberadamente omitidas e β é visto como o efeito de forma reduzida das circunstâncias – diretamente e por meio do esforço.

Contudo, como este trabalho analisa apenas crianças do 5º ano do ensino fundamental, que têm em média 11 anos, considera-se que as diferenças nos resultados dos testes não refletem suas escolhas. Deste modo, seguindo Procópio, Freguglia e Chein (2015), na parcela da desigualdade chamada de esforço, estão características relacionadas a habilidades inatas, fatores genéticos, sorte e outras variáveis não disponíveis no banco de dados (que aqui são interpretadas como outras circunstâncias). Assim, esse grupo de fatores é interpretado como desigualdade residual e revela que a desigualdade de oportunidades calculada é inferior à verdadeira desigualdade de oportunidades.

Deve-se destacar que essas estimativas não podem ser interpretadas como causais, uma vez que componentes do vetor C podem estar correlacionados com outras variáveis de circunstâncias não observadas (omitidas). Entretanto, se o interesse estiver na parcela da variação em y , resultado educacional – que é explicada causalmente pelo efeito conjunto das circunstâncias –, o R quadrado produz um limite inferior válido, sendo uma medida apropriada da desigualdade de oportunidade. Como as únicas variáveis ausentes no modelo são outras circunstâncias, se alguma nova circunstância for adicionada, o R quadrado pode aumentar, mas não cair.

Outra característica importante do R quadrado como medida de desigualdade de oportunidade é que ele permite o uso de um maior número de variáveis de circunstâncias em relação às utilizadas em estudos anteriores, o que atende bem ao interesse desta pesquisa, já que se utiliza um grande número de variáveis.

Por fim, uma das vantagens mais importantes de se utilizar o R quadrado é a possibilidade de decomposição dessa medida nos componentes individuais (ou por grupo) do vetor de circunstâncias. Assim, é possível calcular a contribuição relativa de cada circunstância (ou grupo) na desigualdade de oportunidade educacional. Os autores sugerem a aplicação da decomposição de Shapley-Shorrocks sobre a equação (6) para obter a contribuição de cada circunstância ou grupo.¹⁵ Logo, a desigualdade de oportunidades pode ser reescrita como:

$$IOP = (var y)^{-1} \left[\sum_j \beta_j^2 var C_j + \frac{1}{2} \sum_k \sum_j \beta_k \beta_j cov(C_k, C_j) \right], \quad (9)$$

que, por sua vez, pode ser escrita como a soma da contribuição de todos os elementos j do vetor de circunstâncias C :

$$IOP = \sum_j IOP^j = \sum_j (var y)^{-1} \left[\beta_j^2 var C_j + \frac{1}{2} \sum_k \beta_k \beta_j cov(C_k, C_j) \right]. \quad (10)$$

15. Prova formal em Ferreira, Gignoux e Aran (2011).

4 RESULTADOS

Esta seção apresenta os resultados da pesquisa. Na subseção 4.1, têm-se as estimativas e análise da relação entre a proficiência escolar e as circunstâncias da infância a partir do emprego do GAM. A subseção 4.2 apresenta os resultados da estimação da medida de desigualdade de oportunidade educacional e a contribuição de cada grupo de circunstâncias para ela, com base em Ferreira e Gignoux (2014).

4.1 Estimação do GAM: efeitos das circunstâncias da infância no desempenho escolar

As tabelas 2 e 3 apresentam os resultados do GAM para as proficiências em português e matemática, respectivamente. A partir das referidas tabelas, nas regressões que consideram o cenário base, verifica-se que todas as variáveis são estatisticamente significativas e apresentam os sinais esperados.

A variável referente ao sexo do estudante indica que os meninos apresentam pior resultado em português e melhor resultado em matemática relativamente às meninas, tal como indica a literatura em economia da educação e sobre desigualdade de oportunidades. No que tange à raça, conforme esperado, no grupo mais favorecido em termos de resultado educacional encontram-se os alunos que se autodeclararam brancos, relativamente aos que se autodeclararam pretos, pardos, amarelos e indígenas.

Quanto à escolaridade dos pais, o grau de instrução superior completo da mãe e do pai está associado positivamente com as proficiências em ambas as matérias. A este respeito, de modo geral, a literatura sobre desigualdade de oportunidades educacionais, determinantes do desempenho escolar e transmissão intergeracional da educação aponta para a importância do contexto familiar no aprendizado das crianças, em que pais com maior grau de instrução têm filhos com maior nível educacional.

Segundo Black, Devereux e Salvanes (2005), isso ocorreria por seleção pura – por exemplo, pais com nível de escolaridade mais elevado têm um maior nível de renda e podem investir mais na educação de seus filhos –, ou por efeito de causalidade, em que ter mais educação faz de você um tipo de pai diferente e, portanto, leva seus filhos a obterem melhores resultados educacionais. Deste modo, para os autores a educação pública teria um papel importante no sentido de aumentar a igualdade de oportunidades, e políticas nesse sentido, como ter cidadãos mais instruídos, poderiam ter efeitos prolongados, melhorando os resultados de seus filhos.

A variável correspondente ao índice socioeconômico também apresenta correlação positiva com o desempenho nas avaliações. Esse resultado converge com a literatura, a qual enfatiza a renda ou o nível socioeconômico como uma das principais

variáveis de circunstâncias determinantes da desigualdade de oportunidades educacionais ou dos resultados escolares, tal como verificado em Aguirreche (2012), Foguel e Veloso (2012), Procópio, Freguglia e Chein (2015), Nogueira (2015), Carvalho e Waltenberg (2015) e Souza, Oliveira e Anegues (2018).

Outra variável que apresenta associação positiva com o desempenho nos testes é a relativa a se o aluno cursou creche e/ou pré-escola. Essa evidência é compatível com a maioria dos estudos que analisam a importância dessa variável, como Curi e Menezes-Filho (2009), que encontraram efeitos positivos da creche e pré-escola na conclusão de ciclos escolares posteriores, nos anos de escolaridade e no desempenho escolar. Estudos internacionais têm enfatizado que os primeiros anos de vida são importantes para a formação de habilidades, mais maleáveis nessa fase. Diante disso, argumenta-se que intervenções nessa etapa do ciclo da vida podem promover o aprendizado, refletindo-se em efeitos futuros positivos em resultados escolares, na renda e em comportamentos sociais (Carneiro e Heckman, 2003; Kautz *et al.*, 2014).

Na direção oposta, a variável *tamanho da família* apresentou correlação negativa com os testes de proficiência. Como explicação para isso, sugere-se que famílias maiores realizam escolhas mais restritas quanto à alocação da renda e do tempo na educação das crianças, o que implicaria indiretamente um menor desempenho escolar dos filhos. Outros dois importantes resultados revelam que alunos que já foram reprovados ou abandonaram a escola apresentam piores notas. O efeito negativo do histórico de reprovação do aluno é evidenciado por Aguirreche (2012), porém considerando essa característica como variável de esforço. Ao estudar os determinantes do desempenho escolar, Almeida (2014) verifica efeito negativo de -6,7% do abandono escolar nas notas dos alunos com piores escores nos anos iniciais do ensino fundamental.

TABELA 2

Resultados estimados do modelo aditivo generalizado para o 5º ano do ensino fundamental: portugueses (2017)

Conjuntos de circunstâncias	Cenários						
	1ª	2ª	3ª	4ª	5ª	6ª	7ª
	Base						
Aluno homem	-8,30*** (0,15)	-8,05*** (0,15)	-8,03*** (0,15)	-8,14*** (0,15)	-8,10*** (0,15)	-8,09*** (0,15)	-8,10*** (0,15)
Aluno branco	3,98*** (0,16)	3,90*** (0,16)	3,73*** (0,16)	3,44*** (0,16)	3,34*** (0,16)	3,25*** (0,16)	3,26*** (0,16)
Mãe com curso superior	1,81*** (0,20)	1,73*** (0,20)	1,96*** (0,19)	1,84*** (0,19)	1,78*** (0,19)	1,62*** (0,19)	1,63*** (0,19)
Pai com curso superior	1,59*** (0,20)	1,74*** (0,20)	2,01*** (0,19)	1,81*** (0,19)	1,74*** (0,19)	1,65*** (0,19)	1,65*** (0,19)

(Continua)

(Continuação)

Conjuntos de circunstâncias	Cenários						
	1º	2º	3º	4º	5º	6º	7º
Índice socioeconômico	5,08*** (0,10)	4,72*** (0,10)	4,64*** (0,10)	4,30*** (0,10)	4,21*** (0,10)	4,08*** (0,10)	4,06*** (0,10)
Frequentou creche/pré-escola	4,33*** (0,19)	3,82*** (0,19)	3,35*** (0,19)	3,27*** (0,19)	3,26*** (0,19)	3,23*** (0,19)	3,23*** (0,19)
Tamanho da família	-2,49*** (0,07)	-2,53*** (0,07)	-2,39*** (0,07)	-2,29*** (0,07)	-2,27*** (0,07)	-2,24*** (0,07)	-2,24*** (0,07)
Já foi reprovado	-29,85*** (0,24)	-28,74*** (0,24)	-27,56*** (0,24)	-27,39*** (0,23)	-27,36*** (0,23)	-27,12*** (0,23)	-27,12*** (0,23)
Já abandonou a escola	-7,69*** (0,38)	-6,14*** (0,38)	-5,81*** (0,38)	-5,92*** (0,38)	-5,80*** (0,38)	-5,72*** (0,38)	-5,70*** (0,38)
Escola pública	-12,68*** (0,20)	-12,58*** (0,20)	-12,15*** (0,20)	-12,19*** (0,21)	-11,09*** (0,22)	-7,43*** (0,23)	-7,55*** (0,24)
Escola urbana	8,49*** (0,29)	8,53*** (0,29)	8,55*** (0,29)	7,00*** (0,29)	6,66*** (0,29)	6,82*** (0,30)	6,75*** (0,30)
Trabalho fora de casa	-22,29*** (0,29)	-21,38*** (0,28)	-20,11*** (0,28)	-19,68*** (0,28)	-19,55*** (0,28)	-19,42*** (0,28)	-19,41*** (0,28)
Norte	-9,14*** (0,29)	-8,20*** (0,29)	-6,60*** (0,29)	-4,25*** (0,31)	-3,17*** (0,32)	-3,03*** (0,32)	-2,79*** (0,32)
Nordeste	-11,78*** (0,21)	-10,99*** (0,21)	-10,47*** (0,21)	-8,24*** (0,23)	-7,11*** (0,24)	-6,40*** (0,24)	-6,20*** (0,24)
Sul	-2,29*** (0,22)	-2,09*** (0,22)	-0,97*** (0,22)	-1,94*** (0,23)	-2,27*** (0,23)	-1,96*** (0,23)	-1,75*** (0,23)
Centro-Oeste	-3,14*** (0,29)	-2,37*** (0,29)	-0,80** (0,29)	-0,79** (0,30)	-0,22 (0,30)	0,68* (0,30)	0,86** (0,30)
Relação entre pais e filhos							
Mora com mãe e pai		1,98*** (0,17)	1,83*** (0,17)	1,76*** (0,17)	1,67*** (0,17)	1,65*** (0,16)	1,64*** (0,16)
Vê mãe lendo		2,74*** (0,32)	2,41*** (0,32)	2,46*** (0,32)	2,39*** (0,32)	2,35*** (0,31)	2,34*** (0,31)
Vê pai lendo		-1,04*** (0,23)	-0,77*** (0,23)	-0,69** (0,23)	-0,62** (0,23)	-0,52* (0,23)	-0,52* (0,23)
Frequência às reuniões de pais		4,76*** (0,23)	4,45*** (0,23)	4,35*** (0,23)	4,35*** (0,23)	4,29*** (0,23)	4,28*** (0,23)
Incentivo a estudar		11,80*** (0,74)	10,60*** (0,73)	10,64*** (0,73)	10,58*** (0,72)	10,32*** (0,72)	10,32*** (0,72)
Incentivo a fazer dever de casa		9,38*** (0,56)	8,54*** (0,55)	8,28*** (0,55)	8,22*** (0,55)	8,34*** (0,55)	8,31*** (0,55)
Incentivo a ler		1,17** (0,40)	0,56 (0,40)	0,49 (0,40)	0,60 (0,39)	0,73' (0,39)	0,76' (0,39)
Incentivo a ir à escola		16,82*** (0,52)	16,28*** (0,52)	15,82*** (0,52)	15,82*** (0,51)	15,78*** (0,51)	15,78*** (0,51)

