

INDICADORES EDUCACIONAIS E RESPONSABILIZAÇÃO ESCOLAR: UM ESTUDO DO “PRÊMIO ESCOLA NOTA DEZ”

MARIANE CAMPELO KOSLINSKI

EDUARDO RIBEIRO

LUIXA XAVIER DE OLIVEIRA

RESUMO

O artigo focaliza o impacto do uso do Indicador de Desempenho Escolar – Alfabetização (IDE-Alfa) pela política “Prêmio Escola Nota Dez” do estado do Ceará, empregado para estabelecer pontos de corte para premiação e apoio financeiro para escolas públicas. Utilizam-se os métodos de Propensity Score Matching e regressões múltiplas para observar o efeito da política sobre desempenho, equidade e taxa de participação dos alunos na avaliação externa da alfabetização das escolas premiadas e apoiadas de 2009 a 2011, com base nos resultados do Índice de Desempenho Escolar no 2º ano do Ensino Fundamental (IDE-Alfa) de 2008 a 2010. Os resultados das análises sugerem um efeito positivo no desempenho, equidade e taxa de participação das escolas premiadas e no desempenho das escolas com características similares às premiadas. Para as escolas apoiadas, foram observados efeitos significativos positivos somente para a taxa de participação na avaliação externa do estado do Ceará.

PALAVRAS-CHAVE INDICADORES EDUCACIONAIS • RESPONSABILIZAÇÃO • POLÍTICAS EDUCACIONAIS • BONIFICAÇÃO SALARIAL .

INDICADORES EDUCACIONALES Y RESPONSABILIZACIÓN ESCOLAR: UN ESTUDIO DEL “PRÊMIO ESCOLA NOTA DEZ”

RESUMEN

El artículo enfoca el impacto del uso del Indicador de Desempeño Escolar – Alfabetización (IDE-Alfa) por la política “Prêmio Escola Nota Dez” del estado de Ceará, utilizado para establecer puntos de corte para premiación y apoyo financiero a escuelas públicas. Se emplean los métodos de Propensity Score Matching y regresiones múltiples para observar el efecto de la política sobre desempeño, equidad e índice de participación de los alumnos en la evaluación externa de la alfabetización de las escuelas premiadas y apoyadas del 2009 al 2011, en base a los resultados del Índice de Desempeño Escolar en el 2º año de Educación Fundamental (IDE-Alfa) del 2008 al 2010. Los resultados de los análisis sugieren un efecto positivo en el desempeño, equidad e índice de participación de las escuelas premiadas y en el desempeño de las escuelas con características similares a las premiadas. Para las escuelas apoyadas, se observaron efectos significativos positivos tan solo para el índice de participación en la evaluación externa del estado de Ceará.

PALABRAS CLAVE INDICADORES EDUCACIONALES • RESPONSABILIZACIÓN • POLÍTICAS EDUCACIONALES • BONIFICACIÓN SALARIAL.

EDUCATIONAL INDICATORS AND SCHOOL ACCOUNTABILITY: A STUDY ABOUT THE “PRÊMIO ESCOLA NOTA DEZ”

ABSTRACT

The paper focuses on the impact of the use of the School Performance Index – Literacy (IDE-Alfa) by the “Prêmio Escola Nota Dez”, a policy of the State of Ceará, that establishes cut-off points to award and to allocate financial support to public schools. It uses the Propensity Score Matching and multiple regressions methods to observe the effects of the reward/financial support on school performance, equity, and student participation rate on the external assessment of schools awarded and supported from 2009 to 2011, based on schools’ Index of school Performance results (IDE-Alfa) from 2008 to 2010. The results suggest a positive effect on the performance, equity and participation rate of the awarded schools and a smaller, but also positive, effect on performance of the schools with similar characteristics to the awarded ones. The results show a statistically significant impact only on the participation rate of the schools that received financial support.

KEYWORDS EDUCATIONAL INDICATORS • ACCOUNTABILITY • EDUCATIONAL POLICIES • SALARY BONUSES.

INTRODUÇÃO

Nas últimas décadas, observam-se a implementação e a ampliação dos sistemas de informação e de avaliação educacional no contexto brasileiro. Um número crescente de estados adota seus próprios sistemas de avaliação e, especialmente após a elaboração do Índice de Desenvolvimento da Educação Básica (Ideb), verifica-se a multiplicação de indicadores educacionais estaduais utilizados para diversos fins, tais como monitoramento da evolução geral dos sistemas educacionais, alocação de recursos para escolas e/ou municípios, estabelecimento de metas para escolas, instâncias intermediárias de gestão e redes acompanhadas de baixas e/ou altas consequências (responsabilização escolar), entre outros (BROOKE; CUNHA, 2011; GIMENES et al., 2013).

No contexto brasileiro, apesar do avanço dos indicadores educacionais que, por exemplo, passaram a considerar em suas formulações dimensões como fluxo e equidade (BOOKE; CUNHA, 2011; SOARES, 2009), a discussão sobre os cálculos utilizados para estabelecer metas ainda é incipiente. Por exemplo, grande parte das discussões sobre o uso de

indicadores educacionais para políticas de responsabilização escolar no campo da educação, em especial de bonificação de professores, está mais fundada em posições ideológicas e/ou evidências de outros países do que em indícios do impacto de tais políticas no contexto brasileiro (BROOKE, 2013). Somente muito recentemente observamos a proliferação de estudos que podem subsidiar as discussões no âmbito brasileiro a partir de evidências empíricas geradas por trabalhos com desenhos mais robustos e adequados para realizar inferências causais (ALVES et al., 2016; BRESOLIN, 2014; CARNEIRO; IRFFI, 2014; FERRAZ; BURNS, 2012; OHSIRO; SCORZAFAZE; DORIGAN, 2015; SCORZAFAZE; DORIGAN; FERREIRA, 2015).

O presente trabalho pretende contribuir para essa discussão com uma análise exploratória do uso do Indicador de Desempenho Escolar – Alfabetização (IDE-Alfa), do estado do Ceará, como medida para definir pontos de corte para a alocação de recursos, premiando as escolas de alto desempenho e apoiando aquelas de baixo desempenho, por meio do programa “Prêmio Escola Nota Dez”. Mais especificamente, pretende-se observar o efeito marginal dos três primeiros anos do programa sobre o desempenho, as desigualdades da proficiência e o fluxo escolar, bem como sobre a taxa de participação dos alunos nas avaliações externas.

O artigo discute, em sua primeira parte, características dos indicadores educacionais utilizados no contexto brasileiro em políticas de bonificação salarial e alocação de recursos para escolas de baixo desempenho, trazendo ainda evidências observadas por estudos a respeito do impacto de tais usos sobre o desempenho das escolas. Posteriormente, discutem-se as características dos indicadores de qualidade do estado do Ceará, utilizados pelo “Programa Escolas Nota Dez” para alocação de recursos para escolas públicas que oferecem os anos iniciais do ensino fundamental. Por fim, em sua última etapa, o artigo descreve a estratégia metodológica utilizada para analisar o impacto do “Prêmio Escola Nota Dez” (empregando pareamento por escores de propensão) e apresenta os principais resultados encontrados.

USO DE INDICADORES EDUCACIONAIS PARA RESPONSABILIZAÇÃO ESCOLAR: DESENHOS E IMPACTOS

Ao discutirem o uso dos sistemas de avaliação educacional por estados brasileiros, Brooke e Cunha (2011) fazem uma distinção entre “alocação de recursos” e “incentivo salarial”, entre outros usos observados pelo estudo. Segundo os autores, a utilização para alocação de recursos compreende programas ou políticas que destinam ações ou recursos para regiões, municípios e/ou escolas, empregando critérios baseados nos resultados das avaliações externas. No contexto dos estados brasileiros, Brooke e Cunha (2011) observaram desde políticas compensatórias, com foco em municípios e/ou escolas de baixo desempenho, até a alocação de recursos para regiões, municípios e/ou escolas com melhores resultados. Já o uso para incentivo salarial é caracterizado pela alocação de bônus ou prêmios para equipe escolar com base no desempenho dos alunos em avaliações externas. Essa utilização seria uma forma de responsabilização escolar, uma vez que envolve o estabelecimento de metas e consequências para as escolas. Os autores verificaram ampla disseminação nos estados brasileiros do uso de seus próprios sistemas de avaliação e indicadores educacionais para estabelecer metas e distribuir bônus e/ou prêmios para equipe escolar das escolas que atingem as metas estabelecidas.

Brooke e Cunha (2011) observaram que, em alguns casos, o uso para alocação de recursos é seguido pelo estabelecimento de metas e algum grau de responsabilização das escolas. No entanto, políticas de distribuição de recursos para escolas com baixo desempenho não seriam consideradas políticas de responsabilização escolar, uma vez que não pretendem estimular a redução do desempenho das escolas para que recebam o benefício/recurso. Já as políticas de bonificação salarial ou de alocação de recursos para escolas de alto desempenho levariam à competição pelo prêmio, representando um incentivo para a melhora de todas as escolas (BROOKE, 2013). Para Alves *et al.* (2016), políticas de alocação de recursos para escolas de baixo desempenho, tais como o Plano de Desenvolvimento da Escola (PDE-Escola),¹ podem ser acompanhadas por sistemas “suaves” de *accountability*²

1 O Plano de Desenvolvimento da Escola (PDE-Escola) é um programa do governo federal, instituído pelo Plano de Desenvolvimento da Educação (2007), que promove o apoio financeiro e técnico para escolas com baixo desempenho no Ideb. A intenção do programa é de que as escolas elaborem plano de melhoramento e consigam aumentar seu desempenho. Por exemplo, entre 2007 e 2009, o programa priorizou escolas com Ideb inferior a 3,0 para os anos iniciais e até 2,8 para os anos finais. O montante de recursos recebidos pelas escolas variava de acordo com o número de alunos matriculados.

2 Diversos estudos dividem as políticas de responsabilização escolar a partir da seguinte tipologia: responsabilização forte ou com alta consequência (*high stakes*) que impõe consequências materiais; e responsabilização branda ou com baixa consequência (*low stakes*) que abarca apenas consequências simbólicas. Bonamino e Sousa (2012) e Brooke (2006) descrevem diversos exemplos desses diferentes tipos de responsabilização escolar implementados no contexto brasileiro.

ou de baixas consequências. Por exemplo, o apoio financeiro, acompanhado por medidas de divulgação de resultados das escolas com baixo desempenho que, por sua vez, podem gerar tanto pressão pela comunidade escolar (por melhores resultados), como pressões decorrentes do estigma de “escola de baixa qualidade”, e, ainda, conduzir exigências de que tais escolas elaborem (e cumpram) um plano de melhoramentos – acompanhado por assistência técnica e financeira.

O uso de indicadores educacionais para responsabilização escolar, bonificação salarial e/ou alocação de recursos para escolas de baixo desempenho seria uma forma de estimular certas práticas escolares e, conseqüentemente, gerar maior eficácia das escolas e qualidade dos sistemas de educação. Por exemplo, a alocação de recursos para escolas de baixo desempenho poderia, potencialmente, trazer maior pressão para melhorarem seus resultados e maior apoio técnico e recursos financeiros. Por outro lado, tal uso geraria, eventualmente, efeitos negativos, como estigmatização das instituições e perda de motivação dos atores escolares (ALVES et al., 2016).

Já a utilização dos resultados do desempenho dos alunos em sistemas de avaliação externa para distribuição de bônus salarial, a partir da mensuração e divulgação do desempenho das escolas e adoção de consequências positivas/negativas para aquelas que alcançam/não alcançam as metas estabelecidas, encorajaria professores a mudar em suas práticas ou métodos de ensino (FLIGLIO; LOEB, 2011). Ou seja, iniciativas de remuneração salarial por mérito alinhariam interesses dos professores e dos alunos e poderiam gerar incentivos para o reconhecimento de boas práticas docentes e de gestão e, dessa forma, contribuiriam para o aumento da aprendizagem dos alunos (PONTUAL, 2008).

No entanto, para Ladd (2001), sistemas de responsabilização seriam instrumentos poderosos para mudar comportamento de atores escolares tanto na direção intencionada quanto na não intencionada. Por exemplo, alguns estudos enumeram possíveis efeitos não intencionados de tais políticas, como incentivar práticas de “gaming”³ – por exemplo, seleção ou exclusão de alunos, direcionamento seletivo

3 Estudos sobre políticas de *accountability* na área de educação observaram a proliferação de estratégias adotadas pelas escolas descritas como “*gaming the system*”. Essas práticas remetem ao aprendizado de formas de trapaça ou modos de jogar com as regras e com o sistema, para atingir as metas estabelecidas, sem necessariamente buscar em sua integralidade o objetivo traçado e pretendido por planejadores de políticas e intervenções educacionais.

do planejamento pedagógico da escola com foco em parte dos alunos, limitação do currículo –, o que muitas vezes se traduz em aumento das desigualdades intraescolares (DIAMOND; SPILLANE, 2004; HEILIG; DARLING-HAMMOND, 2008; HSIEH; URQUIOLA, 2006; WEST; PETERSON, 2006). Portanto, Ladd (2001) argumenta que sistemas de responsabilização escolar precisariam ser desenhados de forma cuidadosa para incentivar os comportamentos e alcançar os objetivos desejados.

