



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE APERFEIÇOAMENTO DE ECONOMISTAS DO NORDESTE
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

CAROLINE DE PAULA BRANDÃO DE CARVALHO

O EFEITO DO PRÊMIO APRENDER PARA VALER NAS EXPECTATIVAS
EDUCACIONAIS DE DISCENTES CEARENSES

FORTALEZA

2019

CAROLINE DE PAULA BRANDÃO DE CARVALHO

O EFEITO DO PRÊMIO APRENDER PARA VALER NAS EXPECTATIVAS
EDUCACIONAIS DE DISCENTES CEARENSES

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Economia.

Orientador: Prof. Dr. Leandro Rocco.

FORTALEZA

2019

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

C322e Carvalho, Caroline de Paula Brandão de.
O efeito do prêmio Aprender para Valer nas expectativas educacionais de discentes cearenses /
Caroline de Paula Brandão de Carvalho. – 2020.
42 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Faculdade de Economia, Administração,
Atuária e Contabilidade, Programa de Pós-Graduação em Economia, Fortaleza, 2020.
Orientação: Prof. Dr. Leandro de Almeida Rocco.

1. Expectativas educacionais. 2. SPAECE. 3. Regressão descontínua. 4. Aprender para Valer. I. Título.
CDD 330

CAROLINE DE PAULA BRANDÃO DE CARVALHO

O EFEITO DO PRÊMIO APRENDER PARA VALER NAS EXPECTATIVAS
EDUCACIONAIS DE DISCENTES CEARENSES

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Economia.

Aprovada em: ___/___/_____.

BANCA EXAMINADORA

Prof. Dr. Leandro Rocco (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Guilherme Irffi
Universidade Federal do Ceará (UFC)

Prof. Dr. Ricardo Brito
Universidade Federal do Ceará (UFC/CAEN)

A Deus.

AGRADECIMENTOS

Primeiramente a Deus, pois sem Ele nem viva estaria.

Aos meus pais, pelo apoio durante a trajetória da vida.

Ao professor Leandro Rocco, pela orientação.

Aos professores Ricardo Brito e Guilherme Irffi, pelas relevantes contribuições à este trabalho.

A todos os demais professores, colegas e funcionários do CAEN, que direta ou indiretamente contribuíram para que pudesse concluir mais uma etapa importante em minha vida.

RESUMO

O objetivo deste trabalho é analisar o efeito do prêmio Aprender para Valer sobre as expectativas educacionais de discentes cearenses; mais especificamente, analisar o efeito de um aluno ter o direito de receber o prêmio em determinado ano sobre suas expectativas no ano seguinte de cursar o Ensino Superior. O prêmio Aprender para Valer foi instituído pelo governo do estado do Ceará desde 2009 e visa contemplar com um notebook os alunos do Ensino Médio que obtenham desempenho adequado na avaliação do Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE). As bases de dados do SPAECE utilizadas compreendem os anos de 2013 e 2014, e fornecem informações sobre a proficiência dos alunos em Matemática e Língua Portuguesa, além de atributos demográficos e socioeconômicos dos alunos, bem como a percepção destes sobre os professores e a escola. A metodologia escolhida foi a de regressão descontínua com scores múltiplos, levando em conta os critérios de elegibilidade do prêmio Aprender para Valer. Os resultados mostraram que os alunos que tinham direito a receber o prêmio aumentaram suas expectativas de cursar o Ensino Superior.

Palavras-chave: Expectativas educacionais. Aprender para Valer. Regressão descontínua. SPAECE.

ABSTRACT

The goal of this work is to analyze the effect of the Aprender para Valer award on the educational expectations of students from Ceará; more specifically, it analyzes the effect of a student having the right to receive the award in a given year on his expectations in the following year of attending higher education. The government of the state of Ceará instituted this award in 2009, which aims to give a notebook to the high school students that achieve adequate performance in the evaluation of Permanent System for the Evaluation of Basic Education in Ceará (SPAECE). The SPAECE databases used comprise the years 2013 and 2014, and provide information on students' proficiency in Mathematics and Portuguese, as well as students' demographic and socioeconomic attributes, besides their perception of their teachers and school. The methodology chosen was discontinuity regression with multiple scores, taking into account the eligibility criteria of the Aprender para Valer award. The results showed that students who were entitled to receive the award had risen their expectations of going to college.