(Continua)

(Continuação)

Conjuntos de circunstâncias	Cenários						
	1º	2º	3º	4º	5º	6º	7º
Conversa		-1,29*** (0,23)	-0,83*** (0,23)	-0,71** (0,23)	-0,67** (0,22)	-0,66** (0,22)	-0,66** (0,22)
Hábitos culturais e uso do tempo							
Lê jornais			-4,15*** (0,16)	-3,69*** (0,16)	-3,63*** (0,16)	-3,59*** (0,16)	-3,58*** (0,16)
Lê livros			8,37*** (0,38)	8,52*** (0,38)	8,47*** (0,38)	8,41*** (0,38)	8,39*** (0,38)
Lê revistas em geral			3,15*** (0,16)	3,14*** (0,16)	3,20*** (0,16)	3,19*** (0,16)	3,19*** (0,16)
Lê revistas em quadrinhos			6,48*** (0,24)	6,32*** (0,24)	6,24*** (0,24)	6,16*** (0,24)	6,17*** (0,24)
Lê sites na internet			5,31*** (0,18)	5,09*** (0,18)	5,09*** (0,18)	5,03*** (0,18)	5,04*** (0,18)
Biblioteca			-1,22*** (0,16)	-1,72*** (0,16)	-1,73*** (0,16)	-1,82*** (0,16)	-1,81*** (0,16)
Cinema			-1,91*** (0,19)	-2,08*** (0,19)	-2,22*** (0,19)	-2,25*** (0,19)	-2,25*** (0,19)
Espetáculos			-6,03*** (0,16)	-5,89*** (0,16)	-5,91*** (0,16)	-5,91*** (0,16)	-5,92*** (0,16)
Tempo vendo TV			2,20*** (0,06)	2,25*** (0,06)	2,24*** (0,06)	2,23*** (0,06)	2,23*** (0,06)
Tempo com trabalho doméstico			-1,05*** (0,06)	-1,05*** (0,06)	-1,06*** (0,06)	-1,06*** (0,06)	-1,06*** (0,06)
Características do professor							
Professor homem				-2,54*** (0,33)	-2,12*** (0,33)	-1,91*** (0,33)	-1,83*** (0,33)
Professor branco				3,64*** (0,24)	2,97*** (0,24)	2,85*** (0,24)	2,79*** (0,24)
Curso superior				7,32*** (0,49)	7,01*** (0,49)	6,40*** (0,49)	6,40*** (0,49)
Experiência				1,80*** (0,23)	1,70*** (0,23)	1,46*** (0,23)	1,49*** (0,23)
Formação continuada				4,36*** (0,32)	3,98*** (0,32)	3,56*** (0,32)	3,51*** (0,32)
Pós-graduação				3,76*** (0,23)	3,58*** (0,23)	3,61*** (0,23)	3,65*** (0,23)
Características do diretor							
Diretor homem					-1,04*** (0,19)	-1,08*** (0,19)	-1,09*** (0,19)
Diretor branco					2,82*** (0,17)	2,54*** (0,17)	2,56*** (0,17)

(Continua)

(Continuação)

Conjuntos de circunstâncias	Cenários						
	1º	2º	3º	4º	5º	6º	7º
Curso superior					2,13*** (0,54)	2,02*** (0,54)	1,99*** (0,54)
Experiência					2,21*** (0,16)	2,05*** (0,16)	2,03*** (0,16)
Formação continuada					1,69*** (0,22)	1,54*** (0,22)	1,50*** (0,22)
Pós-graduação					2,21*** (0,19)	2,03*** (0,19)	2,03*** (0,19)
Características da escola							
Qualidade da escola						1,54*** (0,08)	1,52*** (0,08)
Violência						-2,21*** (0,09)	-2,21*** (0,09)
Características da turma							
Homogênea: idade							0,07 (0,17)
Homogênea: rendimento escolar							1,53*** (0,35)
Heterogênea: idade							-1,62*** (0,40)
Heterogênea: rendimento escolar							1,10*** (0,21)

Fonte: Microdados do Saeb 2017. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.

Elaboração dos autores.

Obs.: 1. Erro-padrão entre parênteses.

2. Significância: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; e ' $p < 0,1$.

A variável referente ao tipo de escola, bastante utilizada na literatura sobre desigualdade de oportunidades, revela que estudar em escola pública influencia negativamente os resultados dos testes, em coerência com os trabalhos de Gamboa e Waltenberg (2011) e Nogueira (2015).

No que tange à variável de localização da escola em área urbana, verifica-se efeito positivo, enquanto o trabalho fora de casa apresenta efeito negativo, como esperado. Ressalta-se que o trabalho infantil reduz o tempo disponível para a execução das atividades escolares, refletindo-se negativamente no desempenho escolar. As últimas variáveis do cenário base referem-se às *dummies* de região, cujos sinais dos coeficientes foram negativos comparativamente à categoria de referência, que é a região Sudeste.

TABELA 3
Resultados estimados do modelo aditivo generalizado para o 5º ano do ensino fundamental: matemática (2017)

Conjuntos de circunstâncias	Cenários						
	1ª	2ª	3ª	4ª	5ª	6ª	7ª
	Base						
Aluno homem	6,81*** (0,15)	6,97*** (0,15)	6,79*** (0,15)	6,70*** (0,15)	6,74*** (0,15)	6,74*** (0,15)	6,73*** (0,15)
Aluno branco	4,55*** (0,16)	4,49*** (0,16)	4,31*** (0,16)	3,80*** (0,16)	3,70*** (0,16)	3,59*** (0,16)	3,61*** (0,16)
Mãe com curso superior	3,18*** (0,19)	3,13*** (0,19)	3,28*** (0,19)	3,18*** (0,19)	3,11*** (0,19)	2,94*** (0,19)	2,96*** (0,19)
Pai com curso superior	2,13*** (0,19)	2,29*** (0,19)	2,53*** (0,19)	2,23*** (0,19)	2,12*** (0,19)	2,04*** (0,19)	2,03*** (0,19)
Índice socioeconômico	5,25*** (0,10)	4,96*** (0,10)	4,99*** (0,10)	4,65*** (0,10)	4,54*** (0,10)	4,38*** (0,10)	4,36*** (0,10)
Frequentou creche/pré-escola	6,17*** (0,19)	5,73*** (0,19)	5,29*** (0,19)	5,11*** (0,18)	5,12*** (0,18)	5,07*** (0,18)	5,06*** (0,18)
Tamanho da família	-1,87*** (0,07)	-1,94*** (0,07)	-1,83*** (0,07)	-1,72*** (0,07)	-1,70*** (0,07)	-1,66*** (0,07)	-1,66*** (0,07)
Já foi reprovado	-29,05*** (0,23)	-28,15*** (0,23)	-27,14*** (0,23)	-27,01*** (0,23)	-26,96*** (0,23)	-26,68*** (0,23)	-26,68*** (0,23)
Já abandonou a escola	-7,43*** (0,38)	-6,25*** (0,38)	-6,02*** (0,37)	-6,10*** (0,37)	-5,94*** (0,37)	-5,85*** (0,37)	-5,82*** (0,37)
Escola pública	-14,04*** (0,20)	-14,00*** (0,20)	-13,66*** (0,20)	-13,25*** (0,20)	-11,73*** (0,21)	-7,74*** (0,23)	-8,11*** (0,23)
Escola urbana	6,25*** (0,29)	6,32*** (0,29)	6,46*** (0,29)	5,08*** (0,29)	4,68*** (0,29)	4,81*** (0,29)	4,67*** (0,29)
Trabalho fora de casa	-17,49*** (0,28)	-16,77*** (0,28)	-15,66*** (0,28)	-15,25*** (0,28)	-15,11*** (0,28)	-14,96*** (0,28)	-14,94*** (0,28)
Norte	-12,62*** (0,28)	-11,74*** (0,28)	-10,25*** (0,29)	-6,83*** (0,31)	-5,65*** (0,31)	-5,45*** (0,31)	-4,96*** (0,32)
Nordeste	-15,83*** (0,21)	-15,14*** (0,21)	-14,57*** (0,21)	-11,21*** (0,23)	-9,87*** (0,23)	-9,01*** (0,24)	-8,62*** (0,24)
Sul	-0,97*** (0,22)	-0,77*** (0,22)	0,13 (0,22)	-1,53*** (0,22)	-1,87*** (0,22)	-1,48*** (0,22)	-1,05*** (0,23)
Centro-Oeste	-6,08*** (0,28)	-5,33*** (0,28)	-3,77*** (0,28)	-3,04*** (0,29)	-2,42*** (0,29)	-1,40*** (0,29)	-1,03*** (0,29)
	Relação entre pais e filhos						
Mora com mãe e pai		1,98*** (0,16)	1,81*** (0,16)	1,68*** (0,16)	1,58*** (0,16)	1,55*** (0,16)	1,53*** (0,16)
Vê mãe lendo		2,49*** (0,31)	2,39*** (0,31)	2,5*** (0,31)	2,41*** (0,31)	2,36*** (0,31)	2,35*** (0,31)

(Continua)

(Continuação)