No que diz respeito ao uso de indicadores educacionais para incentivos salariais, existe uma ampla discussão no contexto internacional sobre as medidas utilizadas para estabelecer metas e, conseqüentemente, responsabilizar ou bonificar professores e diretores. Essa discussão torna-se importante uma vez que as medidas e os indicadores trazem diferentes sinais, sendo percebidos como mais ou menos injustos por diversos atores e, por conseguinte, acarretam diferentes reações e adoção de práticas e estratégias de naturezas distintas por parte das escolas (FIGLIO; LOEB, 2011; GORARD; HORDOSY; SIDDIQUI, 2013; LADD, 2001; LIU; CHARLEY; FULLER, 2013; TIMMERMANS; DOOLARD; WOLF, 2011).

Ladd (2001) discute as vantagens e desvantagens e os possíveis comportamentos incentivados por sistemas de responsabilização escolar que utilizam os seguintes tipos de indicadores: medidas baseadas em médias de desempenho ou *status* – porcentagem de alunos em nível adequado de proficiência; medidas baseadas em taxas de aumento do desempenho; e medidas de valor agregado. Para a autora, as duas primeiras métricas refletem mais o perfil dos alunos do que a prática dos professores e são mais passíveis de serem vistas como injustas. Além disso, as medidas baseadas em taxa de aumento frequentemente trazem metas que refletem muito mais o interesse e as agendas dos formuladores do que possibilidades realistas de as escolas aumentarem seus indicadores, sobretudo por não considerarem o perfil do alunado e o contexto no qual as escolas estão inseridas. Segundo Ladd (2001), as medidas de valor agregado que medem ganhos de aprendizagem baseados em *performance* prévia dos alunos seriam mais adequadas, em especial aque-

las de valor agregado contextualizadas (*contextualized value added* – CVA), que controlam a medida por diversas características dos alunos e das escolas (TIMMERMANS; DOOLARD; WOLF, 2011). Entretanto, mesmo essas medidas seriam alvo de críticas, principalmente quando utilizadas em políticas de responsabilização de alta consequência e com punição.⁴ Pesa também contra a aceitação das medidas de valor agregado o fato de que estas dependem de controles estatísticos nem sempre tão facilmente interpretáveis por gestores, diretores e professores.

A seguir, esta seção discute os diferentes indicadores utilizados no contexto brasileiro, em especial pelos estados e pelo governo federal, para a alocação de recursos para as escolas e/ou para a bonificação de professores⁵ ou incentivos salariais. Também são abordadas as recentes evidências sobre o impacto de tais usos sobre desempenho e desigualdades de proficiências dentro e entre escolas.

Os indicadores de qualidade da educação elaborados por estados brasileiros não utilizam medidas de valor agregado porque, em sua maioria, os sistemas de avaliação empregados para elaboração de indicadores de qualidade não possuem desenho longitudinal. No entanto, mesmo provenientes de avaliações transversais,⁶ observamos recentes mudanças nos indicadores educacionais brasileiros considerados para políticas de responsabilização escolar e alocação de recursos, em especial após a implementação do Ideb. Se uma primeira geração de políticas de bonificação salarial, desenvolvidas a partir de 2001, utilizava a média do desempenho escolar dos alunos no ano anterior como critério para o pagamento coletivo de incentivos aos profissionais das escolas,⁷ uma segunda geração, na tentativa de superar as críticas à primeira, foi fortemente influenciada pela elaboração do Ideb, que combina o desempenho médio dos estudantes em matemática e leitura com informações sobre o fluxo escolar, como taxas de aprovação (BROOKE; CUNHA, 2011).

Mais recentemente, os indicadores estaduais passaram a incluir não somente dimensões de fluxo, como também de equidade, taxa de participação dos alunos nas avaliações externas e controles referentes ao contexto das escolas.

⁴ Para mais detalhes ver Filho e Loeb (2011) e Gorard, Hordosy e Siddiqui (2013).

⁵ A responsabilização escolar não se restringe à premiação ou bonificação salarial, podendo incluir punições. No entanto, no Brasil, as experiências de responsabilização escolar utilizam, em sua maioria, a bonificação salarial ou incentivos salariais coletivos aos profissionais das escolas. Como exemplos, citamos o programa “Escola do Novo Milênio” do estado do Ceará, o programa “Nova Escola” do estado do Rio de Janeiro, o sistema de bonificação salarial do estado de São Paulo, o programa “Prêmio Anual do Desempenho” do município do Rio de Janeiro, entre outros. Para mais detalhes ver Brooke (2006), Brooke e Cunha (2011) e Koslinski *et al.* (2015).

⁶ Para desenhos transversais é observado o desempenho dos estudantes aferidos em determinado momento (FRANCO, 2004; LEE, 2004).

⁷ São exemplos dessas políticas: a Nova Escola no estado do Rio de Janeiro, o sistema de bônus salarial do estado de São Paulo, o prêmio Escola do Novo Milênio do estado do Ceará, e o bônus salarial para professores alfabetizadores do município de Sobral (BROOKE, 2006, 2008; BROOKE; CUNHA, 2011).

Assim, de um lado trazem diferentes concepções de qualidade da educação e, de outro, podem incentivar diferentes ações e estratégias por parte dos atores escolares.

Alguns indicadores utilizados para bonificação salarial trazem dimensões de equidade e/ou controles referentes à taxa de participação de alunos nas avaliações externas, visando coibir práticas de *gaming* ou estratégias de atores escolares que promovam desigualdades. Por exemplo, a Secretaria de Educação do Estado de São Paulo elaborou, em 2009, o Índice de Desenvolvimento da Educação do Estado de São Paulo (Idesp), que, além de fluxo, considera a distribuição dos alunos em níveis de desempenho (abaixo do básico, básico, adequado e avançado). Portanto, para que uma escola alcance um alto Idesp, é necessário ter o maior número de alunos distribuídos pelas faixas de desempenho adequado e avançado (SOARES, 2009). A utilização de um indicador que somente leva em conta a média do desempenho dos alunos para a bonificação salarial ou premiação das escolas poderia incentivar a priorização dos alunos com maior desempenho em detrimento daqueles com baixo desempenho ou abaixo do nível adequado. Ao não focalizar no desempenho médio da escola, como é o caso do Idesp, esse tipo de indicador pune a escola que não focar nos alunos que estão nos níveis básico e abaixo do básico (SOARES, 2009). Assim, o indicador de qualidade da educação proposto por Soares (2009), ao considerar um padrão claro de expectativa de aprendizagem e uma distribuição ideal de desempenho dos alunos, estaria alinhado a concepções de equidade baseadas em “igualdade de base” ou “igualdade de conhecimento adquirido como princípio de justiça”⁸ (CRAHAY, 2013; RIBEIRO, 2014).

⁸ Para Ribeiro (2014), a falta de definição clara de um padrão do que as crianças precisariam aprender em cada etapa da educação básica seria uma barreira para o monitoramento da distribuição equitativa do conhecimento.

Além disso, alguns indicadores utilizados em políticas de responsabilização também incorporam mecanismos para inibir práticas de exclusão dos alunos com baixo desempenho nas avaliações externas para elevar a *performance* média da escola. Por exemplo, o programa “Bônus de Desempenho Educacional”, adotado pelo estado de Pernambuco, estabelece um limite de faltas no dia da avaliação, fixado em 10%. Caso esse limite seja superado, a cada aluno adicional ausente

é imputada a nota zero na prova, diminuindo o desempenho médio da escola (FERRAZ, 2009).

Por fim, indicadores mais recentes incluem ainda outros controles relacionados ao contexto da escola. Por exemplo, a Secretaria Estadual da Educação do Espírito Santo utiliza como base para o cálculo do Indicador de Desenvolvimento da Escola (IDE) medidas de nível socioeconômico dos estudantes, aferidas a partir dos questionários contextuais do Programa de Avaliação da Educação Básica do Espírito Santo (Paebes), e medidas do esforço da escola. Assim, o indicador é elaborado com base na premissa de que ensinar os estudantes do ensino médio, com maior trajetória escolar, é mais complicado do que ensinar estudantes com menor tempo de escola. Em tese, tal metodologia permite que escolas com nível socioeconômico e segmento educacional distintos tenham as mesmas chances de receber a bonificação.⁹

⁹ Disponível em: <www.educacao.es.gov.br/web/bonus_ide.htm>. Acesso em: fev. 2014.

Mesmo nesse contexto de evolução dos indicadores, somente recentemente observamos estudos sobre o impacto do uso desses indicadores educacionais para alocação de recursos e/ou bonificação salarial, o desempenho médio das escolas ou as desigualdades intraescolares ou interescolares. No que diz respeito ao emprego de indicadores para a alocação de recursos, o estudo de Alves *et al.* (2016) observou o impacto do PDE-Escola sobre o desempenho das escolas na Prova Brasil do 5º e 9º anos em 2009 e 2011. Os autores buscaram verificar se o programa, a partir de apoio financeiro e/ou técnico às escolas priorizadas, era capaz de reverter os resultados educacionais de escolas de baixo desempenho, bem como um possível efeito de longo prazo em escolas que participaram do programa e que deixaram de receber o recurso, depois de implementado o plano de melhoramento. Para tanto, o estudo comparou escolas que nunca foram priorizadas pelo PDE-Escola (grupo controle) com aquelas que foram priorizadas pelo programa, mas não receberam recursos do PDE-Escola (tratamento 1), as que foram priorizadas e receberam recursos do PDE-Escola (tratamento 2) e as que já haviam deixado de receber recursos do PDE-Escola (tratamento 3).¹⁰ No que diz respeito ao tratamento 2, o estudo observou impacto pequeno e negativo para matemática

¹⁰ Como estratégia analítica, o estudo utilizou modelo de efeitos-fixos (nível das escolas e por ano) e diversos controles relacionados a características das escolas (infraestrutura, escolaridade e experiência do diretor, escolaridade e experiência dos docentes) e do alunado (escolaridade dos pais/responsáveis dos alunos, raça, posse de bens, entre outros) para reduzir o problema de heterogeneidade não observada, isto é, o fato de as escolas priorizadas e não priorizadas terem características sistematicamente diferentes (ALVES *et al.*, 2016).

no 5º ano e não significativo para o desempenho em língua portuguesa. Para o 9º ano, as análises indicaram impacto pequeno e positivo para o desempenho em matemática, mas não significativo para língua portuguesa. As análises ainda observaram impacto negativo mais acentuado para os tratamentos 1 e 3 para matemática no 5º ano e efeito positivo e pequeno para o tratamento 1 e negativo para o tratamento 3 para língua portuguesa no 9º ano. Por fim, o estudo encontrou efeitos heterogêneos para diferentes Unidades da Federação. Por exemplo, efeitos positivos do tratamento 2 no desempenho de matemática e língua portuguesa no 5º ano foram detectados para escolas localizadas nos estados de São Paulo, Minas Gerais e Mato Grosso do Sul, enquanto efeitos negativos foram registrados para escolas localizadas no estado do Rio de Janeiro (ALVES et al., 2016).

Alves *et al.* (2016) propõem algumas hipóteses para impacto pequeno e heterogêneo do apoio técnico e financeiro trazido pelo PDE-Escola, bem como para o impacto negativo do tratamento 3 (escolas que deixaram de receber o PDE-Escola), tais como inexistência de consequências para escolas que não melhoram resultados e/ou não cumprem o plano de melhoramento proposto, falta de apoio externo para elaboração do planejamento das escolas e autonomia das secretarias municipais e estaduais para selecionar escolas participantes do PDE-Escola, que podem adotar critérios que não necessariamente seguem aqueles sugeridos pelas diretrizes do programa.

Os estudos de Carneiro (2015) e Carneiro e Irfi (2014) avaliaram o impacto marginal do “Prêmio Escola Nota Dez”, do estado do Ceará, sobre o desempenho das escolas apoiadas e premiadas na Prova Brasil de matemática e língua portuguesa. Os estudos partiam do pressuposto de que a introdução de mecanismos de competição, como estabelecimento de metas, monitoramento de resultados e distribuição de incentivos financeiros, e a disseminação de boas práticas de gestão e pedagógicas a partir de estratégias de cooperação técnico-pedagógica seriam capazes de produzir melhoria de aprendizado dos alunos, em especial das escolas contempladas com recursos do programa.

Para tanto, Carneiro e Irffi (2014) utilizaram como grupo de tratamento as escolas que receberam o prêmio referente aos resultados de 2008, e como grupo de controle escolas públicas do estado do Ceará que não receberam prêmio ou apoio nas edições do programa de 2008, 2009 e 2010. As análises estimaram modelos de diferenças-em-diferenças com controle de características das escolas, tais como rede, localização da escola (rural ou urbana), quantidade de alunos matriculados, infraestrutura (existência de laboratórios de informática e ciências, quadra de esportes, bibliografia, computadores, acesso à internet) e quantidade de funcionários da escola. O modelo observou impacto positivo sobre o desempenho de língua portuguesa (9,5 pontos) e matemática (12,75) das escolas premiadas. Para as escolas apoiadas, os modelos estimaram impacto negativo, mas não significativo, na proficiência de língua portuguesa e matemática.

Além de modelos de diferenças-em-diferenças, Carneiro (2015) utilizou *propensity score matching* para criar grupos de controle com escolas mais parecidas com aquelas que receberam prêmio ou apoio (grupo de tratamento). As análises indicaram resultados similares aos modelos anteriores: impacto positivo sobre o desempenho das escolas premiadas em língua portuguesa (12,2 pontos) e matemática (14,3 pontos). Novamente, para as escolas apoiadas, as análises indicaram um efeito negativo sobre ambos os desempenhos, mas estatisticamente não significativo.