Keywords: Educational expectations. Aprender para Valer. Discontinuity regression. SPAECE.

LISTA DE QUADROS

Quadro 1 – Descrição das variáveis	39
--	----

LISTA DE GRÁFICOS

Gráfico 1 – Áreas de tratamento e controle.....	26
Gráfico 2 – Histograma da Proficiência em Língua Portuguesa em 2013.....	33
Gráfico 3 – Histograma da Proficiência em Matemática em 2013.....	34

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas	28
Tabela 2 – Resultados da estimação	30
Tabela 3 – Proporções de alunos segundo o desempenho em 2013 e a expectativa em 2014	31
Tabela 4 – Proporções de cada grupo de alunos segundo o desempenho em ambas as provas do SPAECE em 2013	32
Tabela 5 – Resultados da estimação com dois bins distintos	34
Tabela 6 – Estimativas com cutoffs placebo	35
Tabela 7 – Teste de robustez	35
Tabela 8 – Proporções de cada grupo de alunos em 2014 segundo o desempenho no SPAECE em 2014	43

LISTA DE ABREVIATURAS E SIGLAS

EJA	Educação de Jovens e Adultos
ENEM	Exame Nacional do Ensino Médio
ES	Ensino Superior
IBGE	Instituto Brasileiro de Geografia e Estatística
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
OCDE	Organização para Cooperação e Desenvolvimento Econômico
SEDUC	Secretaria de Educação do Ceará
SPAECE	Sistema Permanente de Avaliação da Educação Básica do Ceará
TRI	Teoria de Resposta ao Item

SUMÁRIO

1	INTRODUÇÃO	14
2	ASPECTOS TEÓRICOS	16
3	O SPAECE E O PRÊMIO APRENDER PARA VALER	19
3.1	O SPAECE	19
3.2	Lei 14.483/09	20
4	DADOS	21
4.1	Fonte e cruzamento dos dados	21
4.2	Construção das variáveis	21
5	ESTRATÉGIA EMPÍRICA	25
6	RESULTADOS	28
6.1	Estatísticas descritivas	28
6.2	Regressão descontínua com scores múltiplos	32
6.3	Testes de validação	33
6.4	Testes de robustez	35
7	CONSIDERAÇÕES FINAIS	36
	REFERÊNCIAS	38
	APÊNDICE A – DESCRIÇÃO DAS VARIÁVEIS	40
	APÊNDICE B – DESCRITIVAS DOS DISCENTES EM 2014 DE ACORDO COM DESEMPENHO NO SPAECE EM 2014	42

1 INTRODUÇÃO

O sucesso de uma criança – considerando inclusive resultados no nível educacional - depende de três fatores: o quanto o governo (ou a sociedade) investe nas crianças, o quanto os pais investem em seus filhos e as próprias escolhas da criança. Entre os aspectos que afetam as decisões do pais e do próprio indivíduo estão suas expectativas, ou aspirações (Haveman e Wolfe, 1995).

As expectativas dos indivíduos são tema central na teoria microeconômica; contudo, os estudos na área se concentram nas expectativas financeiras e exploram questões relacionadas à poupança e investimento, por exemplo (Brown, Nuñez e Taylor, 2011). É notório quão pouco se sabe sobre a relação entre expectativa educacional e sucesso escolar. Muitos pesquisadores assumem que existe uma relação causal, apesar da pouca evidência encontrada na literatura (Jacob e Wilder, 2010).

Sem entender quais são os determinantes das expectativas, é difícil entender como estas podem afetar o desempenho dos alunos (Manski, 1993). A importância de compreender melhor as expectativas educacionais reside também no fato de que é possível propor políticas públicas de custo relativamente baixo que teriam impacto duradouro sobre o desempenho escolar, como em Jensen (2010), Oyserman (2013) e Bernard et. al. (2014).