Conjuntos de circunstâncias	Cenários						
	1º	2º	3º	4º	5º	6º	7º
Vê pai lendo		-0,92*** (0,23)	-0,55* (0,23)	-0,50* (0,23)	-0,43' (0,23)	-0,33 (0,22)	-0,31 (0,22)
Frequência às reuniões de pais		4,63*** (0,23)	4,41*** (0,23)	4,27*** (0,23)	4,27*** (0,22)	4,19*** (0,22)	4,18*** (0,22)
Incentivo a estudar		7,34*** (0,72)	6,37*** (0,72)	6,44*** (0,71)	6,35*** (0,71)	6,08*** (0,71)	6,06*** (0,71)
Incentivo a fazer dever de casa		4,72*** (0,55)	4,12*** (0,55)	3,92*** (0,54)	3,83*** (0,54)	3,95*** (0,54)	3,91*** (0,54)
Incentivo a ler		2,54*** (0,39)	2,17*** (0,39)	1,99*** (0,39)	2,15*** (0,39)	2,29*** (0,39)	2,34*** (0,39)
Incentivo a ir à escola		14,21*** (0,51)	13,75*** (0,51)	13,22*** (0,51)	13,23*** (0,51)	13,19*** (0,50)	13,19*** (0,50)
Conversa		-2,44*** (0,22)	-1,84*** (0,22)	-1,79*** (0,22)	-1,74*** (0,22)	-1,73*** (0,22)	-1,73*** (0,22)
Hábitos culturais e uso do tempo							
Lê jornais			-3,56*** (0,16)	-3,00*** (0,16)	-2,92*** (0,16)	-2,87*** (0,16)	-2,86*** (0,16)
Lê livros			6,67*** (0,37)	6,78*** (0,37)	6,72*** (0,37)	6,64*** (0,37)	6,63*** (0,37)
Lê revistas em geral			2,56*** (0,16)	2,52*** (0,16)	2,59*** (0,16)	2,59*** (0,16)	2,59*** (0,16)
Lê revistas em quadrinhos			4,70*** (0,24)	4,60*** (0,23)	4,50*** (0,23)	4,41*** (0,23)	4,40*** (0,23)
Lê sites na internet			2,53*** (0,18)	2,33*** (0,18)	2,33*** (0,18)	2,27*** (0,18)	2,28*** (0,18)
Biblioteca			-0,11 (0,16)	-0,68*** (0,16)	-0,70*** (0,16)	-0,82*** (0,16)	-0,80*** (0,16)
Cinema			-2,12*** (0,19)	-2,22*** (0,19)	-2,40*** (0,19)	-2,44*** (0,19)	-2,46*** (0,19)
Espetáculos			-4,99*** (0,16)	-4,88*** (0,15)	-4,90*** (0,15)	-4,90*** (0,15)	-4,90*** (0,15)
Tempo vendo TV			2,31*** (0,06)	2,35*** (0,06)	2,35*** (0,06)	2,33*** (0,06)	2,33*** (0,06)
Tempo com trabalho doméstico			-0,99*** (0,06)	-0,94*** (0,06)	-0,94*** (0,06)	-0,94*** (0,06)	-0,93*** (0,06)
Características do professor							
Professor homem				-1,14*** (0,32)	-0,81* (0,33)	-0,54' (0,32)	-0,37 (0,32)
Professor branco				6,60*** (0,24)	5,82*** (0,24)	5,67*** (0,24)	5,57*** (0,24)

(Continua)

(Continuação)

Conjuntos de circunstâncias	Cenários						
	1º	2º	3º	4º	5º	6º	7º
Curso superior				7,22*** (0,48)	6,74*** (0,48)	6,10*** (0,48)	5,99*** (0,48)
Experiência				3,19*** (0,23)	3,09*** (0,23)	2,83*** (0,23)	2,83*** (0,23)
Formação continuada				6,21*** (0,32)	5,76*** (0,32)	5,24*** (0,32)	5,16*** (0,32)
Pós-graduação				3,79*** (0,22)	3,52*** (0,23)	3,51*** (0,22)	3,61*** (0,22)
Características do diretor							
Diretor homem					-0,25 (0,19)	-0,27 (0,19)	-0,27 (0,19)
Diretor branco					3,08*** (0,17)	2,77*** (0,17)	2,81*** (0,17)
Curso superior					3,94*** (0,53)	3,76*** (0,53)	3,75*** (0,53)
Experiência					3,09*** (0,16)	2,88*** (0,16)	2,86*** (0,16)
Formação continuada					2,49*** (0,21)	2,33*** (0,21)	2,25*** (0,21)
Pós-graduação					2,99*** (0,19)	2,81*** (0,19)	2,82*** (0,19)
Características da escola							
Qualidade da escola						1,75*** (0,07)	1,70*** (0,07)
Violência						-2,58*** (0,09)	-2,57*** (0,09)
Características da turma							
Homogênea: idade							0,10 (0,17)
Homogênea: rendimento escolar							1,23*** (0,34)
Heterogênea: idade							-1,28** (0,39)
Heterogênea: rendimento escolar							2,41*** (0,21)

Fonte: Microdados do Saeb 2017. Disponível em: <https://www.gov.br/inep/pt-br/ acesso-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.

Elaboração dos autores.

Obs.: 1. Erro-padrão entre parênteses.

2. Significância: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; e ' $p < 0,1$.

No segundo cenário, que engloba o conjunto base e o conjunto de características do relacionamento entre pais e filhos, as estatísticas estimadas mostram que os parâmetros das variáveis *base* de modo geral caem, mas continuam significantes, e seus sinais não apresentam mudanças. Deve-se destacar que, além das tradicionais variáveis de *background* familiar empregadas na literatura econômica sobre desigualdade de oportunidades como renda, ocupação e escolaridade dos pais, variáveis de *background* que caracterizam a interação entre pais e filhos são importantes para o desenvolvimento das habilidades e comportamentos presentes e futuros das crianças.¹⁶

Segundo Coleman (1988), se o capital humano dos pais não é complementado pelo capital social incorporado nas relações familiares (por exemplo, se é empregado exclusivamente no trabalho), é irrelevante para o crescimento educacional da criança que os pais tenham uma grande ou pequena quantidade de capital humano. Por isso a importância de se estudar o efeito da relação entre pais e filhos na formação de capital humano das crianças, o que reforça ainda mais o papel das circunstâncias nos resultados escolares dos estudantes.

As variáveis de incentivo – principalmente o incentivo a estudar, fazer dever de casa e ir à escola – apresentaram forte correlação positiva com os testes, sendo o peso destas variáveis ainda maior do que o de outras importantes do conjunto base, como escola urbana, sexo, abandono escolar e índice socioeconômico. Essas evidências do importante papel das variáveis relacionais¹⁷ para o desempenho escolar corroboram Tramonte e Willms (2010) e Souza, Oliveira e Annes (2018).

A estrutura familiar está positivamente correlacionada com a proficiência escolar. Essa variável representa as crianças que moram com ambos os pais relativamente àquelas que moram apenas com a mãe ou o pai. De acordo com a literatura, alguns possíveis canais de transmissão do efeito negativo da família monoparental seriam a restrição de renda, ausência do pai, o estresse na dissolução conjugal, a redução no tempo despendido no monitoramento e interação com os filhos e a alteração nas redes de contatos (McLanahan, 1985; Carlson e Corcoran, 2001; Vasconcelos, Ribeiro e Fernandez, 2017).

As variáveis relativas à leitura dos pais e à frequência às reuniões dos pais representam a valorização da leitura e educação por parte destes, sinalizando a importância da educação no contexto familiar. A frequência às reuniões também é uma forma de acompanhamento e controle dos resultados educacionais dos filhos. Como esperado, a variável *vê mãe lendo e frequência às reuniões de pais* apresentaram

16. Conforme Coleman (1988), o *background* familiar é dividido em pelo menos três componentes diferentes: capital econômico, capital humano e capital social. O capital econômico é dado pela riqueza ou renda da família, o capital humano pela educação dos pais e o capital social pelas relações entre filhos e pais, importantes para o desenvolvimento intelectual das crianças.

17. Em alguns momentos do texto, as variáveis de representam a relação entre pais e filhos serão chamadas de relacionais.

associação positiva com os testes de aprendizagem. Já a variável de leitura do pai mostrou-se significativa, mas com sinal negativo, o que requer maior investigação. Também ao contrário do previsto, a variável *conversa* mostrou-se negativamente correlacionada com o aprendizado do aluno. Neste caso, talvez a forma como os pais abordam os filhos seja mais importante para explicar essa relação, o que os dados disponíveis não permitem captar.

O terceiro cenário inclui os hábitos culturais e de uso do tempo. Bourdieu (2007) relaciona o sucesso escolar das crianças à distribuição do capital cultural, em que o rendimento escolar depende do capital cultural previamente investido pela família,¹⁸ o que significa que a desigualdade desse capital se reflete na desigualdade educacional. Assim, a família é vista como responsável pela acumulação inicial de capital cultural, que promove o desenvolvimento de habilidades e melhores resultados escolares. Considerando que os hábitos de leitura da criança, de frequência à biblioteca, ao cinema, a espetáculos em geral, e a administração do uso do tempo são tipos de capital cultural diretamente ligados ao capital cultural da família, estas variáveis são de significativa relevância para este estudo.

Quando se incorporam essas características na análise, verifica-se pouca alteração nas magnitudes, significância e sinais dos parâmetros. Aqui observa-se principalmente efeitos positivos dos hábitos de leitura de livros, revistas e *sites* na internet¹⁹ nos resultados dos testes, o que confirma a importância do capital cultural. Já o tempo gasto com trabalho doméstico apresentou associação negativa. Neste caso, entende-se que mais tempo dedicado a atividades não relacionadas à educação e menos prazerosas para as crianças influencia negativamente o desempenho escolar.

Os efeitos da frequência à biblioteca, cinema e espetáculos mostraram-se não significantes ou negativos, enquanto o efeito do uso do tempo assistindo TV, jogando e navegando na internet mostrou-se positivo, o que pode ser explicado pela prevalência de alunos de escola pública e menor nível socioeconômico na amostra. No primeiro caso, a oferta limitada dessas atividades no local de moradia é outra possível explicação. No último caso, outra hipótese para o resultado obtido pode ser a de que, em ambientes menos favorecidos, o uso do tempo em atividades prazerosas e desafiadoras como jogos pode contribuir para despertar a criatividade das crianças, afetando positivamente suas habilidades cognitivas.

O trabalho de Jaeger (2011) corrobora estes últimos resultados. O autor verifica que o efeito do capital cultural varia conforme o *status* socioeconômico

18. O autor define três tipos de capital cultural: incorporado, o qual pressupõe inculcação, assimilação e requer tempo, sendo fortemente dependente da acumulação inicial dos membros da família; objetivado, sob a forma de bens culturais (quadros, livros, dicionários, esculturas, pinturas etc); institucionalizado, na forma de certificado escolar (Bourdieu, 2007).