Os estudos anteriormente discutidos buscaram observar o impacto da alocação de recursos baseados em resultados dos sistemas de avaliação somente sobre as escolas que receberam os recursos. Esse tipo de análise de impacto parece adequado quando o foco são escolas de baixo desempenho, se partimos do pressuposto de que os estados que adotam esse uso dos sistemas de avaliação não pretendem estimular a redução do desempenho das escolas para que recebam recursos financeiros e/ou apoio técnico. No entanto, estudos sobre o impacto da alocação de recursos e/ou da bonificação salarial para escolas com alto desempenho ou que atingem uma determinada meta podem adotar dois focos distintos de análise: de um lado, focalizar o impacto marginal,

comparando escolas que receberam o recurso e/ou o bônus salarial (tratamento) com aquelas que não receberam (controle), com ambos os grupos de escolas estando submetidos a pressões de responsabilização escolar e/ou competindo para receber recursos/bônus salarial; e, de outro, enfatizar o impacto global, comparando escolas que foram submetidas às pressões de responsabilização e, portanto, que estão sob o incentivo da política para melhorar resultados para receber bônus salarial e/ou receber recursos (tratamento), com aquelas que não foram submetidas a tais pressões e/ou incentivos de responsabilização (controle).

No que diz respeito ao impacto marginal de políticas de responsabilização escolar, ou de remuneração variada, Ferraz e Burns (2012) observaram o efeito do “Bônus de Desempenho Educacional” adotado pelo governo do estado de Pernambuco a partir de 2008. O estudo utilizou um desenho de descontinuidade de regressão, uma vez que o programa estabelecia pontos de corte arbitrários para as metas das escolas. Tal desenho permitiu comparar, de um lado, escolas com desempenhos de partida similares, mas que, por estarem em faixas diferentes de desempenho, tinham distintas metas de acréscimo e, de outro, escolas que atingiram a meta e receberam o bônus salarial com escolas que aumentaram seu desempenho e não alcançaram a meta por pouco. Entre outras questões, o estudo buscou observar se metas mais ambiciosas estimulavam maior *performance* das escolas e se alcançar as metas e ganhar a bonificação salarial levaria a um maior esforço por parte dos atores escolares no ano seguinte.

Ferraz e Burns (2012) observaram que as escolas que tinham metas mais ambiciosas obtiveram maiores ganhos de proficiência (desvio-padrão de 0,31 em português e de 0,15 em matemática). Os autores também verificaram que as escolas que não alcançaram por pouco a meta, e não receberam o prêmio, no ano subsequente obtiveram um ganho maior se comparado com as escolas que alcançaram a meta e ganharam o bônus.

Oshiro, Scorzafaze e Dorigan (2015) investigaram o efeito global do programa de bonificação de professores e funcionários da Secretaria de Educação do Estado de São Paulo,

implantado em 2008, sobre o desempenho dos alunos na Prova Brasil. Os autores partem do pressuposto de que a bonificação de docentes traria um incentivo para mudanças em práticas escolares (maior esforço e comprometimento com a aprendizagem dos alunos) e conseqüentemente um efeito sobre a nota dos alunos em testes padronizados.¹¹ O estudo utilizou a metodologia *propensity score matching* para selecionar escolas do grupo de controle e de tratamento e levou em conta os seguintes grupos de controle: escolas municipais paulistas; escolas estaduais de outros estados brasileiros; escolas estaduais da região Sudeste; e escolas estaduais dos estados do Rio de Janeiro e do Espírito Santo. Nesse caso as escolas do grupo de controle necessariamente pertenceriam a outras redes, uma vez que todas as escolas da rede estadual de São Paulo estavam submetidas aos incentivos do programa. Os resultados das análises trouxeram evidências de impacto positivo do programa para matemática (0,42 desvio-padrão) e português (0,14 desvio-padrão) para o 5º ano do ensino fundamental entre 2007 e 2009. Porém, esses ganhos não foram sustentados a longo prazo, pois, em 2011, o estudo observou uma diminuição dos efeitos positivos. A análise não encontrou efeito significativo do programa no desempenho do 9º ano e, para alguns grupos de controle utilizados, foram constatados resultados negativos sobre a proficiência dos alunos para 2009 e 2011.

Já o trabalho de Scorzafave, Dorigan e Ferreira (2015) buscou observar o impacto de políticas estaduais de bonificação de professores do ensino fundamental sobre a desigualdade (desvio-padrão do desempenho)¹² e sobre a distribuição de alunos entre os níveis de proficiência no 5º e 9º anos do ensino fundamental. Para tanto, cada rede estadual foi classificada em uma das seguintes categorias: redes que não adotaram nenhum tipo de bonificação desde 2007; aquelas que adotaram políticas de bonificação utilizando indicadores que consideravam a dimensão equidade;¹³ e as que adotaram políticas de bonificação com indicadores baseados no desempenho médio das escolas.¹⁴ Os autores argumentam que estudos já realizados estão mais preocupados em investigar o impacto de políticas de bonificação salarial de professores

11 Oshiro, Scorzafave e Dorigan (2015) observaram que, apesar de na média o sistema de metas adotado pelo estado de São Paulo não configurar entre os mais rígidos adotados pelos estados brasileiros, a distribuição das escolas por cumprimento das metas mostra que, para algumas, as metas são muito difíceis e, para outras, são mais fáceis. A assimetria de incentivos verificada poderia comprometer o impacto do programa.

12 Para medir desigualdades de proficiência intraescolares, os autores utilizaram a medida proposta por Ferreira e Gignoux (2014), o desvio-padrão, por ser mais adequada para a análise de indicadores construídos via Teoria da Resposta ao Item.

13 Indicadores que usam metodologia similar ao Idesp.

14 As análises utilizaram dados por escola da Prova Brasil de 2007, 2009 e 2011 e modelos de efeitos fixos e aleatórios com dados em painel, empregando controles relacionados aos alunos (escolaridades das mães, proporção de alunos que trabalham, trajetória escolar dos alunos, cor e estrutura familiar), aos professores, aos diretores e à infraestrutura da escola.

sobre a elevação de proficiência dos alunos, a partir de medidas da média das proficiências das escolas, e questionam a ausência de estudos de impacto sobre a desigualdade de proficiências. Partem do pressuposto de que as políticas de bonificação que utilizam indicadores com dimensão de equidade seriam capazes de diminuir desigualdades intraescolares.

O estudo observou que, para o 5º ano nos resultados para língua portuguesa, as redes que adotaram políticas com indicadores com dimensão de equidade tiveram aumento de desigualdades de notas (desvio-padrão) maior do que as escolas em estados que não adotaram política de bonificação salarial. Para o 9º ano não foram encontrados resultados significativos.

No entanto, nas análises sobre a composição dos alunos de acordo com o nível de proficiência, verificou-se que as redes com bonificação que utilizavam indicadores com a dimensão de equidade apresentaram uma maior redução dos alunos que estavam nos níveis básico e abaixo do básico. Além disso, um efeito oposto foi observado nas redes que adotaram bônus sem o componente de indução a equidade: aumento da proporção de alunos nos níveis básico e abaixo do básico. Os resultados mostram que programas com indicadores indutores de diminuição de desigualdades talvez não sejam capazes de diminuir desigualdades intraescolares, mesmo conseguindo aumentar a proporção de alunos nos níveis adequados de proficiência. Scorzafave, Dorigan e Ferreira (2015) propõem como hipótese que os incentivos trazidos por indicadores com dimensão de equidade não sejam suficientes para induzir escolas a concentrar esforços na melhoria da aprendizagem de alunos com proficiência insuficiente ou abaixo do adequado.

Se o estudo de Scorzafave, Dorigan e Ferreira (2015) observou o impacto de políticas que utilizam diferentes indicadores com e sem a dimensão de equidade sobre desigualdades intraescolares, poucos trabalhos investigam o impacto dos diferentes usos sobre práticas escolares. Uma das poucas análises que investigam efeitos intermediários dos programas de bonificação salarial sobre práticas escolares foi realizada por Bresolin (2014).¹⁵

Finalmente, cabe salientar que não encontramos estudos com evidências sobre o efeito do uso de distintos indicadores

15 O foco do estudo foram as seguintes práticas escolares: práticas pedagógicas dos professores; absenteísmo e rotatividade dos professores; e interlocução com familiares para garantir a frequência dos alunos.

educacionais para bonificação salarial sobre a taxa de participação dos alunos nas avaliações externas.

INDICADORES EDUCACIONAIS E ALOCAÇÃO DE RECURSOS NO ESTADO DO CEARÁ

O “Prêmio Escola Nota Dez”, adotado pelo governo do estado do Ceará desde 2009,¹⁶ distribui prêmios para 150 escolas públicas com melhor desempenho e apoio para as 150 escolas com pior resultado no IDE-Alfa. Apesar de o prêmio poder ser considerado uma política de responsabilização escolar de alta consequência, seu desenho se difere das demais políticas adotadas por estados brasileiros.

A premiação não ocorre sob a forma de bonificação salarial, mas sim como provisão de recursos destinados às escolas para diversos gastos. Além disso, a política prevê apoio financeiro para escolas de baixo desempenho, pressupondo ainda um apoio técnico-pedagógico das escolas premiadas para as escolas apoiadas.¹⁷ Os recursos referentes ao prêmio e ao apoio são repassados em duas parcelas. Para as escolas premiadas, a primeira parcela equivale a 75% e a segunda a 25% do valor total do prêmio. As orientações para aplicação dos recursos permitem que até 20% dos valores recebidos em ambas as parcelas sejam destinados a bonificação de professores, núcleo gestor e servidores lotados nas escolas. Os recursos para as escolas apoiadas são divididos em duas parcelas iguais, sendo que somente a segunda (até 20% do valor total da parcela) pode ser utilizada para bonificação salarial (CEARÁ, 2009a, 2009b, 2009c, 2011a). Assim, apenas uma pequena parte desses recursos seria utilizada para bonificação salarial.

As escolas que foram premiadas de acordo com os resultados do IDE-Alfa de 2008 a 2010 receberam o montante de R\$ 2.500,00 por aluno avaliado, enquanto aquelas apoiadas nesses anos receberam R\$ 1.250,00 por aluno avaliado no Spaece-Alfa.¹⁸ Dessa forma, o apoio às escolas de baixo desempenho se aproxima do formato do PDE-Escola, dos usos de sistemas de avaliação externa para distribuição de recursos para escolas de baixo desempenho (CEARÁ, 2009a).

16 Estado do Ceará, Lei n. 14.371, de 19 de junho de 2009.

17 Formam-se pares, de modo que cada uma das escolas premiadas fica responsável por acompanhar uma escola apoiada.

18 As escolas premiadas a partir dos resultados do IDE-Alfa de 2008 receberam a primeira parcela dos recursos em 2009 e a segunda em 2010.

19 O valor do repasse do PDE-Escola para escolas com até 99 alunos é equivalente a R\$ 10.000,00. Já para aquelas apoiadas pelo Prêmio Escola Nota 10, esse valor pode variar entre R\$ 25.000,00 (escolas com 20 alunos avaliados) e R\$ 123.750,00 (escolas com 99 alunos avaliados).

No entanto, os valores repassados pelo “Prêmio Escola Nota Dez” para escolas apoiadas são consideravelmente maiores.¹⁹

Outras características singularizam o programa de responsabilização escolar adotado pelo estado do Ceará. Em primeiro lugar, embora o “Prêmio Escola Nota Dez” tenha sido instituído pelo governo do estado, ele premia e apoia escolas públicas tanto estaduais como municipais. Assim, considerando que em 2008 cerca de 90% das matrículas em escolas públicas de ensino fundamental pertenciam às redes municipais, percebe-se que esta configura uma política estadual que distribui recursos diretos principalmente para escolas municipais.

Adicionalmente, o indicador de qualidade utilizado como medida para distribuição de recursos, o IDE-Alfa, até 2011, levava em conta duas dimensões, sendo uma comumente empregada – a proficiência média padronizada das escolas – e outra pouco usual – a taxa de participação dos alunos na avaliação. Assim, diferentemente de indicadores adotados por outros estados brasileiros, o IDE-Alfa cearense não continha uma dimensão de fluxo,²⁰ mas sim uma componente intimamente relacionada à própria realização da avaliação externa, força motriz da política de responsabilização.

A proficiência empregada na primeira componente do indicador, adotada com intuito de avaliar a alfabetização, tem como ponto de corte o valor 200, pontuação indicativa de que um determinado aluno apresenta as competências desejadas para leitura. O limite inferior da escala é igual a zero, considerando que esta é uma série inicial. Já a participação seria medida pela diferença entre o número de alunos matriculados, registrados oficialmente pelo Censo Escolar do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), e o número de alunos presentes na avaliação.²¹

Para serem elegíveis ao prêmio ou ao apoio, as escolas precisavam ter ao menos 20 alunos avaliados e, para receberem o prêmio, deveriam obter um IDE-Alfa maior ou igual a 8,5. Entre as que alcançavam o ponto de corte, somente as 150 escolas com maior desempenho recebiam o prêmio.²² Dessa forma, o “Prêmio Escola Nota Dez” estabelece uma

20 O Ideb de anos iniciais, por exemplo, ao incluir uma dimensão de fluxo em seu cálculo, pune estratégias de aumento da reprovação em séries anteriores para aumentar a proficiência do 5º ano. Já o IDE-Alfa, na ausência de um componente de fluxo, não seria capaz, por exemplo, de punir escolas com maior reprovação no 1º ano, anterior à série avaliada.