O presente trabalho avalia o impacto do prêmio Aprender para Valer sobre as expectativas do estudantes cearenses, na tentativa de esclarecer um pouco o processo de formação das expectativas educacionais dos indivíduos. Para tanto, foi utilizado a base de dados do Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE) para os anos de 2013 e 2014. O SPAECE é uma avaliação padronizada aplicada aos alunos da rede pública do estado do Ceará. Esta base fornece não apenas a proficiência dos alunos nesta avaliação, mas também informações demográficas e socioeconômicas destes, suas expectativas em relação ao futuro acadêmico, além de medidas da percepção do estudante em relação ao próprio esforço, aos professores e à escola.

O Governo do Estado do Ceará implementou a Lei 14.483/09, segundo a qual todos os alunos do Ensino Médio que obtivessem um padrão de desempenho adequado no SPAECE teriam direito a um prêmio, a saber, um notebook. O prêmio Aprender para Valer, como é chamado, é um incentivo do governo para os alunos que participam da avaliação. Considerando isto, este trabalho avaliou o impacto de ter direito ao prêmio em 2013 sobre as expectativas dos alunos em 2014.

A metodologia escolhida foi a de regressão descontínua com scores múltiplos, pois, para ter direito ao prêmio, o aluno precisaria ter desempenho adequado nas duas avaliações que compõe o SPAECE; portanto, era possível que alguns alunos obtivessem a nota mínima em apenas um dos exames e não recebessem o prêmio. O prêmio foi visto como um indicador para o estudante de sua própria habilidade; considerou-se, ainda, que os alunos cuja nota estivesse muito próxima do ponto de corte seriam iguais em termos de habilidade acadêmica. As expectativas educacionais foram vistas em relação a uma única dimensão, a saber, a relacionada a cursar o Ensino Superior, escolha que seguiu a literatura internacional.

Esta dissertação se divide em mais seis seções. A próxima seção traz uma breve revisão da literatura; a terceira discorre um pouco sobre o SPAECE e a Lei 14.483/09; a seção seguinte apresenta os dados; a quinta seção discute a estratégia empírica, seguida pela apresentação e discussão dos resultados; enfim, são tecidas as considerações finais.

2 ASPECTOS TEÓRICOS

Esta seção se baseia principalmente no artigo *Adolescents Econometricians: How do Youth Infer the Returns to Schooling?* de Manski (1993). Este artigo trata da formação das expectativas, em particular expectativas educacionais e de ganhos ocupacionais, e dos problemas advindos da falta de informação sobre como os indivíduos formam suas expectativas.

Antes, porém, cabe definir o termo. Pode-se designar expectativa como o que a pessoa acredita que irá acontecer; de outra forma, a expectativa envolve uma estimativa do que é possível (Oyserman, 2013). O termo aspiração, embora diferente, está fortemente ligado à expectativa. As aspirações representam preferências, ou seja, o que a pessoa gostaria que acontecesse. As aspirações dos indivíduos são influenciadas por suas expectativas (Bernard et al., 2014).

Alguns estudos utilizam ambas as definições indistintamente. Este trabalho foca as expectativas dos estudantes, mas também são apresentados estudos que tratam das aspirações, uma vez que não tenha sido observada diferença entre conceitos no desenvolvimento dos estudos.

Segundo Manski (1993), os economistas, quando analisam decisões de investimento em capital humano, assumem que os jovens comparam as opções disponíveis e escolhem a melhor possível; em outras palavras, assumem que estes calculam os retornos à educação para cada etapa escolar e então escolhem a melhor opção.

O problema é que, ao conduzir estudos sobre as escolhas educacionais dos indivíduos, muitos assumem que a formação das expectativas é homogênea (que todos os jovens condicionam suas expectativas às mesmas variáveis e processam informações da mesma forma). Ao mesmo tempo, cada estudo traz variáveis independentes diferentes e hipóteses diferentes para o processamento de informações. O autor chega a citar três trabalhos que apresentam hipóteses distintas sobre como as pessoas formaram suas expectativas.

Para Manski (1993), isto traz dois problemas de identificação. Primeiro: sem saber como os jovens percebem os retornos à educação, o pesquisador não tem como inferir o processo de tomada de decisão a partir das decisões do aluno. Segundo: sem saber como esse processo ocorre, não se pode inferir os retornos à educação a partir de dados sobre os resultados.