19. A variável *lê jornais* apresentou sinal negativo, talvez pelo fato de se tratar de um tipo de leitura e veículo de informação que não desperta interesse e empolgação na criança.

(*socioeconomic status* – SES). Consta que indicadores de participação cultural (como frequência a museus, eventos musicais, teatro) e hábitos de leitura têm efeitos mais fortes no desempenho acadêmico em ambientes com alto SES do que em ambientes com baixo SES. Indicadores como número de livros, atividades extracurriculares e o incentivo à prática de *hobbies*, por sua vez, têm efeitos mais fortes em ambientes com baixo e médio SES do que em ambientes com alto SES. Assim, o autor argumenta que o investimento em aspectos práticos do capital cultural produz uma vantagem comparativa em ambientes de baixo SES, nos quais os pais têm menor probabilidade de investir em capital cultural.²⁰

Sabendo que, além do contexto da família, o ambiente escolar é mais um determinante dos resultados educacionais das crianças, essa pesquisa incluiu variáveis características da escola como circunstâncias determinantes da desigualdade de oportunidades educacionais. Isso se explica porque a escolha de uma escola com certas características não é feita pelo aluno, mas pela família, e porque a qualidade da escola é dada pelo lado da oferta do serviço educacional. As variáveis referentes ao professor, ao diretor, à escola e à turma compõem os cenários quatro, cinco, seis e sete, respectivamente.

Quando se incluem as características do professor também, não há variação significativa nas variáveis dos demais conjuntos considerados. Observa-se que o fato de o professor ser do sexo masculino impacta negativamente a proficiência dos alunos. Como em torno de 51% dos estudantes são do sexo feminino, isso implica que a diferença de gênero professor-aluno pode estar afetando sua motivação, expectativa e desempenho. Ressalta-se ainda que a grande maioria dos professores do 5º ano no Brasil são mulheres (87% da amostra), e essa percepção também pode afetar a motivação dos estudantes. Com relação à raça, o professor ser de cor branca afeta positivamente o desempenho nos testes. Como boa parte da amostra (63%) é composta por alunos que não se autodeclararam brancos, pode-se dizer que a diferença de raça em relação ao aluno não diminui sua motivação e seus resultados escolares.

O primeiro resultado corrobora – e o segundo se contrapõe a – pesquisas sobre o assunto, pois a literatura indica que as correspondências de gênero e raça entre aluno e professor podem influenciar positivamente os resultados educacionais, como visto em Dee (2005) e Egalite, Kisida e Winters (2015). Analogamente, as diferenças influenciam esses resultados negativamente. De acordo com Dee (2005), a explicação para isso envolve o que a literatura chama de efeitos passivos e ativos do professor. Os efeitos passivos surgem da identidade racial, étnica e de gênero entre professor e aluno. Eles ocorrem quando a simples presença de um professor da

20. O autor explica esses resultados com base no modelo de reprodução cultural de Bourdieu e no modelo de mobilidade cultural de DiMaggio (1982). Para mais detalhes, ver Jaeger (2011, p. 284).

mesma raça, por exemplo, eleva a motivação e as expectativas acadêmicas do aluno. Por seu turno, os efeitos ativos constituem vieses nas interações com alunos que têm diferentes características étnico-raciais e de gênero. Nesse caso, as percepções dos professores podem influenciar o acesso dos alunos a futuras oportunidades educacionais e dificultar a aprendizagem.

No que tange ao fato de o professor ter curso superior, dez anos ou mais de experiência, participar de cursos de formação continuada e ter pós-graduação, constata-se impacto positivo sobre os testes de aprendizagem. Os trabalhos de Hanushek *et al.* (2005), Buddin e Zamarro (2009) e Canales e Maldonado (2018) mostram relação positiva entre o desempenho dos alunos e a experiência do professor, e alguma evidência do efeito da formação superior e do treinamento em serviço pode ser encontrada em Harris e Sass (2007).

O cenário cinco engloba as características do diretor. Os resultados são semelhantes aos do professor. Observa-se que o diretor do sexo masculino afeta negativamente a proficiência em português, e este ser de cor branca tem correlação positiva com ambos os testes. As variáveis relativas à formação superior, experiência, formação continuada e pós-graduação apresentam parâmetros com sinais positivos, evidências que fazem sentido na medida em que formação superior e pós-graduação significam mais instrução, podendo implicar maior produtividade, e experiência e formação continuada são formas de treinamento que também podem resultar em maior produtividade.

Algumas características da escola que independem do aluno e podem impactar seu desempenho são inseridas no sexto cenário. O índice que representa a qualidade da escola apresentou sinal positivo e estatisticamente significativo, o que significa que quanto melhor a qualidade da escola, melhor o desempenho em português e matemática. Já o índice que corresponde à violência mostra associação negativa, indicando que quanto maior a violência, menor o desempenho do aluno. Os resultados de Nogueira (2015) apontam a qualidade da escola como um dos principais fatores para a desigualdade de oportunidade no Brasil. Quanto à violência na escola, alguns estudos recentes têm focado em seu efeito sobre o desempenho escolar, como Grogger (1997), Teixeira e Kassouf (2015), Lombardi Filho e Oliveira (2017), Cittadin e França (2018), entre outros.

O sétimo cenário introduz variáveis relativas ao critério de formação de turmas com o objetivo de analisar os efeitos dos pares (*peer effects*) sobre o desempenho acadêmico individual do aluno. Verifica-se que turmas homogêneas quanto ao rendimento escolar têm maior correlação positiva com o desempenho em português, e turmas heterogêneas, em matemática. Por sua vez, turmas heterogêneas quanto à idade associam-se negativamente com ambas as matérias, mas mais intensamente com português. Deve-se enfatizar que a importância dos efeitos dos pares sobre o

desempenho do aluno é confirmada pela literatura empírica em vários trabalhos como Hoxby (2000), Hanushek *et al.* (2001), Kang (2007), Jales (2010), Firpo e Jales (2011) e Burke e Sass (2013).

Com relação à magnitude dos parâmetros, no primeiro cenário observa-se que as circunstâncias de maior associação com o desempenho no exame de português e matemática são os fatos de o aluno já ter sido reprovado, trabalhar fora de casa e estudar em escola pública. Quando se olha para o segundo cenário, adicionam-se variáveis importantes do conjunto referente à relação entre pais e filhos, como o incentivo a ir à escola, estudar e fazer dever de casa. No terceiro cenário, destaca-se a forte correlação das variáveis de capital cultural relativas à leitura de livros e revistas. Nos cenários quatro e cinco, as variáveis de formação superior do professor e do diretor são as mais importantes dentre as avaliadas. Quanto às características da escola, ressalta-se o efeito negativo da violência escolar. Já com relação aos efeitos dos pares, se sobressai o efeito negativo da heterogeneidade da turma quanto à idade para a proficiência em português, e o efeito positivo da heterogeneidade quanto ao rendimento escolar para matemática.

É importante destacar que, no geral, as variáveis relacionais e culturais têm maior associação com a proficiência em português, enquanto as variáveis ligadas a escola apresentam maior correlação com o desempenho em matemática. O fato de que a habilidade de leitura seja estimulada mais precocemente e intensamente em casa, e a habilidade em matemática, mais tardiamente e intensamente na escola, constitui uma possível explicação para isso.

Pode-se dizer que as circunstâncias da infância estudadas têm um papel fundamental na determinação da desigualdade de oportunidades educacionais, com as novas variáveis de *background* familiar referentes à relação entre pais e filhos e aos hábitos culturais, e as variáveis ligadas à escola, revelando-se de grande importância. Observa-se que, quanto melhor o conjunto de oportunidades, maior o desempenho educacional do estudante, e quanto pior o conjunto de oportunidades, menor o desempenho educacional do aluno.

4.2 Estimação do IOP educacional

Esta subseção apresenta as estimativas do IOP educacional, bem como sua decomposição, segundo a metodologia proposta por Ferreira e Gignoux (2014). Essa medida é o R quadrado da regressão e representa a parcela da variância total (desigualdade total) no desempenho do estudante no Saeb que é explicada pelas variáveis de circunstâncias.

A tabela 4 mostra a medida de desigualdade de oportunidades a partir de uma estimação para cada conjunto de variáveis de não responsabilidade isoladamente. Com relação ao desempenho em português, o IOP é maior quando se considera

o conjunto base de circunstâncias da infância (24,2%). Em seguida, encontram-se os conjuntos referentes aos hábitos culturais e uso do tempo (7,09%), à escola (6,01%), ao diretor (5,1%), à relação entre pais e filhos (4,79%), ao professor (4,48%) e à turma (0,46%).

Considerando a proficiência em matemática, novamente o conjunto base aparece em primeiro lugar no *ranking* da desigualdade injusta (24,6%), seguido pelas características da escola (7,4%), pelo conjunto relativo aos hábitos culturais e uso do tempo (6,78%), às características do diretor (6,71%), do professor (6,58%), ao conjunto referente à relação entre pais e filhos (4,09%) e às características da turma (0,65%). Estes *rankings* dão uma boa sinalização da importância dos grupos estudados como determinantes da parcela injusta da desigualdade educacional.

TABELA 4

Desigualdade de oportunidades por conjunto de circunstâncias

Conjuntos de circunstâncias	Português	Matemática
Base	0,2420	0,2460
Relação entre pais e filhos	0,0479	0,0409
Hábitos culturais e uso do tempo	0,0709	0,0678
Características do professor	0,0448	0,0658
Características do diretor	0,0510	0,0671
Características da escola	0,0601	0,0740
Características da turma	0,0046	0,0065

Fonte: Microdados do Saeb 2017. Disponível em: <https://www.gov.br/inep/pt-br/ acesso-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.

Elaboração dos autores.

Verifica-se ainda que os grupos de variáveis relacionais e culturais apresentam maiores IOPs para o desempenho em português do que em matemática, enquanto os grupos de variáveis de professor, diretor, escola e turma são mais relevantes para matemática, o que corrobora as estimações do modelo GAM.

Dado que a desigualdade de oportunidades é maior quando se considera o grupo de variáveis base, o exercício de decomposição do R quadrado referente a este grupo é relevante. Conforme a tabela A.1 do apêndice A, a variável reprovação apresenta a maior contribuição (28,88% para português e 25,96% para matemática), seguida pela participação conjunta da educação dos pais e do índice socioeconômico – português (15,90%) e matemática (19,67%).

A contribuição da educação dos pais e do índice socioeconômico está em conformidade com a literatura sobre o tema, como verificado em Foguel e Veloso (2012) para a educação infantil; Procópio, Freguglia e Chein (2015) para o ensino fundamental; Nogueira (2015) para o ensino médio; Carvalho e Waltenberg (2015) para o ensino superior; e Gamboa e Waltenberg (2011) com base em dados do Pisa.