21 “É acrescido no resultado número de matrículas de alunos ocorridas até 31 de agosto e subtraído o número de alunos que foram transferidos da escola, até a referida data” (CEARÁ, 2012, p. 88).

22 Somente na edição que tomou como base os resultados do IDE-Alfa de 2008, menos de 150 escolas alcançaram o ponto de corte e, portanto, somente 148 foram premiadas.

meta única para todas escolas para o recebimento do prêmio. Já a maioria dos estados que utilizam indicadores de qualidade para bonificação salarial estabelece metas múltiplas, levando em conta o ponto de partida do desempenho das escolas em períodos anteriores.²³

A partir de 2012, o programa passou a premiar e apoiar escolas por meio de indicadores que consideram o desempenho das escolas no 2º ano (IDE-Alfa) e no 5º ano (IDE-5) do ensino fundamental (CEARÁ, 2011b). A inclusão da dimensão da taxa de participação no IDE-Alfa e IDE-5 tinha o objetivo de coibir práticas de *gaming*, tais como a exclusão de alunos de menor desempenho. No entanto, somente a partir de 2011 os indicadores de qualidade incorporaram uma dimensão de equidade com o acréscimo de um “fator de ajuste”²⁴ que leva em conta a porcentagem de alunos em cada nível de proficiência.²⁵ Assim, apesar de já existir antes de 2011 uma referência pedagógica do nível desejável de alfabetização, foi somente a partir da incorporação do fator de ajuste que o IDE-Alfa e o IDE-5²⁶ passaram a incorporar um componente indutor de equidade. De acordo com documento oficial a inclusão desse fator teria por finalidade estimular as escolas a incluírem um maior percentual de alunos nos níveis adequados. A necessidade dessa dimensão “torna-se relevante pelo fato de a média de proficiência não expressar devidamente o grau de universalização do aprendizado” (CEARÁ, 2012, p. 34). Assim, o indicador passa a considerar não somente a média padronizada, mas também a distribuição de alunos entre diversos níveis de proficiência e, de forma similar ao Idesp, pune escolas com maior concentração de alunos com proficiência abaixo dos níveis adequados.

BASE DE DADOS

O banco de dados a partir do qual as análises dispostas neste estudo foram realizadas tomou como cadastro a listagem de escolas públicas municipais e estaduais presentes no Censo Escolar da educação básica do Inep, entre 2007 e 2011. Essa listagem foi utilizada como base à qual foram agregadas informações fornecidas pelo governo do Ceará

23 Por exemplo, o programa de bônus por desempenho adotado pela Secretaria de Educação do Estado de Pernambuco divide as escolas em três grupos, de acordo com seus desempenhos em cada série e disciplina avaliada (baixo, médio e alto). As metas estipuladas são iguais dentro de cada grupo de referência. Dessa forma, uma escola pode ter metas distintas para leitura e matemática, assumindo que o desempenho em cada disciplina estaria inserido em dois grupos diferentes. Já em São Paulo, de acordo com as regras do programa de Qualidade na Escola, as metas anuais para cada escola são calculadas de forma similar às metas municipais do Ideb, a partir de uma função de crescimento logística (FERRAZ, 2009).

24 O prêmio é recebido em 2012 a partir dos resultados do IDE-Alfa de 2011. Portanto, os indicadores referentes aos resultados de 2011 já incorporam o fator de ajuste.

25 O fator de ajuste assume um valor percentual de 0% a 1,00%, a partir da distribuição das crianças em cada nível de aprendizado da classificação utilizada: 0,00% x percentual de crianças no nível considerado não alfabetizado; 0,25% x percentual de crianças no nível considerado alfabetização incompleta; 0,50% x percentual de crianças no nível considerado intermediário; 0,75% x percentual de crianças no nível considerado suficiente; 1,00% x percentual de crianças no nível considerado desejável (CEARÁ, 2011c).

26 O IDE-5, de forma similar ao IDE-Alfa, é composto pelas seguintes dimensões: proficiência padronizada para a avaliação do 5º ano; taxa de participação na avaliação do 5º ano (Space); e fator de ajuste para língua portuguesa e matemática para o 5º ano (CEARÁ, 2011c).

(sobre o Spaece-Alfa e o IDE-Alfa, além da identificação das escolas premiadas e apoiadas), assim como dados provenientes de outras fontes do próprio Inep, como a Prova Brasil.

Como um dos nossos intuítos era observar graficamente a evolução dos indicadores escolares no período analisado, captando potenciais mudanças de escala em sua série histórica, implementou-se um conjunto de filtros. Consideraram-se, por exemplo, apenas as escolas ativas e que haviam sido avaliadas em todos os anos de interesse, de 2007 a 2011, e aquelas que tivessem as informações necessárias às análises nesses mesmos anos. Além disso, tomou-se como ponto de corte o limite de pelo menos 15 alunos avaliados no Spaece-Alfa entre 2007 e 2011, já que escolas muito pequenas (com poucos alunos nas séries contempladas pela política) não podem ser premiadas ou apoiadas. Finalmente, não foram consideradas nas análises duas escolas que, com base nos resultados do IDE-Alfa de 2008 a 2010, receberam entre 2009 e 2011, alternadamente, tanto uma premiação como um apoio. Apenas a seleção de “escolas em atividade” em todos os anos do período estudado já reduziu as pouco mais de 12.700 escolas, que constavam na listagem original proveniente do Censo Escolar (Inep), para 6.850 escolas. Após a aplicação dos demais filtros e dos procedimentos de seleção de casos para compor os grupos de controle, chegou-se a uma base final que continha um total de 1.357 escolas públicas.

Segundo as informações cedidas pelo governo do Ceará, entre 2008 e 2010, 355 escolas foram premiadas e 404 foram apoiadas no âmbito do “Prêmio Escola Nota Dez”. Dessas, 207 escolas premiadas e 223 apoiadas permaneceram na base de dados final, por apresentarem dados para todos os anos da avaliação e por registrarem ao menos 15 alunos avaliados em cada um dos anos considerados. Embora o número de escolas apoiadas/premiadas anualmente seja fixado em 150, algumas escolas foram beneficiadas em mais de um ano. Das 207 escolas premiadas, 51 receberam o prêmio duas vezes e outras oito foram premiadas nos três anos observados. Entre as 233 escolas apoiadas, 73 receberam apoio duas vezes e dez o receberam três vezes.

ESTRATÉGIA METODOLÓGICA E DEFINIÇÃO DOS GRUPOS DE CONTROLE

Como visto anteriormente, este trabalho tem o propósito de investigar os efeitos marginais do uso do IDE-Alfa e da política de responsabilização à qual esse indicador está vinculado – o “Prêmio Escola Nota Dez” – sobre distintos resultados escolares. Não ignoramos que os incentivos do prêmio podem gerar um efeito global para todas as escolas com mais de 20 alunos avaliados que, a princípio, concorreriam para receber tanto o prêmio como o apoio. Mas o foco da presente análise é observar possíveis efeitos do recebimento do recurso para escolas apoiadas e premiadas, bem como um possível incentivo diferenciado para escolas com desempenho próximo ao daquelas premiadas.

Nesse sentido, nossa intenção foi acompanhar a evolução de uma série de indicadores educacionais (desempenho médio das escolas, medidas de equidade e taxas de participação dos alunos na avaliação Spaece-Alfa), verificando variações em seu comportamento entre os momentos anteriores e os posteriores ao início da política.

Uma noção fundamental para a elaboração das análises realizadas aqui corresponde à ideia de *contrafactual*. Resumidamente, considerando uma intervenção, política ou, mais genericamente, uma causa cujo impacto se deseja averiguar, representada por uma variável binária (indicando presença ou ausência da intervenção), o modelo *contrafactual*, sobre o qual se baseia a lógica dos estudos experimentais e quase-experimentais,²⁷ pressupõe a existência de dois estados causais bem definidos, para os quais todos os membros de uma população de interesse estão expostos – o tratamento, quando um caso é afetado pela intervenção; e o controle, quando o caso não é afetado (MORGAN; WINSHIP, 2012).

Idealmente, um único caso, ou grupo de casos, deveria assumir, num mesmo ponto do tempo, ambos os estados (tratamento e controle) para que o efeito ou impacto da intervenção pudesse ser calculado pela diferença simples entre os valores obtidos para os mesmos casos, simultaneamente expostos e não expostos à causa avaliada. Na prática, como é impossível atingir essa situação de comparabilidade ideal,

27 Para mais detalhes sobre desenhos experimentais e quase-experimentais, ver Campbell e Stanley (1963), Cook e Campbell (1979) e Shadish, Cook e Campbell (2002).

buscamos que dois ou mais grupos de comparação sejam equivalentes; alguns recebem a intervenção (grupos experimentais) outros não, sendo considerados grupos de controle (CANO, 2002; RAMOS, 2009).

No caso concreto das análises relacionadas ao “Prêmio Escola Nota Dez”, existem claramente dois grupos experimentais: o das *escolas premiadas*, com melhores resultados no IDE-Alfa; e o das *escolas apoiadas*, que receberam recursos por apresentarem os piores resultados nesse mesmo indicador. Basicamente, interessa avaliar se as *escolas premiadas*, assim como as *apoiadas*, apresentam evolução significativamente diferente de outras escolas não contempladas pela política, nos indicadores supracitados.

Assim, nosso grande desafio metodológico consiste em obter ao menos dois grupos de controle que sejam bem similares, em algumas características-chave, respectivamente, aos grupos de escolas premiadas e de escolas apoiadas.

Quando não é possível a solução metodologicamente ideal, que consiste em compor os grupos experimental e controle por atribuição aleatória, uma opção para obter grupos comparáveis é utilizar técnicas de pareamento estatístico (*matched sampling*), que procuram associar a unidades já definidas do grupo experimental uma ou mais unidades com características similares, formando o grupo de controle. São utilizadas uma ou mais variáveis percebidas como relevantes para a análise, para definir grupos de controle equalizados, construídos a partir da seleção de casos de uma amostra maior de candidatos (MORGAN; WINSHIP, 2007; RUBIN, 2006).

28 Ver, por exemplo, o livro de Rubin (2006).

29 Ver, por exemplo, o livro de Guo e Fraser (2009).

Existem inúmeras técnicas de pareamento estatístico.²⁸ Atualmente, os procedimentos mais utilizados consistem em pareamentos baseados nos chamados escores de propensão ou escores de propensidade (*propensity score matching*),²⁹ que agrupam controles a unidades do grupo experimental segundo uma probabilidade de pertencimento ao grupo experimental estimada *a priori*, com base em um conjunto determinado de variáveis (GUO; FRASER, 2009; RAMOS, 2009).

Com base nos trabalhos de Guo e Fraser (2009), Rosenbaum e Rubin (1983) e Sekhon (2011), realizamos um pareamento de acordo com os procedimentos apresentados a seguir, para

a formação dos grupos de escolas a serem utilizados como controle.

ESTIMATIVA DOS ESCORES DE PROPENSÃO

Tomando como variáveis resposta duas variáveis indicadoras que apontavam escolas que tinham ou não recebido premiação em algum dos três anos de interesse (2009, 2010 e 2011, com base nos seus resultados no IDE-Alfa em 2008, 2009 e/ou 2010), ou tinham ou não recebido apoio no mesmo período, foram ajustadas duas regressões logísticas estimando, respectivamente, as probabilidades de ser premiada ou apoiada, de cada uma das escolas presentes na base. Entre as variáveis explicativas empregadas nos modelos (ver Quadro 1),³⁰ constavam medidas de pré-teste, tais como as proficiências no Spaece-Alfa 2007 e 2008, localização e contexto territorial (microrregião da escola, ou se configura uma escola urbana ou rural) e outros indicadores escolares como a média do número de alunos matriculados no 2º ano e previstos para avaliação do Spaece-Alfa (2007-2012) e infraestrutura escolar. Foram também utilizados indicadores de contexto municipal, como a porcentagem de pais com alta escolaridade no município em que a escola está localizada e a renda *per capita* do município.³¹

PAREAMENTO DAS ESCOLAS

Com base nos escores de propensão estimados na regressão logística, a seleção das escolas para o grupo de controle foi realizada a partir de um algoritmo de pareamento que associa a cada escola premiada ou apoiada uma escola que lhe é semelhante, mas que não foi contemplada pela política entre 2009 e 2011, com base em seus resultados no IDE-Alfa entre 2008 e 2010, de acordo com o escore de propensão mais semelhante possível. Trata-se do pareamento denominado 1:1 – uma escola no grupo de controle associada a cada escola do experimental. O algoritmo escolhido usa o método do vizinho mais próximo (*nearest-neighbor matching*), sem reposição das unidades pareadas.³² As escolas de ambos os grupos foram primeiramente ordenadas pelo valor dos escores de propensão. A seguir, a primeira das escolas premiadas foi

³⁰ Não utilizamos variáveis referentes à composição do alunado no nível da escola (sexo, cor, escolaridade dos pais, nível socioeconômico), porque os alunos do 2º ano não respondem questionários contextuais. Tentamos utilizar a composição do alunado do 5º ano (cor, nível socioeconômico e escolaridade dos pais), no entanto, essas variáveis não estavam associadas ao desempenho médio dos alunos do 2º ano.

³¹ As variáveis no nível do município foram retiradas da amostra do Censo Demográfico 2010/IBGE.