Para o autor, parece que os jovens se deparam com o mesmo problema dos econométricos. A literatura sobre retornos à educação varia bastante em relação às variáveis independentes utilizadas, aos resultados observados e na forma de lidar com o problema de

autoseleção. Portanto, se especialistas mudam tanto a forma de inferir este problema, é possível que os jovens mudem também.

Em suma, é importante entender a formação das expectativas educacionais porque estas podem afetar as decisões e o comportamento dos indivíduos. Oyserman (2013) apontou vários resultados da literatura que relaciona as expectativas dos estudantes com o desempenho escolar e verificou que: crianças que tem a expectativa de cursar uma faculdade são mais propensas a se tornarem universitárias em relação àquelas que não tem essa expectativa; aquelas que esperam alcançar certa etapa escolar não apenas tem melhor desempenho, mas também são menos propensas a engajar em atividades delinquentes; intervenções nesta parte podem afetar não apenas o comportamento presente da criança, mas continua impactando seu desempenho escolar por mais tempo.

Considerando essa crítica, e o tempo transcorrido, alguém poderia pensar que houve grandes avanços no conhecimento sobre a formação das expectativas. De fato, trabalhos recentes como os de Nguyen (2008) e Jensen (2010) apontam fatos interessantes sobre as expectativas dos alunos.

Nguyen (2008) realizou um experimento em 640 escolas públicas de Madagascar com três objetivos: saber se os agentes estimam os retornos à educação “corretamente”, ou seja, se acertam a média dos retornos educacionais no país; entender se eles corrigem suas expectativas ao receber novas informações; e qual seria o melhor meio de fornecer essas informações. O modelo de incerteza sobre retornos à educação propõe que a estimativa dos alunos pode ser incorreta por falta de informações ou pela relação entre a própria habilidade e o retorno esperado.

Para o experimento, as escolas foram selecionadas aleatoriamente para que cada uma recebesse uma intervenção. Algumas receberam uma pessoa bem-sucedida da comunidade que compartilhou com os estudantes e seus pais sua história de superação; em outras, pais e estudantes receberam informações estatísticas sobre a renda média esperada para homens e mulheres caso os alunos completassem certas etapas escolares; e a última intervenção combinava as duas estratégias.

Os resultados mostraram que, nas escolas que receberam o tratamento com as informações estatísticas, os participantes atualizaram suas crenças, reduzindo a diferença entre o retorno médio esperado da educação para a população e o retorno médio esperado da educação para o estudante. Por consequência, as decisões de investimento em educação também mudaram – a nota média dos estudantes e a frequência à escola aumentaram significativamente.

O trabalho de Jensen (2010) é bastante semelhante ao de Nguyen (2008), porém voltado para garotos da República Dominicana. Os resultados mostram que a percepção da própria habilidade está relacionada à expectativa de ganhos futuros.

Um outro estudo que analisa a relação entre resultados acadêmicos e expectativas educacionais é o de Jacob e Wilder (2010). Estes autores analisam a evolução das expectativas dos alunos americanos desde os anos 1970 até os anos 2000. Ao investigar possíveis determinantes das expectativas dos alunos, eles verificaram que a habilidade acadêmica está positivamente correlacionada com as mesmas.

Outras hipóteses foram levantadas sobre os possíveis determinantes das aspirações e expectativas dos estudantes: o status socioeconômico parece ser importante (Jacob e Wilder, 2010; Jensen, 2010); etnia/cor da pele também podem influenciar a formação das aspirações (Pasquier-Doumer e Brandon, 2015; Vázquez e Cortina, 2018); o próprio esforço do aluno, o qual apenas ele conhece, também se mostrou correlacionado com a expectativa de sucesso escolar (Del Siegle et al., 2010). Alguns atributos que aparecem de forma recorrente na literatura sobre os determinantes do bom desempenho dos discentes, tais como características dos professores, também podem estar associados a maiores expectativas (Brown, Nuñez e Taylor, 2011).