A variável *reprovação* é tratada, em geral, como esforço, mas sua importância é documentada pela literatura sobre os determinantes do desempenho escolar.

A relevância do conjunto base está bem consolidada e foi ratificada. Com relação à escola, as evidências são mais escassas. Tavares, Camelo e Paciência (2018), por exemplo, mostram que grande parcela da desigualdade de notas está sob o controle de políticas educacionais estabelecidas pelos órgãos centrais e diretores escolares, com o restante estando relacionado às circunstâncias fora do controle dessas políticas. Já quanto aos conjuntos referentes à relação entre pais e filhos e aos aspectos culturais, não há, na literatura sobre desigualdade de oportunidade, evidências do estudo de sua importância, tratando-se portanto de uma contribuição significativa desta pesquisa.

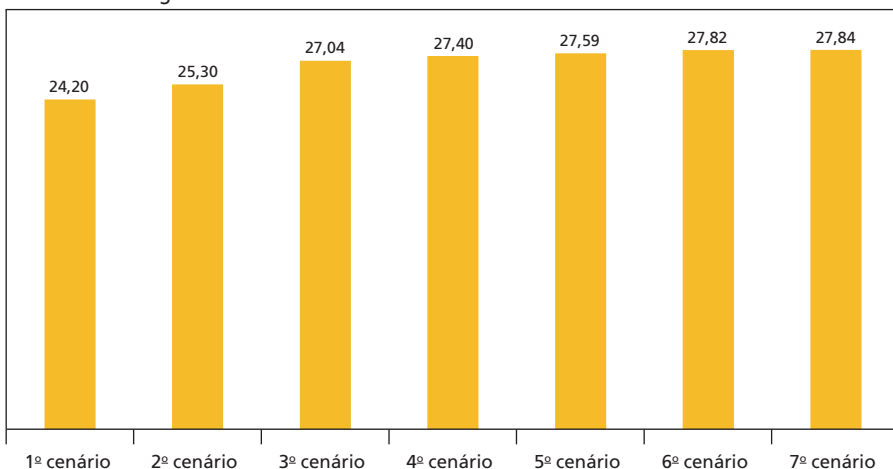
O gráfico 1 aponta como as estimativas do IOP para as habilidades em português e matemática variam à medida que se introduzem, sequencialmente, os conjuntos de variáveis definidos neste estudo. Acrescentar novas circunstâncias aumenta em 15% e 13%, respectivamente, o IOP para português (de 24,20% para 27,84%) e matemática (24,60% para 27,87%) considerando o primeiro e último cenários. No último cenário, verifica-se que o IOP é um pouco maior para matemática do que para português.

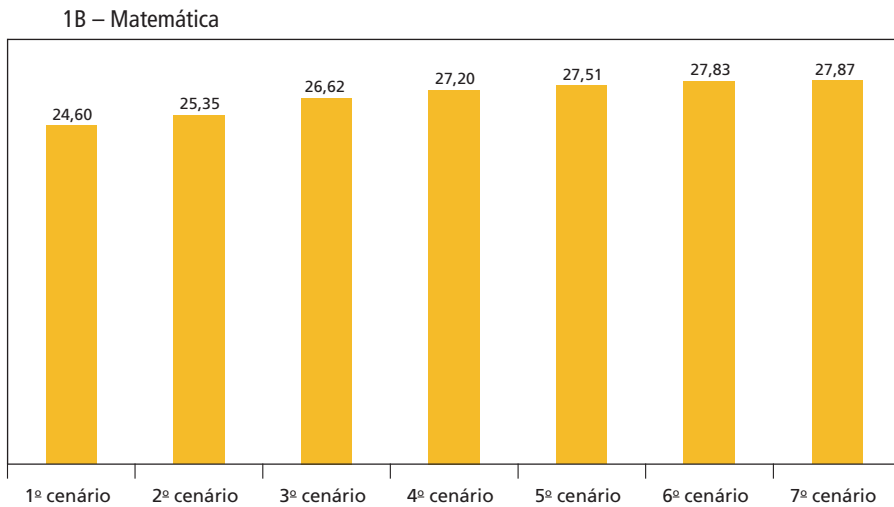
Outro resultado importante é a evidência de que os maiores aumentos ocorrem entre o primeiro e segundo, e entre o segundo e terceiro cenários, que representam a inclusão dos conjuntos de variáveis de relação entre pais e filhos e culturais, ratificando a importância deles. Esses aumentos são de 4,5% e 6,9% para português e de 3,1% e 5,0% para matemática.

GRÁFICO 1

IOP com diferentes cenários de circunstâncias da infância: português e matemática (Em %)

1A – Português





Fonte: Microdados do Saeb 2017. Disponível em: <https://www.gov.br/inep/pt-br/ acesso-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.
Elaboração dos autores.

Já as tabelas 5A e 5B apresentam a contribuição de cada conjunto de variáveis para a medida de desigualdade de oportunidades em cada cenário, calculada pela decomposição de Shapley-Shorroks. As referidas tabelas correspondem aos resultados para português e matemática, respectivamente.

Os resultados indicam que o conjunto base representa predominantemente as variáveis mais importantes para a desigualdade injusta. Considerando todos os cenários analisados e ambas as avaliações, sua participação relativa é sempre maior que 55%. No entanto, as contribuições dos demais grupos também são significativas. Na sequência, para a proficiência em português, encontram-se os conjuntos de variáveis de *background* correspondentes aos hábitos culturais e à relação entre pais e filhos, cujos pesos relativos são superiores a 15% e 8%. Para matemática, encontram-se os hábitos culturais (superior a 10%) e o conjunto de características da escola (superior a 9%).

Adicionalmente, observa-se novamente que o conjunto de variáveis culturais e da relação entre pais e filhos contribuem mais para o desempenho em português do que em matemática, enquanto os grupos das variáveis de escola têm maior contribuição para a proficiência em matemática.

As análises realizadas revelam a importância das circunstâncias da infância para desigualdade de oportunidades educacionais, particularmente a relevante contribuição de novas variáveis de não responsabilidade – culturais, de relação entre pais e filhos, e de escola. As evidências mostram que a não utilização das variáveis aqui estudadas como circunstâncias da infância pode subestimar a verdadeira

desigualdade de oportunidades, atribuindo responsabilidade às crianças por algo que está fora do seu controle e que pode ser objeto de políticas públicas no sentido de reduzir a desigualdade injusta.

TABELA 5

Peso de cada conjunto sobre o efeito total da desigualdade de oportunidades por cenário: português e matemática

5A – Português

Conjuntos de circunstâncias	Cenários						
	1ª	2ª	3ª	4ª	5ª	6ª	7ª
IOP	0,2420	0,2530	0,2704	0,2740	0,2759	0,2782	0,2784
Base (%)	-	88,36	74,63	68,92	64,24	59,25	59,02
Relação entre pais e filhos (%)		11,64	9,83	9,19	8,78	8,40	8,37
Hábitos culturais e uso do tempo (%)			15,54	14,49	13,67	12,89	12,83
Características do professor (%)				7,40	6,25	5,57	5,51
Características do diretor (%)					7,06	6,05	6,00
Características da escola (%)						7,84	7,78
Características da turma (%)							0,49

5B – Matemática

Conjuntos de circunstâncias	Cenários						
	1ª	2ª	3ª	4ª	5ª	6ª	7ª
IOP	0,2460	0,2535	0,2662	0,2720	0,2751	0,2783	0,2787
Base (%)	-	90,44	77,31	68,27	62,06	56,02	55,70
Relação entre pais e filhos (%)		9,56	8,25	7,46	7,00	6,61	6,56
Hábitos culturais e uso do tempo (%)			14,44	12,94	11,92	11,01	10,93
Características do professor (%)				11,33	9,56	8,56	8,45
Características do diretor (%)					9,46	8,13	8,05
Características da escola (%)						9,67	9,58
Características da turma (%)							0,74

Fonte: Microdados do Saeb 2017. Disponível em: <https://www.gov.br/inep/pt-br/acao-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.

Elaboração dos autores.

5 CONSIDERAÇÕES FINAIS

O objetivo deste artigo foi analisar como as circunstâncias da infância explicam as desigualdades de oportunidades educacionais no Brasil para alunos do quinto ano do ensino fundamental. Para isso, assumiu-se que toda desigualdade de resultado escolar é decorrente de fatores de não responsabilidade, dado que durante a infância o comportamento das crianças é influenciado basicamente por fatores ambientais. Ademais, incluíram-se novas variáveis de circunstâncias da infância relativas à relação entre pais e

filhos, aos hábitos culturais e à escola. Para que esse objetivo fosse alcançado, utilizou-se um modelo não paramétrico, o GAM, e calculou-se um IOP.

A análise dos determinantes do desempenho educacional, representado pelo resultado do Saeb em português e matemática, revelou principalmente que alunos típicos de um ambiente desfavorecido já foram reprovados, trabalham fora de casa, estudam em escola pública, moram nas regiões Nordeste e Norte, em áreas rurais, já abandonaram a escola, são meninos no caso do desempenho em português, e meninas no de matemática, possuem pais que não incentivam a ir à escola, estudar e fazer dever de casa, não têm o hábito de ler, possuem professores e diretores escolares que não têm capacitação adequada para a função, estudam em escolas de pior qualidade, sofrem influência da violência escolar e estudam em turmas heterogêneas quanto à idade. O contrário é observado para o grupo de alunos mais favorecidos, o que comprova o papel fundamental das circunstâncias da infância na desigualdade de oportunidades educacionais.

Deve-se destacar que o conjunto base é o mais importante, mas os grupos de variáveis relacionais e culturais mostraram-se fortes determinantes dos resultados de ambos os testes, apresentando maior associação com o desempenho em português que em matemática. Os grupos de variáveis da escola também se mostraram relevantes, principalmente para a nota de matemática.

Estes resultados são ratificados pelo cálculo do IOP educacional para cada grupo e sua contribuição em cada cenário. Considerando cada grupo isoladamente, o conjunto de circunstâncias da infância com maior IOP em ambos os testes é de fato o conjunto base, mas todos os demais apresentam índices consideráveis. Na análise por cenário, a decomposição dos índices de desigualdade por conjunto de circunstâncias novamente ratifica o conjunto base como o mais importante, seguido pelos grupos de hábitos culturais e de variáveis relacionais para o desempenho em português, e pelo de hábitos culturais e características da escola para matemática.

Comparando a importância dos conjuntos de circunstâncias para cada matéria de conhecimento, as variáveis culturais e de relação entre pais e filhos contribuem mais para o desempenho em português do que em matemática, enquanto os grupos das variáveis de escola têm maior contribuição para a proficiência em matemática. De modo geral, acrescentar novas circunstâncias aumenta em 15% e 13%, respectivamente, o IOP para português e matemática.

