

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 66^a posição em ciências, na 57^a em leitura e na 70^a 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

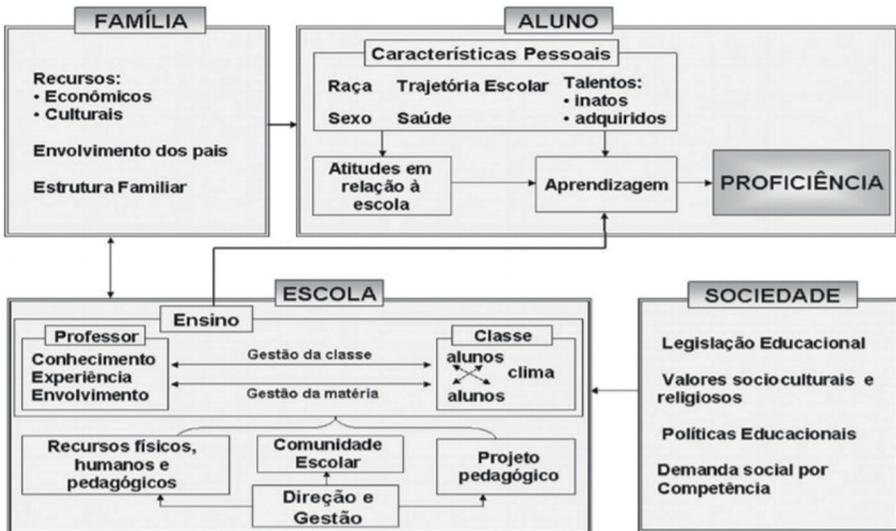
comparados a sistemas de remuneração coletiva, porém, em um contexto de trabalho em equipe, podem gerar sensação de injustiça e competitividade entre docentes, e, no limite, desmotivar professores a trabalharem em equipe ou prejudicar o ambiente de trabalho, prejudicando também o aluno. Em contrapartida, 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 da gestão e cópias de documentos comprobatórios especificados no edital, além de posteriores vistórias às escolas.	Não houve necessidade de realização de inscrição, apenas do 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.	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.	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 preditor 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

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

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/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.

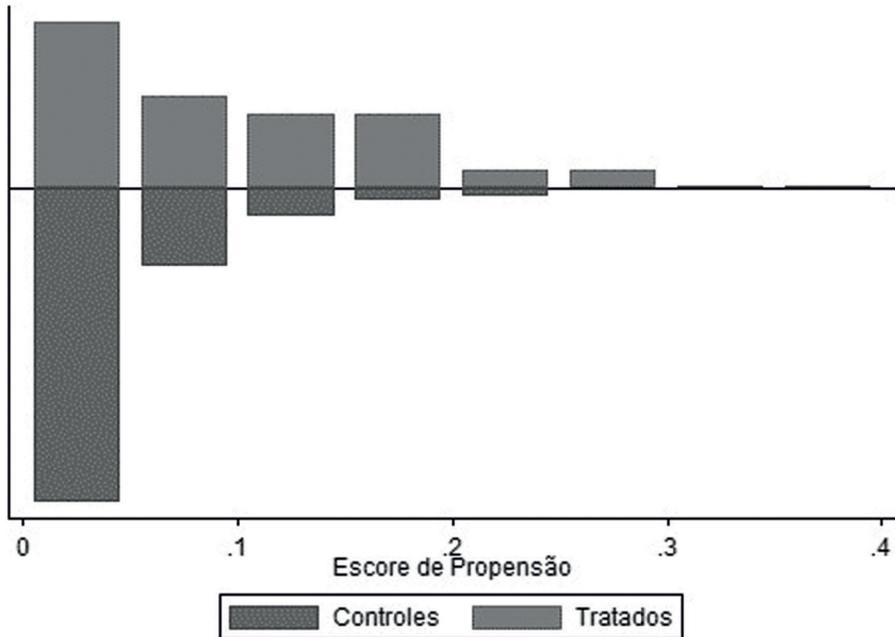
Obs.: *** Estatisticamente significante a 1%.

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.