Visto que, no Brasil, não há trabalhos publicados na área de Economia que tratem das expectativas educacionais dos estudantes, o presente trabalho se justifica ao verificar o efeito do prêmio Aprender para Valer sobre as expectativas acadêmicas de discentes cearenses. Mais especificamente, o objetivo deste trabalho é analisar se o fato de um aluno ter direito a receber o prêmio influencia sua expectativa de cursar o Ensino Superior. O prêmio foi visto como uma informação para o aluno de sua própria habilidade acadêmica. Assim, é possível extrair alguns *insights* sobre a formação das expectativas educacionais dos alunos cearenses.

3 O SPAECE E O PRÊMIO APRENDER PARA VALER

Nesta seção será apresentado um pouco do que é o SPAECE e de como é conduzido, além de detalhes sobre a Lei 14.483/09.

3.1 O SPAECE

O Sistema Permanente de Avaliação da Educação Básica do Ceará foi implementado em 1992, como uma ferramenta que permite avaliar a qualidade do ensino no estado. No início, a avaliação incluía apenas os alunos da 4ª e 8ª séries do Ensino Fundamental nas escolas de Fortaleza. A avaliação do 3º ano do Ensino Médio começou a ser feita em 2001. O exame foi universalizado em 2004, ao abranger todos os municípios cearenses e escolas da rede municipal e estadual. A partir de 2008, passou também a incluir as demais séries do Ensino Médio e alunos do 2º ano do Ensino Fundamental (SPAECE-ALFA). Desde 2010, a avaliação inclui alunos da Educação de Jovens e Adultos (EJA), permitindo o acompanhamento destes alunos com a finalidade de desenvolver intervenções adequadas para eles.

No ano de 2013, o SPAECE foi aplicado de forma censitária para alunos do 2º e 5º ano do Ensino Fundamental, 1º ano do Ensino Médio e EJA, e de forma amostral para alunos do 9º ano do Ensino Fundamental e 2º e 3º ano do Ensino Médio. Já em 2014, o teste foi censitário para as séries avaliadas no Ensino Fundamental, para o 1º do Ensino Médio e para o EJA, sendo amostral para os demais.

Tanto em 2013 quanto em 2014 as amostras foram construídas a partir dos dados do Censo Escolar do mesmo ano. Esses dados amostrais mantêm a série histórica do SPAECE. Apesar de ser uma avaliação amostral, ela conta com a participação de alunos de todas as escolas estaduais.

O SPAECE é dirigido pela Matriz de Referência para Avaliação, organizada pela Secretaria da Educação do Estado do Ceará (SEDUC), e abrange os conhecimentos de Língua Portuguesa e Matemática. As questões são formuladas utilizando a Teoria de Resposta ao Item (TRI), o que possibilita comparar resultados entre anos diferentes. A escala de proficiência do SPAECE vai de 0 a 500 pontos.

Os padrões de desempenho são categorias definidas a partir de cortes numéricos que agrupam os níveis da escala de proficiência com base nas metas educacionais estabelecidas pela avaliação. Esses cortes dão origem a quatro padrões de desempenho: Muito Crítico, Crítico, Intermediário e Adequado – que correspondem ao perfil da performance dos alunos.

Além disso, são aplicados questionários contextuais aos alunos, professores e diretores. Especificamente, o questionário do aluno permite conhecer atributos socioeconômicos deste, alguns hábitos de estudo, e também vislumbra alguns aspectos do ambiente escolar.

Neste trabalho, foram utilizados os dados de 2013 para alunos do 1º e 2º do Ensino Médio e de 2014 para alunos do 1º, 2º e 3º ano do Ensino Médio. O EJA não foi considerado para a análise.

3.2 Lei 14.483/09

Esta Lei, conhecida como Prêmio Aprender pra Valer, instituiu a premiação dos alunos das três séries do Ensino Médio. O governo do estado doaria um notebook para aqueles que conseguissem alcançar o nível de aprendizagem correspondente ao padrão de desempenho Adequado na escala de proficiência das duas provas do SPAECE¹. A nota de corte é de 325 pontos para Língua Portuguesa e 350 pontos para Matemática. Portanto, todos os alunos que obtivessem as referidas notas ou qualquer nota acima em ambas as provas teriam direito ao prêmio.

Segundo o próprio governo, a intenção do prêmio é incentivar os alunos a se dedicarem aos estudos, valorizar o ensino público, e ainda proporcionar àqueles com melhor desempenho uma ferramenta que irá facilitar seus estudos.