Pode-se dizer que os resultados obtidos enfatizam, além do já constatado para as variáveis base da literatura, a necessidade do aprofundamento de discussões acerca do papel do capital social e cultural das famílias, bem como da escola, na busca de alternativas para a melhoria no desempenho dos alunos e redução da desigualdade educacional. Políticas que incentivem uma maior interação entre pais e filhos, principalmente no que tange aos aspectos educacionais,

que possibilitem um maior acesso a atividades culturais em casa e fora de casa por parte das crianças, desde o início da infância, que visem maior qualificação dos professores e diretores, redução da violência nas escolas e melhoria na sua infraestrutura, têm o potencial de fazer diferença nos resultados educacionais das crianças e em seus resultados econômicos futuros, nivelando um pouco mais o ponto de partida e tornando o resultado final mais justo, ou menos injusto.

Por fim, deve-se ressaltar que, tendo em vista que as circunstâncias dos primeiros anos de vida são importantes preditoras do desenvolvimento da criança e de diferentes resultados na vida adulta, não controlar variáveis relevantes desta fase é uma limitação deste trabalho, dada a indisponibilidade de dados, o que se configura num problema para a melhor exploração do tema. Neste sentido, uma sugestão para pesquisas futuras seria utilizar os microdados da ANA como forma de verificar se, para idades mais jovens, os resultados do estudo são mais robustos. Estes dados estão disponíveis para os anos de 2014 e 2016. Outra sugestão seria comparar esses resultados para diferentes anos escolares, no intuito de observar como os efeitos das circunstâncias analisadas variam de acordo com a idade dos alunos.

REFERÊNCIAS

- AGUIRRECHE, A. L. **Inequality of opportunity in education**. 2012. 31 f. Dissertation (Master's Degree) – Universidad del País Vasco, Comunidad Autónoma del País Vasco, 2012.
- ALMEIDA, A. T. C. de. Determinantes dos piores e melhores resultados educacionais dos alunos da rede pública de ensino fundamental no Brasil. **Planejamento e Políticas Públicas**, n. 42, p. 147-187, jan.-jun. 2014.
- BARROS, R. P. *et al.* **Measuring inequality of opportunities in Latin America and the Caribbean**. Washington: The World Bank; Nova York: Palgrave Macmillan, 2009.
- BARROS, R. P.; MENDONÇA, R. S. P. **Os determinantes da desigualdade no Brasil**. Rio de Janeiro: Ipea, jul. 1995. (Texto para Discussão, n. 377).
- BLACK, S. E.; DEVEREUX, P. J.; SALVANES, K. G. Why the apple doesn't fall far: understanding intergenerational transmission of human capital. **The American Economic Review**, v. 95, n. 1, p. 437-449, mar. 2005.
- BOURDIEU, P. Os três estados do capital cultural. *In*: NOGUEIRA, M. A.; CATANI, A. (Org.). **Escritos de educação**. 9. ed. Petrópolis: Vozes, 2007. p. 71-79.

BOURGUIGNON, F.; FERREIRA, F. H. G.; MENÉNDEZ, M. Inequality of opportunity in Brazil. **The Review of Income and Wealth**, New Haven, v. 53, n. 4, p. 585-618, 2007.

BRASIL. Ministério da Educação. **Microdados da Aneb e da Anresc 2017**. Brasília: Inep, 2018. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.

BRASIL. Ministério da Educação. **Relatório Brasil no Pisa 2018**: versão preliminar. Brasília: Inep/MEC, 2019. Disponível em: http://download.inep.gov.br/acoes_internacionais/pisa/documentos/2019/relatorio_PISA_2018_preliminar.pdf. Acesso em: 22 mar. 2020.

BUDDIN, R.; ZAMARRO, G. Teacher qualifications and student achievement in urban elementary schools. **Journal of Urban Economics**, v. 66, n. 2, p. 103-115, 2009.

BUJA, A.; HASTIE, T.; TIBSHIRANI, R. Linear smoothers and additive models. **The Annals of Statistics**, v. 17, n. 2, p. 453-510, jun. 1989.

BURKE, M. A.; SASS, T. R. Classroom peer effects and student achievement. **Journal of Labor Economics**, v. 31, n. 1, p. 51-82, 2013.

CANALES, A.; MALDONADO, L. Teacher quality and student achievement in Chile: linking teachers' contribution and observable characteristics. **International Journal of Educational Development**, v. 60, p. 33-50, 2018.

CARLSON, M. J.; CORCORAN, M. E. Family structure and children's behavioral and cognitive outcomes. **Journal of Marriage and Family**, v. 63, n. 3, p. 779-792, ago. 2001.

CARNEIRO, P.; HECKMAN, J. **Human capital policy**. Cambridge, Estados Unidos: NBER, fev. 2003. (Working Paper, n. 9495).

CARVALHO, M. M.; WALTENBERG, F. D. Desigualdade de oportunidades no acesso ao ensino superior no Brasil: uma comparação entre 2003 e 2013. **Economia Aplicada**, v. 19, n. 2, p. 369-396, 2015.

CITTADIN, I.; FRANÇA, M. T. A. A violência interna e externa como fator de influência no desempenho de alunos do primeiro ciclo do ensino fundamental. *In*: SACHSIDA, A. (Org.). **Políticas públicas**: avaliando mais de meio trilhão de reais em gastos públicos. Brasília: Ipea, 2018. p. 555-582.

COLEMAN, J. S. Social capital in the creation of human capital. **American Journal of Sociology**, v. 94, p. 95-120, 1988.

COLÓN-GONZÁLEZ, F. J. *et al.* The effects of weather and climate change on dengue. **PLOS Neglected Tropical Diseases**, v. 7, n. 11, p. 1-9, nov. 2013.

CURI, A. Z.; MENEZES-FILHO, N. A. A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil. **Estudos Econômicos**, São Paulo, v. 39, n. 4, p. 811-850, out.-dez. 2009.

DEE, T. S. A teacher like me: does race, ethnicity, or gender matter? **The American Economic Review**, v. 95, n. 2, p. 158-165, 2005.

DIAZ, M. D. M. (Des)igualdades de oportunidades no ensino médio brasileiro: escolas públicas e privadas. **Economia**, Brasília, v. 13, n. 3a, p. 553-568, set.-dez. 2012.

DIMAGGIO, P. Cultural capital and school success: the impact of status culture participation on the grade of U.S. high school students. **American Sociological Review**, v. 47, p. 189-201, 1982.

EGALITE, A. J.; KISIDA, B.; WINTERS, M. A. Representation in the classroom: the effect of own-race teachers on student achievement. **Economics of Education Review**, v. 45, p. 44-52, 2015.

FERREIRA, F. H. G. Os determinantes da desigualdade de renda no Brasil: luta de classes ou heterogeneidade educacional? *In*: HENRIQUES, R. (Org.). **Desigualdade e pobreza no Brasil**. Rio de Janeiro: Ipea, 2000. p. 131-158.

FERREIRA, F. H. G.; GIGNOUX, J. **The measurement of inequality of opportunity**: theory and an application to Latin America. Washington: The World Bank, jul. 2008. (Policy Research Working Paper, n. 4659).

FERREIRA, F. H. G.; GIGNOUX, J. **The measurement of educational inequality**: achievement and opportunity. Washington: The World Bank, nov. 2011. (Policy Research Working Paper, n. 5873).

FERREIRA, F. H. G.; GIGNOUX, J. The measurement of educational inequality: achievement and opportunity. **World Bank Economic Review**, v. 28, n. 2, p. 210-246, 2014.

FERREIRA, F. H. G.; GIGNOUX, J.; ARAN, M. Measuring inequality of opportunity with imperfect data: the case of Turkey. **Journal of Economic Inequality**, v. 9, n. 4, p. 651-680, 2011.

FERREIRA, F. H. G.; PERAGINE, V. **Equality of opportunity**: theory and evidence. Washington: World Bank Group, mar. 2015. (Policy Research Working Paper, n. 7217).

FIRPO, S.; JALES, H. B. **Measuring peer effects in the Brazilian school system**. São Paulo: FGV-EESP, Nov. 2011. (Working Paper, n. 17).

FLEURBAEY, M. **Fairness, responsibility, and welfare**. Oxford: Oxford University Press, 2008.

FLEURBAEY, M.; SCHOKKAERT, E. Unfair inequalities in health and health care. **Journal of Health Economics**, v. 28, n. 1, p. 73-90, 2009.

FOGUEL, M. N.; VELOSO, F. A. **Inequality of opportunity in daycare and preschool services in Brazil**. Ibre/FGV, abr. 2012. (Texto para Discussão).

GAMBOA, L. F.; WALTENBERG F. D. **Inequality of opportunity in educational achievement in Latin America: evidence from Pisa 2006-2009**. Palma de Mallorca: Ecineq, 2011. (Working Paper Series, n. 206).

GAMBOA, L. F.; WALTENBERG, F. D. Inequality of opportunity for educational achievement in Latin America: evidence from Pisa 2006-2009. **Economics of Education Review**, v. 31, n. 5, p. 694-708, 2012.

GROGGER, J. Local violence and educational attainment. **The Journal of Human Resources**, v. 32, n. 4, p. 659-682, 1997.

HANUSHEK, E. A. *et al.* **Does peer ability affect student achievement?** Cambridge, Estados Unidos: NBER, out. 2001. (Working Paper, n. 8502).

HANUSHEK, E. A. *et al.* **The market for teacher quality**. Cambridge, Estados Unidos: NBER, fev. 2005. (Working Paper, n. 11154).

HARDLE, W.; LINTON, O. Applied nonparametric methods. *In*: ENGLE, R. F.; McFADDEN, D. L. **Handbook of econometrics**. [s.l.]: Elsevier, 1994. p. 2295-2339.

HARRIS D. N.; SASS, T. R. **Teacher training, teacher quality and student achievement**. Arlington: Calder, mar. 2007. (Working Paper, n. 3).

HASTIE, T.; TIBSHIRANI, R. Generalized additive models. **Statistical Science**, v. 1, n. 3, p. 297-310, 1986.

HASTIE, T.; TIBSHIRANI, R. Generalized additive models: some applications. **Journal of the American Statistical Association**, v. 82, n. 398, p. 371-386, jun. 1987.

HASTIE, T.; TIBSHIRANI, R. (Ed.). **Generalized additive models**. Londres: Chapman and Hall/CRC, 1990.

HOFFMANN, R. Distribuição de renda e crescimento econômico. **Estudos Avançados**, v. 15, n. 41, p. 67-76, 2001.

HOROWITZ, J. L.; MAMMEN, E. Nonparametric estimation of an additive model with a link function. **The Annals of Statistics**, v. 32, n. 6, p. 2412-2443, 2004.

HOXBY, C. **Peer effects in the classroom: learning from gender and race variation**. Cambridge, Estados Unidos: NBER, ago. 2000. (Working Paper, n. 7867).