³² Os procedimentos de pareamento foram realizados com a biblioteca “*matching*”, utilizando o aplicativo estatístico R. Para mais detalhes, ver Sekhon (2011).

selecionada e associada à escola não premiada que possui o valor mais próximo do escore de propensão. Se duas escolas do grupo de controle possuem o mesmo valor ocorre uma atribuição aleatória. Como o procedimento é sem repetição, todas as escolas já pareadas são removidas do conjunto de candidatas. Esse procedimento é realizado repetidas vezes, até que todas as escolas premiadas tenham seus respectivos controles. A mesma técnica foi realizada para as escolas apoiadas. Outra restrição imposta ao processo de seleção era que uma escola não poderia ser controle simultaneamente de uma escola premiada e de uma escola apoiada.

QUADRO 1 - Descrição das variáveis utilizadas no pareamento

VARIÁVEL	FONTE	DESCRIÇÃO
Proficiência 2º ano (2007)	Spaace (2007)	Proficiência média da escola - Spaace-Alfa 2007
Proficiência 2º ano (2008)	Spaace (2008)	Proficiência média da escola - Spaace-Alfa 2008
Número médio de alunos previstos para avaliação	Spaace (2007-2012)	Número médio de alunos matriculados no 2º ano e previstos para avaliação do Spaace-Alfa entre 2007 e 2012
Infraestrutura da escola	Censo Escolar (2008)	Medida contínua padronizada (média 0 e desvio-padrão 1) obtida a partir de análise fatorial das informações sobre a presença de sala de professores, laboratório de informática, laboratório de ciências, quadra de esportes, biblioteca/sala de leitura e equipamentos. Alpha de Conbrach = 0,857
Escolas urbanas	Censo Escolar (2008)	Localização da escola (1 = urbana, 0 = rural)
Microrregião	IBGE	Variável categórica com as microrregiões geográficas em que as escolas estão localizadas
Escolaridade dos chefes de domicílio (município)	Censo IBGE (2010)	Porcentagem de crianças no município matriculadas em escolas públicas no 2º ano cujo chefe do domicílio tem ao menos ensino médio completo
Renda <i>per capita</i> (município)	Censo IBGE (2010)	Renda <i>per capita</i> do município em que a escola está localizada

Fonte: Elaboração dos autores.

NOVOS GRUPOS DE CONTROLE

Além dos dois grupos experimentais e dos dois grupos de controle obtidos via pareamento, as escolas não pareadas foram divididas em outros dois grupos com base em suas probabilidades (escores de propensão) de receberem o prêmio ou serem apoiadas. Quando a probabilidade estimada de receber o prêmio era maior do que a de receber um apoio, a escola era classificada na categoria “Não pareada próxima às premiadas”; caso contrário, inseria-se a escola no grupo das

“Não pareadas próximas às apoiadas”. Esses grupos também foram utilizados nas análises.

A Tabela 1 traz um panorama das diferenças existentes entre os grupos de comparação constituídos, observando variáveis utilizadas no pareamento e também a proficiência média das turmas de 5º ano, na Prova Brasil. Fica nítida a maior proximidade dos grupos pareados entre si, bem como sua distinção em relação aos demais grupos. As escolas premiadas partem de um nível de proficiência maior, considerando-se os resultados do Spaece-Alfa em 2007 e 2008 (antes do início da política). Além disso, as escolas premiadas possuem menos alunos nas turmas de 2º ano, são menos urbanas e estão em municípios mais pobres. Ao contrário do que se poderia supor, essas escolas possuem ainda índices de infraestrutura menores do que as apoiadas. Tais resultados não são necessariamente novos e já haviam sido observados por Oliveira (2016). A autora constatou que municípios cearenses maiores e com alta renda (em especial a capital, Fortaleza) tinham mais escolas apoiadas do que premiadas no “Escola Nota Dez”.

TABELA 1 – Descrição das variáveis utilizadas no pareamento

	ESCOLAS						TOTAL
	PAREADAS APOIADAS	PAREADAS NÃO APOIADAS	NÃO PAREADAS PRÓXIMAS ÀS APOIADAS	NÃO PAREADAS PRÓXIMAS ÀS PREMIADAS	PAREADAS NÃO PREMIADAS	PAREADAS PREMIADAS	
VARIÁVEIS/ PAREAMENTO	MÉDIAS						
Proficiência - Spaece (2007)	108,3	111,7	122,0	129,8	133,9	140,7	123,5
Proficiência - Spaece (2008)	102,5	109,0	120,3	139,9	154,0	174,8	131,0
Número de alunos	74,4	77,0	57,6	53,3	48,4	48,5	60,3
Infraestrutura escolar	0,78	0,88	0,47	0,31	0,40	0,38	0,55
Escola em área urbana	79,8%	78,5%	78,2%	81,3%	72,9%	75,8%	77,7%
Escol. chefes nos municípios	12,4%	12,6%	11,8%	11,9%	11,1%	10,9%	11,8%
Renda <i>per capita</i> municipal	408,3	417,8	295,4	270,6	249,1	249,7	317,5
TAMANHO DA AMOSTRA	223	223	358	139	207	207	1357
	16,4%	16,4%	26,4%	10,2%	15,3%	15,3%	100,0%
PROFIC. 5º ANO/ PROVA BRASIL	MÉDIAS						
Matemática (2007)	155,4	160,8	157,4	158,7	157,4	159,8	158,1
Língua Portuguesa (2007)	171,6	175,6	172,2	173,1	171,8	175,0	173,1
TAMANHO DA AMOSTRA	154	163	258	93	137	141	946
	16,3%	17,2%	27,3%	9,8%	14,5%	14,9%	100,0%

Fonte: Elaboração dos autores com base nos dados do Spaece-Alfa, Censo Escolar/Inep e Prova Brasil/Inep.

Outra constatação interessante é que as diferenças observadas entre os grupos na proficiência do 2º ano não foram encontradas nas médias calculadas para o 5º ano, com dados da Prova Brasil, independentemente de vislumbrarmos os resultados de matemática ou língua portuguesa. Para esses indicadores, os pontos de partida da série histórica (no ano de 2007) são bastante similares entre os grupos de comparação, sejam experimentais ou de controles.

EFEITOS DO IDE-ALFA E “PRÊMIO ESCOLA NOTA DEZ” SOBRE INDICADORES

Basicamente, o objetivo das análises foi observar o impacto do “Prêmio Escola Nota Dez” (premiação e apoio), nos primeiros anos da política dessa responsabilização – 2009, 2010 e 2011 (com base nos resultados das escolas em 2008, 2009 e 2010). Nesses anos, o prêmio estava atrelado ao IDE-Alfa, ou seja, ao desempenho dos alunos do 2º ano do ensino fundamental. Esse período é metodologicamente interessante pelo fato de o impacto dos recursos destinados às escolas (via premiação ou apoio) se restringir ao 2º ano. A partir de 2011, a política passou também a destinar recursos a escolas com base nos resultados do 5º ano, o que deixa as análises mais complexas e as inferências mais confusas de interpretar. Além disso, os valores dos recursos distribuídos para as escolas foram mais baixos a partir de 2011, o que introduz a possibilidade de trabalhar com uma lógica de dose-resposta. Tais análises deverão ser realizadas em um próximo trabalho.

Considerando-se as características supramencionadas do IDE-Alfa e do “Prêmio Escola Nota Dez”, bem como as evidências observadas em estudos anteriores, esperávamos encontrar algumas tendências e associações e, assim, pudemos antecipar algumas hipóteses.

No que diz respeito ao desempenho, esperávamos encontrar um efeito positivo tanto para escolas premiadas quanto para as apoiadas. Em suma, presumimos que o recebimento de recursos pelas escolas, via premiação ou apoio, se traduza numa posterior elevação da proficiência média dessas escolas. No entanto, supõe-se um impacto mais acentuado nas escolas premiadas, já que estas recebem uma quantia maior de recursos e podem, já com a primeira parcela do prêmio, bonificar seus professores. Adicionalmente, tais recursos deveriam afetar a rotina e o clima de toda a escola, não apenas das séries contempladas pela avaliação, de modo que efeitos devem ser constatáveis não apenas no 2º ano, mas também no 5º ano do ensino fundamental – aparecendo nas análises dos dados da Prova Brasil. Esperávamos um efeito ainda maior para as escolas que, durante o período analisado, receberam a premiação e o apoio duas ou mais vezes.

Também esperamos encontrar efeitos no desempenho das escolas com características próximas às escolas premiadas, se comparadas com aquelas que têm menor expectativa de receber recursos, ou seja, escolas com desempenho de partida intermediário. Como somente as 150 escolas com maior desempenho recebem prêmio, podemos supor mudanças de práticas ou maior incentivo nas escolas que estão próximas ao ponto de corte (IDE-Alfa maior ou igual a 8,5), o que não deveria ocorrer entre aquelas que apresentam um indicador muito distante desse ponto e que talvez não julguem possível alcançar tal meta.

Uma lógica similar à colocada anteriormente, mas com efeitos negativos, poderia ocorrer com escolas próximas de serem apoiadas: a adoção da estratégia de piorar seu desempenho para receber o apoio destinado às piores escolas. No entanto, acreditamos que seja uma hipótese menos plausível pelo ônus que implicaria tal decisão. Primeiramente, estar entre as piores escolas traz a possibilidade de uma rotulagem negativa, passando atestado de má qualidade do trabalho desenvolvido pelo corpo docente e administrativo e, possivelmente, resultando na estigmatização da instituição e de seus alunos e profissionais. Ademais, as escolas apoiadas têm necessariamente que receber, além do apoio financeiro, um apoio técnico, passando a ser acompanhada por uma das escolas premiadas. É plausível pensar que tal situação seja percebida por diretores e diretoras como intromissão e comprometimento da autonomia escolar. Por fim, o recebimento da totalidade do apoio financeiro está vinculado a uma série de condicionalidades. Assim, a decisão de jogar desse modo com a regra da política, baixando o desempenho da escola, não seria uma decisão tão fácil de ser tomada, numa avaliação entre custos e benefícios.

Esperávamos ainda observar um aumento das desigualdades no desempenho das escolas premiadas e apoiadas. Como nos três primeiros anos o indicador utilizado no programa não continha uma dimensão de equidade, ensinar somente uma parte dos alunos, os mais próximos da média, poderia ser uma estratégia adotada pelas escolas para aumentar seus indicadores, garantindo o recebimento da segunda

parcela do prêmio. A mesma estratégia estaria disponível para escolas com características e desempenho próximos aos das escolas premiadas, na tentativa de ganhar o prêmio. Por fim, como taxa de participação é uma das dimensões do IDE-Alfa, esperamos observar um aumento nesse indicador, o que deve ocorrer especialmente para as escolas premiadas, para as apoiadas e para aquelas com características e desempenho próximos aos das escolas premiadas.

As próximas seções trazem gráficos com séries históricas elaboradas para os indicadores educacionais já mencionados, visando mostrar como a evolução dessas diferentes medidas se comporta para os seis grupos de comparação constituídos. Além dos gráficos, foram ajustados modelos de regressão múltipla para estimar o impacto de uma escola ter sido premiada ou apoiada. A especificação desses modelos é relativamente simples, uma vez que incorporam apenas uma variável de pré-teste (o resultado do indicador defasado para um ano anterior à implementação da política) e variáveis indicadoras do grupo (de tratamento ou controle) ao qual as escolas pertencem.³³

Para cada um dos cinco indicadores educacionais observados,³⁴ foram obtidos dois conjuntos de modelos de regressão para avaliar o impacto do “Prêmio Escola Nota Dez” no que tange à premiação (amostra 1), depois em relação ao apoio (amostra 2). As amostras utilizadas são formadas pelas escolas que se beneficiaram diretamente da política, premiadas ou apoiadas, pelas escolas que compunham seus respectivos grupos de controle (selecionadas por pareamento estatístico), além de escolas não pareadas – separadas, com base em seus escores de propensão, em escolas com perfil mais próximo das escolas premiadas ou das apoiadas. Variáveis dicotômicas indicam o pertencimento a uma das três categorias: grupo experimental (premiadas ou apoiadas); grupo de controle pareado; e grupo de controle não pareado. As escolas presentes no grupo de controle formado por pareamento têm sua variável dicotômica omitida, funcionando como categoria de referência. Estas são comparadas tanto com as escolas beneficiadas (dando uma estimativa do efeito da política) como com as não pareadas.