¹ Os alunos do 2º e 3º ano do Ensino Médio que obtivessem certa nota no Exame Nacional do Ensino Médio (ENEM) também teriam direito ao prêmio. Este caso não é analisado neste trabalho.

4 DADOS

Nesta seção, serão apresentados: a fonte dos dados, a condução do cruzamento das informações dos alunos e a construção das variáveis utilizadas.

4.1 Fonte e cruzamento dos dados

A base de dados foi cedida pela SEDUC. Os dados do SPAECE tem caráter longitudinal, por isso permitem acompanhar um mesmo aluno ao longo do tempo. No presente trabalho, o objetivo é saber se o desempenho do aluno em uma prova padronizada afeta sua expectativa de cursar o Ensino Superior.

A primeira etapa deste trabalho consistiu em cruzar as informações contextuais do aluno com sua proficiência em português e matemática dentro do mesmo ano. Como o código do aluno não muda dentro do mesmo ano, após a exclusão de casos repetidos, esta variável foi utilizada para realizar o cruzamento dessas informações em 2013, resultando em 136.667 observações. Já para o ano de 2014 só foi possível cruzar as informações do aluno pelo nome deste, após excluir casos repetidos, que resultou em 210.103 observações.

O cruzamento entre os anos foi feito pelo nome do aluno, após a exclusão de casos repetidos em ambas as bases. Devido ao caráter amostral da prova para o 2º e 3º ano, muitos estudantes que apareciam na base em 2013 não estavam na base de 2014 e vice-versa. Ao final, restaram apenas 52.815 observações com informações para os dois anos. Decidiu-se manter na base apenas os alunos que pudessem ser acompanhados nos dois anos, pois não é possível saber o que aconteceu com o aluno que não aparece na base no ano seguinte.

4.2 Construção das variáveis

O questionário contextual do SPAECE traz informações sobre as expectativas dos estudantes quanto à concluir o Ensino Médio, cursar o Ensino Superior, cursar um curso técnico e ter boas chances no mercado de trabalho. Neste trabalho, analisa-se a expectativa de cursar o Ensino Superior por questão de comparação com os resultados da literatura internacional.

No questionário contextual de 2013, a pergunta que captura a expectativa dos discentes de chegar ao Ensino Superior é a seguinte: “Com base em suas expectativas, você acredita que irá: Ingressar em uma universidade pública / Ingressar em uma universidade

particular”. O aluno deveria responder apenas uma das quatro alternativas possíveis: 1 - Sim, 2 - provavelmente sim, 3 - provavelmente não e 4 - não.

Como havia dois subitens distintos, um concernente ao ingresso numa universidade pública e outro ao ingresso numa universidade particular, ambos foram combinados em uma única variável, nomeada *Superior*, que consistia no valor mínimo das respostas dos discentes nos dois subitens, e mede a expectativa de ingressar no Ensino Superior independentemente da instituição. Em 2014 havia apenas um subitem, referente a Cursar uma faculdade, e esta variável foi nomeada como *Superior14*, assumindo que ela também mede a expectativa do aluno de entrar no Ensino Superior.

Manski (1993; 2004) faz duras críticas a essa forma de mensuração das expectativas, nomeadamente, em uma escala verbal e subjetiva. Segundo o autor, é muito melhor propor em um questionário uma escala de probabilidade onde as chances de o evento x acontecer sejam avaliadas numericamente. Porém, dada a limitação da base de dados, segue-se utilizando a escala subjetiva.

Esta escala foi recodificada, para que os valores maiores correspondessem a maior expectativa de o evento acontecer. Portanto, *Superior* e *Superior14* são variáveis categóricas que assumem os seguintes valores: 1 – Não; 2 – Provavelmente Não; 3 – Provavelmente Sim; 4 – Sim.

A proficiência do aluno é medida em uma escala de 0 a 500 pontos, e sua performance é avaliada de acordo com os padrões de desempenho. Para ser considerado como tendo um desempenho Adequado em Língua Portuguesa, o aluno precisaria obter pontuação igual ou superior a 325 pontos. Para Matemática, a nota de corte é de 350 pontos. Estas variáveis permaneceram na forma contínua, com os nomes *Proficiência LP* e *Proficiência MT*. Portanto, para cada aluno, havia quatro variáveis relacionadas a sua proficiência – duas para 2013 e duas para 2014.