HUFE, P. *et al.* **Inequality of income acquisition**: the role of childhood circumstances. Mannheim: ZEW, dez. 2015. (Discussion Papers, n. 15-084).

HUNGER, M. *et al.* Multimorbidity and health-related quality of life in the older population: results from the German KORA-Age study. **Health and Quality of Life Outcomes**, v. 9, n. 53, p. 1-10, 2011.

INEP – INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Ideb**: resultados e metas. Brasília: Inep, 2020. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/ideb/resultados>. Acesso em: 15 set. 2020.

JAEGER, M. M. Does cultural capital really affect academic achievement? New evidence from combined sibling and panel data. **Sociology of Education**, v. 84, n. 4, p. 281-298, 2011.

JALES, H. B. **Peer effects na educação no Brasil**: evidência a partir dos dados do Saeb. 2010. 82 f. Dissertação (Mestrado) – Fundação Getulio Vargas, São Paulo, 2010.

KANG, C. Classroom peer effects and academic achievement: quasi-randomization evidence from South Korea. **Journal of Urban Economics**, v. 61, n. 3, p. 458-495, 2007.

KAUTZ, T. *et al.* **Fostering and measuring skills**: improving cognitive and non-cognitive skills to promote lifetime success. Cambridge, Estados Unidos: NBER, dez. 2014. (Working Paper, n. 20749).

LANGONI, C. G. **Distribuição da renda e desenvolvimento econômico do Brasil**: uma reafirmação. Rio de Janeiro: Ibre/FGV, 1973. (Ensaio Econômico da EPGE, n. 8).

LINTON, O.; NIELSEN, J. P. A kernel method of estimating structured non-parametric regression based on marginal integration. **Biometrika**, v. 82, n. 1, p. 93-100, 1995.

LOMBARDI FILHO, S. C.; OLIVEIRA, A. M. H. C. de. O impacto da violência nas escolas sobre a proficiência estudantil: uma análise para as capitais nordestinas. **Revista Econômica do Nordeste**, Fortaleza, v. 48, n. 3, p. 95-112, jul.-set. 2017.

McLANAHAN, S. Family structure and the reproduction of poverty. **American Journal of Sociology**, v. 90, n. 4, p. 873-901, jan. 1985.

MÉNDEZ, Y. S.; WALTENBERG, F. D. Desigualdade de renda e demanda por redistribuição caminham juntas na América Latina no período 1997-2015. **Dados – Revista de Ciências Sociais**, Rio de Janeiro, v. 61, n. 4, p. 341-384, 2018.

NOGUEIRA, L. C. B. **Ensaio sobre desigualdade de oportunidades educacionais e de renda**. 2015. 109 f. Tese (Doutorado) – Universidade Federal da Paraíba, João Pessoa, 2015.

PISTOLESI, N. Inequality of opportunity in the land of opportunities, 1968-2001. **The Journal of Economic Inequality**, v. 7, n. 4, p. 411-433, 2009.

PRATA, D. N.; RODRIGUES, W.; BERMEJO, P. H. Temperature significantly changes covid-19 transmission in (sub)tropical cities of Brazil. **Science of the Total Environment**, v. 729, p. 1-7, 2020.

PROCÓPIO, I. V.; FREGUGLIA, R. da S.; CHEIN, F. Desigualdade de oportunidades na formação de habilidades: uma análise com dados longitudinais. **Economia Aplicada**, v. 19, n. 2, p. 326-348, 2015.

ROEMER, J. E. A pragmatic theory of responsibility for the egalitarian planner. **Philosophy & Public Affairs**, v. 22, n. 2, p. 146-166, 1993.

ROEMER, J. E. (Ed.). **Equality of opportunity**. Nova York: Harvard University Press, 1998.

SALEHI-ISFAHANI, D.; HASSINE, N. B.; ASSAAD, R. Equality of opportunity in educational achievement in the Middle East and North Africa. **The Journal of Economic Inequality**, v. 12, n. 4, p. 489-515, 2014.

SILVEIRA, L. F. de V. **A avaliação do impacto de um treinamento utilizando propensity score matching**: uma abordagem não-paramétrica e semiparamétrica. 2015. 71 f. Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2015.

SOUZA, W. P. S. de F.; OLIVEIRA, V. R.; ANNEGUES, A. C. Desigualdade de oportunidades na educação fundamental brasileira: novas evidências a partir de um índice para o Saeb. **Ensaio FEE**, Porto Alegre, v. 38, n. 2, p. 329-356, set. 2017.

SOUZA, W. P. S. de F.; OLIVEIRA, V. R.; ANNEGUES, A. C. *Background* familiar e desempenho escolar: uma abordagem não paramétrica. **Pesquisa e Planejamento Econômico**, v. 48, n. 2, p. 133-162, ago. 2018.

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

TEIXEIRA, E. C.; KASSOUF, A. L. Impacto da violência nas escolas paulistas sobre o desempenho acadêmico dos alunos. **Economia Aplicada**, v. 19, n. 2, p. 221-240, 2015.

TRAMONTE, L.; WILLMS, J. D. Cultural capital and its effects on education outcomes. **Economics of Education Review**, v. 29, n. 2, p. 200-213, 2010.

UNDP – UNITED NATIONS DEVELOPMENT PROGRAMME. **Human Development Report 2019: beyond income, beyond averages, beyond today – inequalities in human development in the 21st century.** Nova York: UNDP, 2019. Disponível em: https://reliefweb.int/report/world/human-development-report-2019-beyond-income-beyond-averages-beyond-today-inequalities?gclid=Cj0KCQjwj5mpBhDJARIsAOVjBdo5NKXHAIHTj_yk9hY-2p8f3vfHcxccDPmh4QwwO-8qP20FzHjI7iuUaAiNiEALw_wcB. Acesso em: 21 mar. 2020.

VASCONCELOS, A. M.; RIBEIRO, F. G.; FERNANDEZ, R. N. O efeito da estrutura familiar na educação dos filhos. **Análise Econômica**, Porto Alegre, v. 35, n. especial, p. 289-315, jul. 2017.

VENABLES, W. N.; RIPLEY, B. D. (Ed.). **Modern applied statistics with S.** 4. ed. Nova York: Springer, 2002. 504 p.

WALTENBERG, F. D.; VANDENBERGHE, V. What does it take to achieve equality of opportunity in education? An empirical investigation based on Brazilian data. **Economics of Education Review**, v. 26, n. 6, p. 709-723, 2007.

BIBLIOGRAFIA COMPLEMENTAR

CHECCHI, D.; PERAGINE, V. **Regional disparities and inequality of opportunity: the case of Italy.** Bonn: IZA, 2005. (Discussion Papers, n. 1874).

COLEMAN, J. S. *et al.* (Ed.). **Equality of educational opportunity.** Washington: U.S. GPO, 1966.

FERREIRA, F. H. G. *et al.* Ascensão e queda da desigualdade de renda no Brasil: uma atualização para 2005. *In*: BARROS, R. P.; FOGUEL, M. N.; ULYSSEA, G. (Org.). **Desigualdade de renda no Brasil: uma análise da queda recente.** Brasília: Ipea, 2006. v. 1, p. 359-378.

FIGUEIREDO, E. A.; SILVA, C. R. da F.; REGO, H. de O. Desigualdade de oportunidades no Brasil: efeitos diretos e indiretos. **Economia Aplicada**, v. 16, n. 2, p. 237-254, 2012.

LINTON, O. B.; HÄRDLE, W. Estimating additive regression models with known links. **Biometrika**, v. 83, n. 3, p. 529-540, 1996.

PERAGINE, V.; SERLENGA, L. **Higher education and equality of opportunity in Italy.** Bonn: IZA, nov. 2007. (Discussion Paper, n. 3163).

RAWLS, J. **Uma teoria da justiça.** São Paulo: Martins Fontes, 2000.

SAMPAIO, C. E. M. **Situação educacional dos jovens brasileiros na faixa etária de 15 a 17 anos.** Brasília: Inep, 2009. (Textos para Discussão, n. 33).

SILVA FILHO, G. A.; CARVALHO, M. R. V. **O efeito da formação inicial do professor sobre o desempenho escolar em matemática nos anos iniciais do ensino fundamental**. Brasília: Inep/MEC, 2017. (Texto para Discussão, n. 43).

SOARES, F. V. *et al.* **Programas de transferência de renda no Brasil: impactos sobre a desigualdade**. Brasília: Ipea, out. 2006. (Texto para Discussão, n. 1228).

SOARES, S. S. D. Análise de bem-estar e decomposição por fatores da queda na desigualdade entre 1995 e 2004. **Econômica**, v. 8, n. 1, p. 83-115, 2006.

APÊNDICE A

QUADRO A.1
Descrição das variáveis

	Variáveis	Descrição
Variáveis dependentes	Proficiência em língua portuguesa e matemática	Proficiência em língua portuguesa e matemática em escala única do Sistema Nacional de Avaliação da Educação Básica (Saeb) para o ano 2017.
Conjuntos de circunstâncias		
Base	Sexo	<i>Dummy</i> que assume valor um para estudantes do sexo masculino, e zero caso contrário.
	Raça	<i>Dummy</i> que assume valor um para estudantes que se autodeclararam branco, e zero caso contrário.
	Escolaridade da mãe	<i>Dummy</i> que assume valor um para mães com ensino superior completo, e zero caso contrário.
	Escolaridade do pai	<i>Dummy</i> que assume valor um para pais com ensino superior completo, e zero caso contrário.
	Índice socioeconômico	O índice foi obtido por meio da técnica de análise de componentes principais para variáveis categóricas. Foram utilizadas as respostas a sete perguntas: se no domicílio tem televisão, geladeira, máquina de lavar roupa, carro, computador, dormitório, e se trabalha empregado(a) doméstico(a) em casa.
	Frequentou creche e/ou pré-escola	<i>Dummy</i> que assume valor um se o estudante frequentou creche e/ou pré-escola, e zero caso contrário.
	Tamanho da família	Variável que indica o número de pessoas no domicílio.
	Já foi reprovado	<i>Dummy</i> que assume valor um se o estudante já foi reprovado, e zero caso contrário.
	Já abandonou a escola	<i>Dummy</i> que assume valor um se o estudante já abandonou a escola, e zero caso contrário.
	Escola pública ou privada	<i>Dummy</i> que assume valor um se o estudante estuda em escola pública, e zero caso contrário.
	Escola urbana ou rural	<i>Dummy</i> que assume valor um se o estudante estuda em escola urbana, e zero caso contrário.
	Trabalho fora de casa	<i>Dummy</i> que assume valor um se o estudante trabalha fora de casa, e zero caso contrário.
Região	Cinco <i>dummies</i> para as regiões Norte, Nordeste, Sudeste, Sul e Centro-Oeste.	