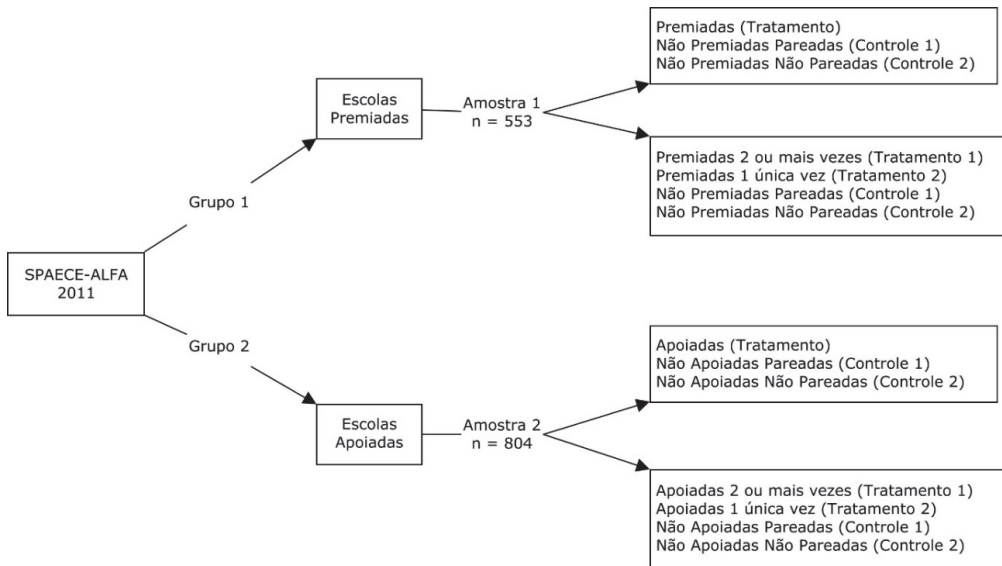
33 Antes de avaliar o efeito médio dessa política sobre a totalidade das escolas públicas do estado do Ceará, as análises se preocuparam em elucidar o fato específico de uma escola receber os recursos relacionados à premiação ou ao apoio. No campo da avaliação de políticas públicas, programas e projetos são comuns e já consolidadas as análises que abarcam apenas dois pontos no tempo (antes e após a implementação de uma dada intervenção), seja com desenhos experimentais ou quase-experimentais. Nesse sentido, consideramos que os modelos de regressão utilizados suprem os objetivos do artigo: explorar e apresentar evidências sobre a eficácia de certos componentes da política de responsabilização escolar cearense. Acreditamos que modelos estatísticos específicos para dados longitudinais poderiam oferecer novas evidências sobre os efeitos e mecanismos associados ao funcionamento da política. Esses serão realizados em análises posteriores.

34 Desempenho no 2º ano (Spaace-Alfa 2011) e no 5º ano (Prova Brasil 2011 - língua portuguesa e matemática), desvio-padrão da proficiência do Spaace-Alfa e taxa de participação na mesma avaliação.

Adicionalmente, foram estimados modelos levando em conta o número de vezes que se recebeu um prêmio ou apoio. As escolas que receberam a intervenção foram divididas em outros dois grupos, segundo o número de prêmios (ou apoios) recebidos: uma única vez ou mais de uma vez. Tais modelos dão às análises uma perspectiva de dose-resposta e oferecem evidências sobre os efeitos de diferentes níveis de intensidade da intervenção. Em suma, foram ajustadas quatro regressões para cada indicador,³⁵ como mostra a Figura 1, que exemplifica os modelos estimados para a proficiência no 2º ano (Spaace-Alfa 2011).

35 Considerando as limitações de espaço do presente trabalho, nem todos os modelos estatísticos serão dispostos. Seus resultados, entretanto, serão comentados e discutidos nas seções posteriores.

FIGURA 1 – Modelos de regressão ajustados para Spaace-Alfa 2011



Fonte: Elaboração dos autores com base nos dados do Censo Escolar e Spaace-Alfa.

IMPACTOS SOBRE DESEMPENHO ESCOLAR: SPAECE-ALFA E PROVA BRASIL

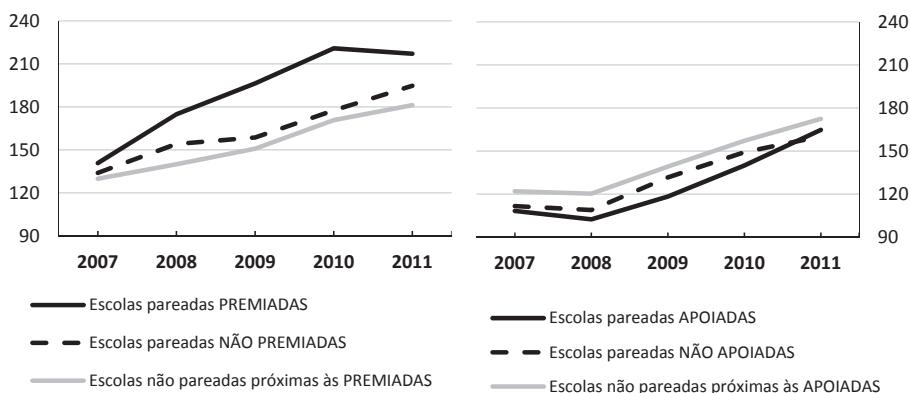
O “Prêmio Escola Nota Dez” começou a vigorar em 2009. Nesse ano, a primeira parcela dos recursos destinados às premiações e apoios foi repassada às escolas com base nos resultados do IDE-Alfa de 2008. A segunda parcela foi repassada em 2010. Nos anos posteriores, mais escolas receberam os

recursos, acumulando novas unidades ao conjunto de escolas contempladas pela política.

Como as escolas recebem os recursos em momentos distintos, mudanças no nível ou padrão das séries históricas não são tão simples de serem visualizadas, como num desenho clássico de séries temporais interrompidas (SHADISH; COOK; CAMPBELL, 2002). O padrão que se pretende detectar, na busca por evidências sobre o efeito da política, é um distanciamento progressivo das séries dos grupos experimental e controle com o passar do tempo.

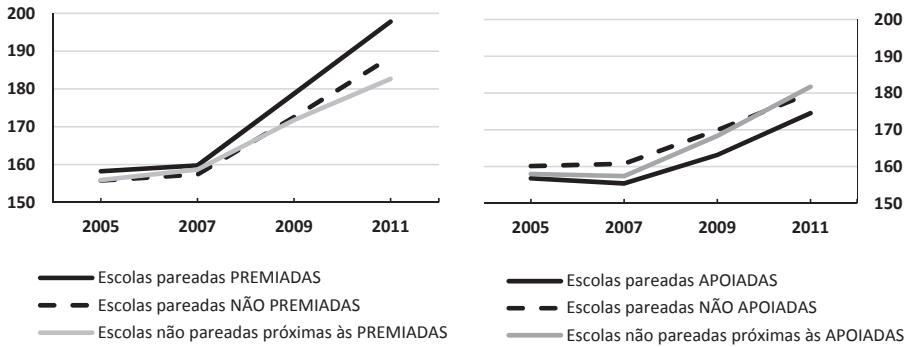
No Gráfico 1, que traz a evolução da proficiência em língua portuguesa (no Spaece-Alfa), é possível ver essa separação das séries históricas para as escolas premiadas. As séries partem de um ponto similar, mas se distanciam com o tempo, embora haja uma desaceleração desse movimento de 2010 para 2011. Há uma tendência mais geral de crescimento do desempenho, que pode, inclusive, corresponder a um efeito global da política de responsabilização, mas o mais relevante é a evolução diferenciada das escolas premiadas. O mesmo não é observado para as escolas apoiadas, cujo padrão da evolução é muito similar em relação ao comportamento dos grupos de controle. No Gráfico 2, o mesmo tipo de padrão é visualizado quando se consideram os resultados da Prova Brasil. O fato de a série ser bienal torna ainda mais evidente as diferenças das escolas premiadas.

GRÁFICO 1 - Proficiência em língua portuguesa no Spaece-Alfa



Fonte: Spaece-Alfa.

GRÁFICO 2 – Proficiência em língua portuguesa na Prova Brasil



Fonte: Prova Brasil/Inep.

36 Os modelos de regressão linear descrevem a associação ou a relação linear entre duas variáveis, uma variável resposta (ou variável dependente) cuja variação estaria condicionada pelo comportamento de uma ou mais variáveis explicativas (ou independentes). Especificamente nas análises que se preocupam com inferências causais, a associação entre as variáveis explicativa e resposta pode ser interpretada como uma estimativa do impacto da intervenção social analisada. Nesse sentido, os coeficientes de regressão calculados descrevem o efeito ou impacto das premiações ou apoios do “Prêmio Escola Nota Dez” sobre os indicadores educacionais analisados – medidas de desempenho, equidade e taxa de participação das escolas.

Para estimar o impacto do “Prêmio Escola Nota Dez” sobre o desempenho escolar, 12 modelos de regressão³⁶ foram ajustados, considerando três indicadores de desempenho, para os grupos de escolas premiadas (Tabela 2) e apoiadas (Tabela 3). As análises avaliaram o impacto das premiações em relação às escolas do grupo de controle obtido por pareamento e também observaram os coeficientes desse último grupo em relação às escolas que não foram pareadas.

Complementando a informação dos gráficos apresentados, com os quais pretendemos observar o comportamento dos indicadores ao longo dos anos e, mais especificamente, verificar quaisquer distanciamentos que possam ser atribuíveis à implementação do “Prêmio Escola Nota Dez”, os modelos de regressão tentam estabelecer o impacto dessa política a partir da comparação dos resultados escolares entre as escolas premiadas (e apoiadas) e seus respectivos grupos de controle, constituídos por pareamento. Uma vez que tais grupos de controle podem ser entendidos como contrafactuais válidos, os modelos de regressão devem oferecer uma boa estimativa do impacto da política nas escolas. Os modelos também incluem como variável explicativa a própria variável dependente defasada num momento anterior à implementação da intervenção (para o ano de 2007). Tal variável serve como uma medida de pré-teste, controlando diferenças entre grupos que existiam num primeiro momento.

As regressões reiteraram o que foi antes constatado nos gráficos, isto é, um efeito positivo³⁷ das premiações, que pode ser verificado tanto para os modelos que trataram do Spaece como para aqueles que utilizaram dados da Prova Brasil. No que diz respeito ao efeito sobre o Spaece-Alfa, o fato de uma escola ter sido premiada eleva sua proficiência, em média, 22 pontos, em relação às escolas do grupo controle. Adicionalmente, modelos similares aos apresentados na Tabela 2, mas separando as escolas que receberam o prêmio uma vez daquelas que o receberam duas ou três vezes, indicaram aumento mais elevado no desempenho das escolas que foram premiadas mais de uma vez no período analisado. Essas últimas apresentaram um incremento médio de 41 pontos, em relação às escolas do grupo de controle, contra 14 pontos das escolas que receberam o prêmio uma única vez.

Uma tendência similar foi observada no que diz respeito aos resultados da Prova Brasil, com um incremento médio para as escolas premiadas de 7 pontos nas provas de língua portuguesa e de 10 pontos em matemática, em comparação com as escolas do grupo de controle. Esses resultados, contudo, parecem restringir-se às escolas que receberam mais de uma premiação. De fato, as escolas premiadas duas ou mais vezes apresentaram um incremento médio de 18 pontos nas provas de língua portuguesa e de 23 pontos nas provas de matemática, enquanto aquelas com uma única premiação não obtiveram coeficientes significativos. Os resultados observados para o 5º ano sugerem certo transbordamento dos efeitos do prêmio para toda a escola, sobretudo, no caso de premiações recorrentes.

Por outro lado, as escolas não pareadas, mas com características mais próximas das escolas premiadas, apresentaram uma proficiência em média 13 pontos inferior àquela das escolas pareadas no grupo de controle (que não receberam o prêmio). Logo, existe alguma evidência para corroborar a hipótese de que a própria expectativa da premiação atua como fator motivador, que pode provocar, nas escolas pareadas do grupo de controle, um crescimento diferencial da proficiência menor do que o registrado pelas escolas premiadas, mas superior àquele que ocorre nas escolas dispostas no meio da distribuição, com menor possibilidade de premiação.

37 Um efeito positivo (representado por uma estatística com valor maior do que zero) indica uma relação diretamente proporcional, ou seja, a implementação do programa eleva o indicador educacional da escola, em relação às escolas que não receberam uma premiação ou apoio. Um efeito negativo reflete uma relação inversamente proporcional entre as variáveis, isto é, a presença da intervenção implica uma redução no indicador educacional observado. O valor do coeficiente de regressão representa o incremento (ou decréscimo) esperado no indicador educacional, devido à implementação da política (o recebimento de uma premiação ou apoio), mantidas constantes as variações dos demais indicadores presentes no modelo, que servem como controles estatísticos.

Os modelos para o impacto do “Prêmio Escola Nota Dez” sobre a proficiência das escolas apoiadas não trouxeram qualquer evidência que ajudasse a corroborar a hipótese de um impacto positivo da política. Nos três indicadores observados, as relações entre a implementação da política e o desempenho escolar não foram significativas.³⁸ Os efeitos não foram significativos independentemente do número de vezes que as escolas foram apoiadas.

TABELA 2 - Regressão para indicadores de proficiência - escolas premiadas

VARIÁVEIS	PROFICIÊNCIA SPAECE-ALFA		PROVA BRASIL ¹ LÍNGUA PORTUGUESA		PROVA BRASIL ¹ MATEMÁTICA	
	B	ERRO PADRÃO	B	ERRO PADRÃO	B	ERRO PADRÃO
Escolas premiadas	21,08***	3,93	6,93**	2,28	10,00**	3,19
Não pareada próxima das premiadas	-13,24**	4,38	-7,19**	2,44	-10,34**	3,41
Desempenho em 2007	-0,078+	0,043	0,405***	0,07	0,396***	0,09
Constante	184,20***	6,35	114,28***	9,76	134,07***	14,67
Número de observações	553		421		421	

Fonte: Elaboração dos autores com base nos dados do Spaece-Alfa e Prova Brasil/Inep.

*** p < 0,001, ** p < 0,01, * p < 0,05, + p < 0,10

[1]: Nas análises realizadas com dados da Prova Brasil, o acesso ao questionário contextual, aplicado em conjunto com a avaliação, permitiu estimar medidas de nível socioeconômico dos alunos e, por conseguinte, das escolas. Assim, os modelos para proficiência em língua portuguesa e matemática na Prova Brasil incluem como variável de controle um indicador que representa o percentual de responsáveis por alunos com alta escolaridade. Como essa variável apresentava casos sem informações, sua inclusão reduziu a amostra de 553 para 421 escolas, no caso das premiadas e de 804 para 669 escolas, no caso das apoiadas.

TABELA 3 - Regressão para indicadores de proficiência - escolas apoiadas

VARIÁVEIS	PROFICIÊNCIA SPAECE-ALFA		PROVA BRASIL ¹ LÍNGUA PORTUGUESA		PROVA BRASIL ¹ MATEMÁTICA	
	B	ERRO PADRÃO	B	ERRO PADRÃO	B	ERRO PADRÃO
Escolas apoiadas	4,72	3,47	-1,83	1,55	-1,89	1,87
Não pareada próxima das apoiadas	11,61***	3,15	2,64*	1,34	3,61*	1,63
Desempenho em 2007	0,06	0,04	0,24***	0,05	0,24***	0,06
Constante	153,50***	5,2	133,63***	7,12	144,90***	9,58
Número de observações	804		669		669	

Fonte: Elaboração dos autores com base nos dados do Spaece-Alfa e Prova Brasil/Inep.

*** p < 0,001, ** p < 0,01, * p < 0,05, + p < 0,10

IMPACTOS SOBRE A EQUIDADE

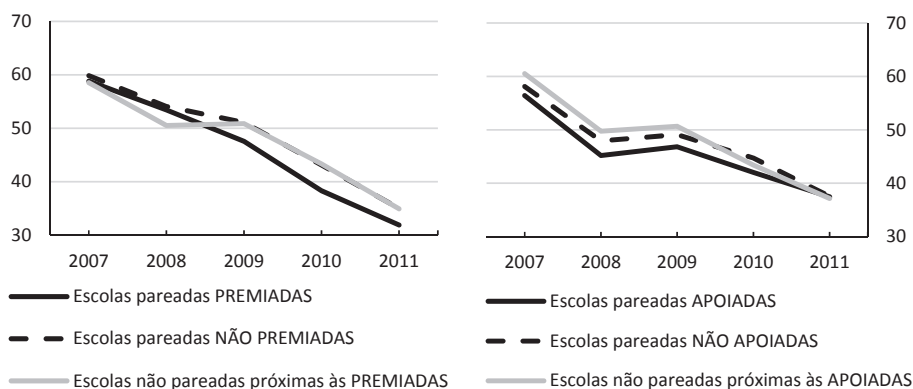
Um dos possíveis efeitos perversos das políticas de responsabilização reside no fato de que, em alguns casos, as tentativas de elevar os índices monitorados podem conduzir gestores, diretores e professores a políticas e práticas geradoras de desigualdades. Tendo em mente essa ameaça, alguns dos indicadores mais recentes procuraram incorporar em suas métricas componentes cujo intuito era controlar esses desvios de finalidade.

No caso do Ceará, no período observado neste estudo, não havia no IDE-Alfa qualquer preocupação com a dimensão equidade, o que nos fez supor que o aumento do desempenho viria acompanhado do crescimento da desigualdade dentro das escolas. Contrariando tais expectativas, o que se percebeu entre as escolas premiadas que, como visto anteriormente, tiveram um aumento significativo na proficiência foi uma redução da desigualdade intraescolar, expressa pela queda relativa do desvio-padrão da proficiência.

No Gráfico 3 é possível observar uma tendência geral de diminuição da variabilidade nas pontuações no interior das escolas. Essa tendência é mais pronunciada entre as escolas premiadas e mais suave nas escolas apoiadas e seus controles. Na comparação com os grupos de controle, percebe-se a partir de 2009 uma diferenciação nos padrões das séries históricas dos grupos experimentais e controles, com padrões diferentes entre premiadas e apoiadas. Para as primeiras há uma redução das desigualdades superior à registrada no grupo de controle, enquanto para as segundas verifica-se leve crescimento que aproxima as séries das escolas apoiadas e não apoiadas.

38 A significância estatística, num modelo de regressão, passa por testar a hipótese nula de que o efeito estimado da intervenção é diferente (maior ou menor) do que zero. Para tanto, verifica-se a probabilidade (estatística "p-valor") de o efeito estimado ter ocorrido ao acaso. Quando as estatísticas "p-valor" são muito elevadas, não há evidências empíricas suficientes para aceitar que o efeito registrado reflete um impacto real da política, diferente de zero. Em contrapartida, valores pequenos do "p-valor" indicam que seria muito pouco provável observar ao acaso certa magnitude de efeito, se este, de fato, não existisse. Os pontos de corte para avaliar a significância estatística são chamados de nível de significância. O valor mais amplamente utilizado é 0,05, de modo que valores inferiores a esse limite são estatisticamente significativos. Podem ser considerados, entretanto, outros limites, como 0,001 ou 0,01 (valores mais conservadores e mais exigentes) ou 0,10 (um valor menos conservador, com o qual é mais fácil de verificar significância estatística). No presente estudo, esses quatro níveis de significância foram sinalizados ao lado das estimativas dos efeitos. A ausência de qualquer marcador indica que o resultado estimado não foi estatisticamente significativo.

GRÁFICO 3 - Desvio-padrão da proficiência no Spaece-Alfa



Fonte: Spaece-Alfa.

As análises multivariadas corroboram os achados relatados acima, mostrando, primeiramente, a redução estatisticamente significativa do desvio-padrão para as escolas premiadas, em relação ao seu grupo de controle, bem como a inexistência de diferenças nas desigualdades entre escolas não premiadas e as não pareadas, que podem ser entendidas como um único grupo de controle. Tomando a dispersão da proficiência como medida de desigualdade dentro das escolas, isso significa dizer que, de modo geral, as escolas premiadas tornaram-se internamente menos desiguais se comparadas com as do grupo de controle. No entanto, nos modelos estimados separando as escolas premiadas mais de uma vez daquelas que receberam o prêmio apenas uma vez, o efeito sobre a equidade somente foi observado para as premiadas mais de uma vez durante os três anos analisados.

No caso das escolas apoiadas parece não haver diferenças significativas em relação aos grupos de controle, indicando a inexistência de impacto dos apoios sobre a desigualdade. A mesma tendência foi observada nos modelos que separaram as escolas que receberam o apoio duas ou três vezes daquelas que receberam somente uma vez.

TABELA 4 - Regressão para o desvio-padrão da proficiência no Spaece-Alfa

VARIÁVEIS	B	ERRO PADRÃO	VARIÁVEIS	B	ERRO PADRÃO
Escolas premiadas	-3,02**	1,05	Escolas apoiadas	-0,011	0,863
Não pareada próxima das premiadas	0,032	1,18	Não pareada próxima das apoiadas	-0,482	0,779
Desempenho em 2007	0,076	0,038	Desempenho em 2007	0,4	0,026
Constante	30,41***	2,39	Constante	35,238***	1,646
Número de observações	553		Número de observações	804	

Fonte: Elaboração dos autores com base nos dados do Spaece-Alfa.

*** p < 0,001, ** p < 0,01, * p < 0,05, + p < 0,10.

IMPACTOS SOBRE A PRESENÇA NAS AVALIAÇÕES EXTERNAS

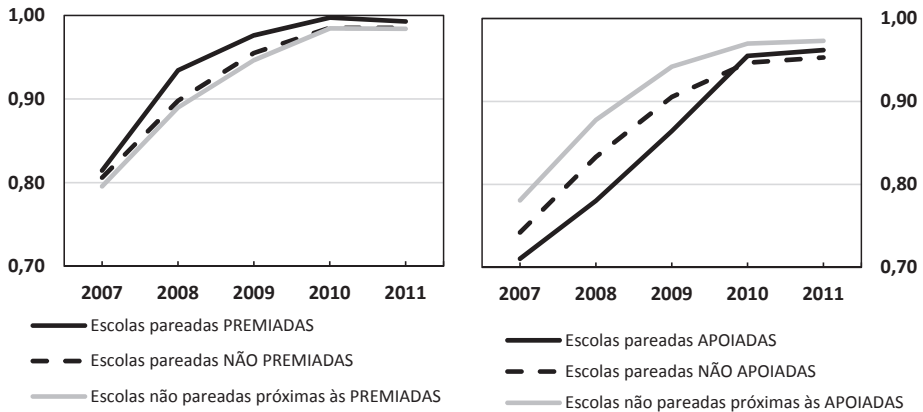
Como visto, a taxa de participação é um dos indicadores que compõem o IDE-Alfa, sendo uma característica que confere singularidade ao indicador cearense. No Gráfico 6, observa-se que a proporção de alunos que participavam das avaliações cresceu rapidamente entre 2007 e 2011.

Para o grupo das escolas premiadas, as diferenças em relação aos grupos de controle surgem em 2008, antes mesmo da implementação da política de premiações. Tal situação poderia indicar certa apropriação, pelas escolas municipais mais eficazes do Ceará, das ferramentas e informações oferecidas pela avaliação educacional e expressar ainda o reconhecimento de suas possibilidades para a gestão dos cotidianos educacionais. Já em 2010, quase a totalidade dos alunos comparecia às avaliações externas nas escolas premiadas, de modo que alguns poucos alunos faltantes seriam suficientes para diminuir a probabilidade de estar ou não entre as escolas premiadas.

Entre as escolas apoiadas, as taxas de participação eram cerca de 10 pontos percentuais menores do que as registradas nas escolas premiadas, em 2007, e se aproximavam de 100% em 2011. Mais uma vez, a tendência de crescimento das taxas de participação é anterior à implementação da política de premiações. No entanto, é possível observar a partir de 2009 uma mudança na tendência das escolas apoiadas que não foi acompanhada pelas séries dos grupos de

controle. Assim, a implementação do “Prêmio Escola Nota Dez” parece ter certo impacto tanto para as escolas premiadas como para as apoiadas, mas o impacto é mais claro para as últimas.

GRÁFICO 4 - Proporção da participação dos alunos nas avaliações externas



Fonte: Spaece-Alfa.

Os modelos de regressão múltipla confirmaram as impressões obtidas pela análise dos gráficos e apontaram um aumento na participação, expressa pela proporção de alunos presentes na avaliação para as premiadas. Novamente, esse impacto foi mais acentuado nas escolas com mais de uma premiação. Um pequeno impacto sobre as taxas de participação também foi detectado no que se refere às escolas apoiadas. Aqui também houve um ligeiro aumento na taxa de participação em relação ao registrado no grupo de controle. Contudo, nos modelos que separam as escolas que receberam o apoio duas ou mais vezes daquelas que receberam somente uma vez, esse impacto restringe-se às escolas apoiadas uma única vez. Além disso, as escolas que não faziam parte do grupo de controle (não pareadas), mas que tinham perfil próximo ao das escolas apoiadas, também acusaram um aumento significativo nas taxas de participação.

TABELA 5 - Modelos de regressão para proporção de alunos presentes na avaliação externa

VARIÁVEIS	B	ERRO PADRÃO	VARIÁVEIS	B	ERRO PADRÃO
Escolas premiadas	0,007*	0,003	Escolas apoiadas	0,012*	0,005
Não pareada próxima das premiadas	-0,001	0,003	Não pareada próxima das apoiadas	0,016***	0,005
Taxa de participação em 2007	-0,040***	0,012	Taxa de participação em 2007	0,103***	0,017
Constante	0,953***	0,01	Constante	0,877***	0,013
Número de observações	553		Número de observações	804	

Fonte: Elaboração dos autores com base nos dados do Spaece-Alfa.

*** p < 0,001, ** p < 0,01, * p < 0,05, + p < 0,10

CONSIDERAÇÕES FINAIS

As análises apresentadas ainda são preliminares e restringem-se aos primeiros anos de funcionamento do “Prêmio Escola Nota Dez”. Os resultados indicam, como esperado, um aumento no desempenho das escolas premiadas e das escolas pareadas não premiadas. Embora o IDE-Alfa não apresentasse uma preocupação com a dimensão de equidade nos primeiros anos do prêmio, os resultados não sugeriram aumento na dispersão das proficiências entre as escolas premiadas (em busca da segunda parcela do prêmio) ou entre as escolas pareadas não premiadas (com maiores chances de concorrer ao prêmio em anos subsequentes). Observou-se uma tendência contrária à esperada: diminuição da dispersão das proficiências para as escolas premiadas.³⁹ No entanto, não encontramos esses efeitos (diminuição ou aumento da dispersão das proficiências) para as escolas que receberam apoio financeiro, quando comparadas com aquelas pareadas não apoiadas.

Uma particularidade do indicador utilizado para alocação de recursos no estado do Ceará – a dimensão referente à taxa de participação – parece inibir as práticas de exclusão de alunos. Observamos um efeito positivo na taxa de participação tanto para as escolas premiadas como para as apoiadas. Assim, mesmo sendo foco de uma política de alta consequência, o IDE-Alfa não parece ser inflado por tal uso.

39 Análise similar foi realizada estimando-se a porcentagem de alunos nos níveis suficiente e adequado de proficiência no Spaece-Alfa. Os resultados indicaram tendências na mesma direção: aumento da porcentagem de alunos nos níveis suficiente e adequado de proficiência nas escolas premiadas e ausência de efeito estatisticamente significativo para escolas apoiadas.