(Continua)

(Continuação)

	Variáveis	Descrição
Relação entre pais e filhos	Mora com mãe e pai	<i>Dummy</i> que assume valor um se o estudante mora com mãe e pai, e zero caso contrário.
	Vê mãe lendo	<i>Dummy</i> que assume valor um se o estudante vê mãe lendo, e zero caso contrário.
	Vê pai lendo	<i>Dummy</i> que assume valor um se o estudante vê pai lendo, e zero caso contrário.
	Frequência às reuniões de pais	<i>Dummy</i> que assume valor um se pais frequentam as reuniões de pais, e zero caso contrário.
	Pais incentivam a estudar	<i>Dummy</i> que assume valor um se pais incentivam a estudar, e zero caso contrário.
	Pais incentivam a fazer dever de casa	<i>Dummy</i> que assume valor um se pais incentivam a fazer dever de casa, e zero caso contrário.
	Pais incentivam a ler	<i>Dummy</i> que assume valor um se pais incentivam a ler, e zero caso contrário.
	Pais incentivam a ir à escola	<i>Dummy</i> que assume valor um se pais incentivam a ir à escola, e zero caso contrário.
	Pais conversam sobre o que acontece na escola	<i>Dummy</i> que assume valor um se pais conversam sobre o que acontece na escola, e zero caso contrário.
Hábitos culturais e uso do tempo	Lê jornais	<i>Dummy</i> que assume valor um se o estudante lê jornais, e zero caso contrário.
	Lê livros	<i>Dummy</i> que assume valor um se o estudante lê livros, e zero caso contrário.
	Lê revistas em geral	<i>Dummy</i> que assume valor um se o estudante lê revistas em geral, e zero caso contrário.
	Lê revistas em quadrinhos	<i>Dummy</i> que assume valor um se o estudante lê revistas em quadrinhos, e zero caso contrário.
	Lê sites na internet	<i>Dummy</i> que assume valor um se o estudante lê sites na internet, e zero caso contrário.
	Costuma ir à biblioteca	<i>Dummy</i> que assume valor um se o estudante costuma frequentar bibliotecas, e zero caso contrário.
	Costuma ir ao cinema	<i>Dummy</i> que assume valor um se o estudante costuma ir ao cinema, e zero caso contrário.
	Costuma ir a espetáculos	<i>Dummy</i> que assume valor um se o estudante costuma ir a espetáculos (museu, teatro, dança ou música), e zero caso contrário.
	Tempo gasto em dias de aula assistindo TV, navegando na internet ou jogando jogos eletrônicos	Variável contínua que indica o tempo gasto em dias de aula assistindo TV, navegando na internet ou jogando jogos eletrônicos.
	Tempo gasto em dias de aula fazendo trabalhos domésticos	Variável contínua que indica o tempo gasto em dias de aula fazendo trabalhos domésticos.

(Continua)

(Continuação)

	Variáveis	Descrição
Características do professor	Sexo	Percentual de professores homens na escola.
	Raça	Percentual de professores brancos na escola.
	Se tem curso superior	Percentual de professores que têm curso superior.
	Experiência	Percentual de professores com dez anos ou mais de experiência.
	Formação continuada	Percentual de professores que participaram de alguma atividade de desenvolvimento profissional nos últimos dois anos.
	Pós-graduação	Percentual de professores que têm pós-graduação.
Características do diretor	Sexo	<i>Dummy</i> que assume valor um para diretores do sexo masculino, e zero caso contrário.
	Raça	<i>Dummy</i> que assume valor um para diretores que se autodeclararam brancos, e zero caso contrário.
	Se tem curso superior	<i>Dummy</i> que assume valor um se diretor tem curso superior (pedagogia, escola normal superior, outras licenciaturas, outros), e zero caso contrário.
	Experiência	<i>Dummy</i> que assume valor um para diretores com mais de cinco anos experiência, e zero caso contrário.
	Formação continuada	<i>Dummy</i> que assume valor um se diretor participou de alguma atividade de desenvolvimento profissional nos últimos dois anos, e zero caso contrário.
	Pós-graduação	<i>Dummy</i> que assume valor um se diretor tem pós-graduação, e zero caso contrário.
Características da escola	Qualidade da escola	Índice obtido por meio da técnica de análise de componentes principais para variáveis categóricas. Foram utilizadas respostas a dez perguntas acerca da infraestrutura escolar: salas de aula, banheiros, cozinha, instalações hidráulicas, instalações elétricas, biblioteca, sala de leitura, quadra de esportes, laboratório de informática e laboratório de ciências.
	Violência	Índice obtido por meio da técnica de análise de componentes principais para variáveis categóricas. Foram utilizadas respostas dos diretores sobre os seguintes acontecimentos na escola: agressão verbal ou física, atentado à vida, ameaça, furto, roubo, se alunos frequentaram a escola sob efeito de bebida alcoólica, drogas ilícitas, portando arma branca, portando arma de fogo.
Características da turma	Homogênea quanto à idade	<i>Dummy</i> que assume valor um se a turma é homogênea quanto à idade, e zero caso contrário.
	Homogênea quanto ao rendimento escolar	<i>Dummy</i> que assume valor um se a turma é homogênea quanto ao rendimento escolar, e zero caso contrário.
	Heterogênea quanto à idade	<i>Dummy</i> que assume valor um se a turma é heterogênea quanto à idade, e zero caso contrário.
	Heterogênea quanto ao rendimento escolar	<i>Dummy</i> que assume valor um se a turma é heterogênea quanto ao rendimento escolar, e zero caso contrário.

Fonte: Microdados do Saeb 2017. Disponível em: <https://www.gov.br/inep/pt-br/acao-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.

Elaboração dos autores.

TABELA A. 1
Peso de cada circunstância na desigualdade de oportunidades do conjunto base

Variáveis de circunstâncias	Português	Matemática
Índice de desigualdade de oportunidades (IOP)	0,2420	0,2460
Aluno homem (%)	4,65	1,20
Aluno branco (%)	3,58	4,58
Educação dos pais e índice socioeconômico (%)	15,90	19,67
Frequentou creche/pré-escola (%)	1,85	2,79
Tamanho da família (%)	3,42	2,58
Já foi reprovado (%)	28,88	25,96
Já abandonou a escola (%)	3,15	2,64
Escola pública (%)	11,79	13,97
Escola urbana (%)	4,17	3,70
Trabalho fora de casa (%)	12,37	7,43
Região (%)	10,23	15,47

Fonte: Microdados do Saeb 2017. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 16 ago. 2019.
 Elaboração dos autores.

Originais submetidos em: out. 2020.

Última versão recebida em: dez. 2021.

Aprovada em: dez. 2021.

NOTA AOS COLABORADORES DE PESQUISA E PLANEJAMENTO ECONÔMICO

1. A revista só analisa, com vistas a eventual publicação, artigos com conteúdo inédito, tanto no país quanto no exterior. Além disso, o seu tema deve se inserir em uma das áreas da ciência econômica, contribuindo de modo significativo ao avanço do conhecimento científico nessa área.

2. Resenhas de livros recentemente publicados poderão ser consideradas para publicação, mas resenhas temáticas e os textos essencialmente descritivos não serão, de um modo geral, aceitos.

3. As contribuições não serão remuneradas, e a submissão de um artigo à revista implica a transferência dos direitos autorais ao Ipea, caso ele venha a ser publicado.

4. Em geral, os artigos submetidos à revista devem ser escritos em português. Em casos excepcionais, poderão ser recebidos textos em língua inglesa para análise, mas, se ele vier a ser aceito para publicação, o autor deverá se responsabilizar por sua tradução.

5. Só serão publicados artigos em português, mas sua versão em inglês poderá ser disponibilizada no sítio da revista na internet. Os anexos muito longos ou complexos para serem publicados, bem como as bases de dados necessárias para reproduzir os resultados empíricos do trabalho, serão também oferecidos aos leitores em versão virtual.

6. Caso o trabalho seja aceito para publicação, cada autor receberá 3 (três) exemplares do número da revista correspondente.

7. Para submeter um trabalho à revista, o autor deve acessar a página de Pesquisa e Planejamento Econômico na internet, em <http://ppe.ipea.gov.br/index.php/pppe/index>, e seguir os procedimentos ali descritos.

8. Os artigos recebidos que estejam de acordo com as instruções acima serão avaliados pelo Corpo Editorial com o auxílio de dois pareceristas escolhidos pelo Editor. O trabalho dos pareceristas é feito observando-se o método duplamente cego: o autor não saberá quem são os pareceristas, nem estes quem é o autor. Dessa análise poderá resultar a aceitação do artigo, condicionada, ou não, à realização de alterações; sua rejeição, com ou sem a recomendação de nova submissão após modificações; ou a rejeição definitiva. No caso de uma segunda submissão, o artigo será novamente avaliado por pareceristas, podendo vir a ser enquadrado em qualquer das situações acima. A rotina de análise se repete até que uma decisão final de rejeição ou aceitação seja alcançada. O processamento do artigo é conduzido pelo Editor, a quem cabe também a comunicação com os autores.

9. A decisão final quanto à publicação dos artigos cabe ao Corpo Editorial, que se reúne ordinariamente para decidir a composição de cada um dos números da revista, por recomendação do Editor. A aprovação do artigo para publicação só então é comunicada aos autores dos artigos respectivos, por escrito.

Ipea – Instituto de Pesquisa Econômica Aplicada

EDITORIAL

Coordenação

Aeromilson Trajano de Mesquita

Assistentes da Coordenação

Rafael Augusto Ferreira Cardoso

Samuel Elias de Souza

Supervisão

Ana Clara Escórcio Xavier

Everson da Silva Moura

Revisão

Alice Souza Lopes

Amanda Ramos Marques Honorio

Barbara de Castro

Brena Rolim Peixoto da Silva

Cayo César Freire Feliciano

Cláudio Passos de Oliveira

Clícia Silveira Rodrigues

Olavo Mesquita de Carvalho

Regina Marta de Aguiar

Reginaldo da Silva Domingos

Jennyfer Alves de Carvalho (estagiária)

Katarinne Fabrizzi Maciel do Couto (estagiária)

Editoração

Anderson Silva Reis

Augusto Lopes dos Santos Borges

Cristiano Ferreira de Araújo

Daniel Alves Tavares

Danielle de Oliveira Ayres

Leonardo Hideki Higa

Natália de Oliveira Ayres

The manuscripts in languages other than Portuguese published herein have not been proofread.

Ipea – Brasília

Setor de Edifícios Públicos Sul 702/902, Bloco C

Centro Empresarial Brasília 50, Torre B

CEP: 70390-025, Asa Sul, Brasília-DF

Missão do Ipea

Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.

ISSN 0100-0551



9 770100 055040

Apoio editorial



ipea Instituto de Pesquisa
Econômica Aplicada

MINISTÉRIO DO
PLANEJAMENTO
E ORÇAMENTO

GOVERNO FEDERAL



UNIÃO E RECONSTRUÇÃO