O presente trabalho focalizou somente o efeito marginal do uso de um dos indicadores de qualidade do estado do Ceará para alocação de recursos/bonificação salarial. Os resultados descritivos apresentados nos gráficos sugerem um aumento no desempenho das escolas de todos os grupos, mas a investigação do efeito global do “Prêmio Escola Nota Dez” não foi foco deste estudo. Os resultados encontrados mostram que o recebimento do prêmio foi acompanhado por um efeito positivo sobre a proficiência média das escolas. Também sugerem um incentivo diferenciado sobre as escolas de acordo com sua proficiência de partida, isto é, os incentivos do programa parecem mais fortes para as escolas mais semelhantes àquelas premiadas e que talvez tenham maiores chances e expectativas de alcançar a meta e de receber o prêmio. No entanto, tal tendência (efeito positivo sobre a proficiência) não foi verificada nas análises que focalizaram as escolas apoiadas. Na mesma direção das análises de Alves *et al.* (2016) e Carneiro (2015), o presente estudo parece indicar que o apoio técnico-financeiro não se mostra suficiente para promover melhoria na proficiência dos alunos das escolas de baixo rendimento.

REFERÊNCIAS

- ALVES, F.; ELACQUA, G.; MARTÍNEZ, M.; SANTOS, H. Efeitos do Plano de Desenvolvimento das Escolas nos resultados escolares. *Estudos em Avaliação Educacional*, São Paulo, v. 27, n. 64, p. 128-159, jan./abr. 2016.
- BONAMINO, A.; SOUSA, S. Z. Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola. *Educação e Pesquisa*, São Paulo, v. 38, n. 2, p. 373-388, abr./jun. 2012.
- BRESOLIN, A. *Análise de resultados intermediários das políticas de bônus em escolas públicas estaduais brasileiras*. 2014. 172 f. Dissertação (Mestrado em Administração Pública e Governo) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2014.
- BROOKE, N. O futuro das políticas de responsabilização educacional no Brasil. *Cadernos de Pesquisa*, São Paulo, v. 36, n. 128, p. 377-401, maio/ago. 2006.
- BROOKE, N. Responsabilização educacional no Brasil. *Revista Electrónica de Investigación Educativa*, Baja Califórnia, v. 1, n. 1, p. 93-109, jan. 2008.
- BROOKE, N. Controvérsias sobre políticas de alto impacto. *Cadernos de Pesquisa*, São Paulo, v. 43, n. 148, p. 336-347, jan./abr. 2013.

BROOKE, N.; CUNHA, M. A. A avaliação externa como instrumento da gestão educacional nos estados. *Estudos & Pesquisas Educacionais*, São Paulo, v. 1, p. 17-79, 2011.

CAMPBELL, D. T.; STANLEY, J. C. Experimental and quasi-experimental designs of research. In: GAGE, N. L. *Handbook of research on teaching*. Chicago: Rand McNally, 1963.

CANO, I. *Introdução à avaliação de programas sociais*. São Paulo: FGV, 2002.

CARNEIRO, D. R. F. *Políticas de incentivo à escola melhoram a proficiência no ensino fundamental? Uma avaliação do Prêmio Escola Nota Dez*. 2015. 41 f. Dissertação (Mestrado em Economia) – Curso de Pós-Graduação em Economia (Caen), Universidade Federal do Ceará, Fortaleza, 2015.

CARNEIRO, D. R. F.; IRFFI, G. Política de incentivos à escola melhora a proficiência no ensino fundamental? Uma avaliação do Prêmio Escola Nota Dez. In: ENCONTRO ECONOMIA DO CEARÁ EM DEBATE, 10., Fortaleza, 2014. *Anais...* Fortaleza: Instituto de Pesquisa e Estratégia Econômica do Ceará – Ipece –, 2014.

CEARÁ. Lei n. 14.371 de 19 de junho de 2009. Cria o Prêmio Escola Nota Dez, destinado a premiar as escolas públicas com melhor resultado no Índice de Desempenho Escolar-Alfabetização (IDE-alfa), e dá outras providências. *Diário Oficial do Estado*, Fortaleza, série 3, ano I, n. 111, 19 de junho de 2009a.

CEARÁ. Decreto n. 29.896 de 16 de setembro de 2009. Regulamenta a Lei n. 14.371, de 19 de junho de 2009, que cria o Prêmio Escola Nota Dez, destinado a premiar as escolas públicas com melhor resultado no Índice de Desempenho Escolar-Alfabetização (IDE-Alfa), e dá outras providências. *Diário Oficial do Estado*, Fortaleza, série 3, ano 1, n. 174, 17 de setembro de 2009b.

CEARÁ. Lei n. 14.580 de 21 de dezembro de 2009. Altera dispositivos da Lei n. 14.371, de 19 de junho de 2009. *Diário Oficial do Estado*, Fortaleza, série 3, ano 1, n. 241, 28 de dezembro de 2009c.

CEARÁ. Lei n. 15.052 de 6 de dezembro de 2011. Institui o Prêmio Escola Nota Dez, destinado a premiar as escolas públicas com melhores resultados de aprendizagem no segundo e quinto anos de ensino fundamental, e dá outras providências. *Diário Oficial do Estado*, Fortaleza, série 3, ano III, n. 235, 12 de dezembro de 2011a.

CEARÁ. Lei n. 14.949 de 27 de junho de 2011. Acrescenta dispositivo à Lei n. 14.371/2009 e dá outras providências. *Diário Oficial do Estado*, Fortaleza, série 3, ano III, n. 127, 5 de julho de 2011b.

CEARÁ. Decreto n. 30.797 de 29 de dezembro 2011. Regulamenta a Lei n. 15.052, de 06 de dezembro de 2011, que cria o Prêmio Escola Nota Dez, destinado a premiar escolas públicas com melhores resultados de aprendizagem no segundo e quinto anos do ensino fundamental, e dá outras providências. *Diário Oficial do Estado*, Fortaleza, série 3, ano III, n. 249, 30 de dezembro de 2011c.

CEARÁ. Secretaria da Educação Básica. Coordenadoria de Planejamento e Políticas Educacionais. Célula de Pesquisa e Avaliação Educacional. *Sistema Permanente de Avaliação da Educação Básica do Ceará – SPAECE-Alfa e +5*, 2012: relatório estadual. Fortaleza: SEB, 2012.

COOK, T. D.; CAMPBELL, D. T. *Quasi-experimentation: design & analysis issues for field settings*. Boston: Houghton Mifflin, 1979.

CRAHAY, M. Como a escola pode ser mais justa e mais eficaz? *Cadernos Cenpec*, São Paulo, v. 3, n. 1, p. 9-40, 2013.

DIAMOND, J. B.; SPILLANE, J. P. High stakes accountability in urban elementary schools: challenging or reproducing inequality? *Teachers College Records*, New York, v. 106, n. 6, p. 1145-1176, 2004.

FERRAZ, C. Sistemas educacionais baseados em desempenho, metas de qualidade, e a remuneração de professores: os casos de Pernambuco e São Paulo. In: GIAMBIAGI, F.; HENRIQUES, R.; PESSOA, S.; VELOSO, F. (Org.). *Educação no Brasil: construindo o país do futuro*. Rio de Janeiro: Campus, Elsevier, 2009.

FERRAZ, C.; BURNS, B. Paying teachers to perform: the impact of bônus pay in Pernambuco, Brazil. In: SREE SPRING 2012 CONFERENCE. Washington, DC, 2012.

FERREIRA, F. H. G.; GIGNOUX, J. The measurement of educational inequality: Achievement and opportunity. *The World Bank Economic Review*, v. 28, n. 2, p. 210-246, June 2014.

FIGLIO, D.; LOEB, S. School accountability. In: HANUSHEK, E. A.; MACHIN, S.; WOESSMANN, L. (Ed.). *Handbook in economics of education*. North-Holland: Elsevier, 2011. v. 3

FRANCO, C. Quais as contribuições da avaliação para as políticas educacionais? In: BONAMINO, A.; BESSA, N.; FRANCO, C. *Avaliação da educação básica: pesquisa e gestão*. São Paulo: Loyola, 2004. p. 45-63.

GIMENES, N.; SILVA, V. G.; PRÍNCIPE, L.; LOUZANO, P.; MARICONI, G. Além da Prova Brasil: investimento em sistemas próprios de avaliação externa. *Estudos em Avaliação Educacional*, São Paulo, v. 24, n. 55, p. 12-32, maio/ago. 2013.

GORARD, S.; HORDOSY, R.; SIDDIQUI, N. How unstable are 'school effects' assessed by a value-added technique? *International Education Studies*, North York, v. 6, n. 1, p. 1-9, Jan. 2013.

GUO, S. Y.; FRASER, M. W. *Propensity score analysis: statistical methods and applications*. Thousand Oaks, CA: Sage, 2009. (Advanced Quantitative Techniques in the Social Sciences Series, 12).

HEILIG, J. V.; DARLING-HAMMOND, L. Accountability Texas-Style: the progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, Los Angeles, v.30, n. 2, p. 75-110, June 2008.

HSIEH, C.; URQUIOLA, M. The effects of generalized school choice on achievement and stratification: evidence from Chile's Voucher Program. *Journal of Public Economics*, v. 90, n. 8, p. 1477-1503, Sept. 2006.

KOSLINSKI, M. C.; CARRASQUEIRA, K.; ANDRADE, F. M.; CUNHA, C. P. Políticas de responsabilização educacional: modelos possíveis e uma análise preliminar da política do município do Rio de Janeiro. In: CAVALIERI, A. M.; SOARES, A. J. *Políticas educacionais no Estado do Rio de Janeiro*. Rio de Janeiro: Mauad, 2015.

LADD, H. F. School-based educational accountability systems: the promise and the pitfalls. *National Tax Journal*, Washington DC, v. 54, n. 2, p. 385-400, June 2001.

LEE, V. Medidas educacionais: avaliando a eficácia das escolas em termos de excelência e de equidade. In: BONAMINO, A.; BESSA, N.; FRANCO, C. *Avaliação da educação básica: pesquisa e gestão*. São Paulo: Loyola, 2004. p. 13-43.

LIU, J.; CHARLEY, E.; FULLER, E. Evaluating principals: an analysis of current state strategies. In: UCEA CONVENTION, 27., 2013. Indianapolis: University Council Educational Administration, 2013.

MORGAN, S. L.; WINSHIP, C. *Counterfactuals and causal inference: methods and principles for social research*. Cambridge: Cambridge University, 2007.

MORGAN, S. L.; WINSHIP, C. Bringing context and variability back into causal analysis. In: KINCAID, H. (Ed.). *The Oxford Handbook of Philosophy of Social Science*. Oxford: Oxford University, 2012.

OLIVEIRA, L. X. de. *Políticas de responsabilização como estratégia para promoção de desempenho escolar: um estudo sobre o Prêmio "Escola Nota Dez"*. 2016. 166 f. Tese (Doutorado em Educação) – Programa de Pós-Graduação em Educação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

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*, Rio de Janeiro, v. 69, n. 2, p. 213-249, abr./jun. 2015.

PONTUAL, T. *Remuneração por mérito, desafio para a educação*. São Paulo: Fundação Lemman, 2008.

RAMOS, M. Aspectos conceituais e metodológicos da avaliação de políticas e programas sociais. *Planejamento e Políticas Públicas*, Brasília, DF, n. 32, p. 95-114, jan./jun. 2009.

RIBEIRO, V. M. Que princípio de justiça para a educação básica? *Cadernos de Pesquisa*, São Paulo, v. 44, n. 154, p. 1094-1109, out./dez. 2014.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, Oxford, v. 70, n. 1, p. 41-55, abr. 1983.

RUBIN, D. B. *Matched sampling for casual effects*. 1. ed. New York: Cambridge University Press, 2006.

SCORZAFAVE, L. G.; DORIGAN, T. A.; FERREIRA, J. N. P. Efeito das políticas de bonificação aos professores sobre a desigualdade de proficiência no Ensino Fundamental. In: REUNIÃO DA ABAVE, 8., 2015, Florianópolis. *Anais...* Florianópolis: Abave, 2015.

SEKHON, J. S. Multivariate and propensity score matching software with automated balance optimization: the matching package for R. *Journal of Statistical Software, Forthcoming*, Innsbruck, v. 42, n. 7, p. 1-52, June 2011.

SHADISH, W. R.; COOK, T. D.; CAMPBELL, D. T. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin, 2002.

SOARES, J. F. O índice de desenvolvimento da educação de São Paulo – IDESP: bases metodológicas. *São Paulo em Perspectiva*, São Paulo, v. 23, n. 1, p. 29-41, jan./jun. 2009.

TIMMERMANS, A. C.; DOOLAARD, S.; WOLF, I. Conceptual and empirical differences among various value-added models for accountability. *School Effectiveness and School Improvement*, Londres, v. 22, n. 4, p. 393-413, Aug. 2011.

WEST, M. R.; PETERSON, P. E. The efficacy of choice threats within school accountability systems: results from legislatively induced experiments. *The Economic Journal*, Medford, v. 116, n. 510, p. C46-C62, Mar. 2006.

MARIANE CAMPELO KOSLINSKI

Professora da Faculdade de Educação e do Programa de Pós-Graduação em Educação da Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Rio de Janeiro, Brasil
mckoslinski@gmail.com

EDUARDO RIBEIRO

Professor do Centro de Ciências Sociais da Universidade Estadual do Rio de Janeiro (UERJ), Rio de Janeiro, Rio de Janeiro, Brasil
eduardoribeiroBR@gmail.com

LUISA XAVIER DE OLIVEIRA

Professora do Centro de Ciências da Educação da Universidade Federal do Piauí (UFPI), Picos, Piauí, Brasil
luisaxavier77@yahoo.com.br

Recebido em: JULHO 2016

Aprovado para publicação em: SETEMBRO 2017