A variável de tratamento, *Ganhou13*, é uma dummy valorada em 1 para os alunos que obtiveram o desempenho adequado em ambas as provas, e 0 para os demais. É necessário alertar sobre a possibilidade de que nem todos os alunos que tinham direito a receber o prêmio em 2013 o tenham realmente recebido. A base de dados considerada não contempla esta informação. No entanto, havia meios de o aluno saber qual a nota obtida na avaliação, portanto, considerou-se o efeito de ter direito a receber o prêmio.

As variáveis independentes, utilizadas na estimação, incluem atributos demográficos, *background* familiar e características da escola, dos professores e dos alunos de acordo com a percepção destes.

Algumas dessas variáveis merecem um olhar mais aprofundado. Começando pelas variáveis demográficas, o questionário contextual do SPAECE traz uma pergunta em que o aluno marcaria a opção correspondente à autoidentificação da cor da pele. As opções incluíam branco, negro, pardo, amarelo e indígena. Optou-se por colocar apenas as dummies referentes aos brancos e negros, visto que a maioria se autodeclarou parda e os autodeclarados amarelos ou índios tiveram pouca representatividade na amostra. Vale ainda mencionar que esta variável foi considerada com as respostas dadas em 2013, tendo apenas os *missings* atualizados com respostas de 2014.

O questionário perguntou também a escolaridade da mãe, considerando seis opções, que iam desde “não completou o primário” até “completou o ensino superior” e ainda uma opção que considerava se o aluno não sabia o nível escolar da mãe. Seguindo Benevides (2016), optou-se por criar uma dummy que considerasse o efeito de a mãe não ter o Ensino Fundamental completo.

Conforme Haveman e Wolfe (1995), decidiu-se considerar algum aspecto do investimento dos pais na educação dos filhos, incluindo na estimação uma variável que captura a frequência com que o aluno costuma ver os pais lendo (livros, revistas, etc.). Esta foi transformada em uma dummy onde 0 corresponde à frequência “raramente” e “nunca”, e 1 corresponde à frequência “às vezes” e “sempre”.

O questionário do SPAECE não pergunta diretamente a renda familiar do aluno; em vez disto, traz algumas opções dos bens que o aluno pode ter em casa, e uma variável que trata do recebimento do benefício Bolsa Família por alguém da casa dele. Considerando o exposto em Haveman e Wolfe (1995) e em Oyserman (2013), a fonte dos recursos econômicos pode influenciar as expectativas dos estudantes – se a fonte é um benefício do governo, isto pode ter um efeito negativo ao invés de positivo, como seria de se esperar no caso de uma renda familiar maior.

A variável *Sentimento escola* consiste na média de 7 perguntas que avaliam a percepção do aluno em relação à escola. As perguntas eram as seguintes:

- Tem participado de coisas interessantes na escola?
- Acha que vale a pena estudar nessa escola?
- Está sempre aprendendo coisas novas na escola?
- Se sente cuidado na escola?
- Se sente valorizado na escola?

- A escola é uma bagunça? Ficam muitos alunos do lado de fora da sala de aula fazendo barulho?
- Quando alguém te ameaça ou agride, não adianta reclamar na direção, pois tudo continua do mesmo jeito?

Essas perguntas permitem entender como o aluno se sente na escola. Considerando a falta de informações mais específicas sobre a escola – tais como violência e estrutura – optou-se por organizar as respostas destas perguntas em dummies onde os alunos que concordavam total ou parcialmente com as perguntas tiveram suas respostas categorizadas com valor 1 e os demais com valor 0.

Por falta de informações mais objetivas sobre a qualidade da escola, foi construída a média das notas dos alunos para cada escola, para ambas as provas. Já no caso da percepção do aluno em relação ao professor, foram criadas duas variáveis. *Metodologia* é a média de quatro perguntas referentes ao fato de o professor passar e corrigir o dever de casa, relacionar conteúdos de disciplinas diferentes e utilizar instrumentos variados na avaliação da aprendizagem. *Conduta* é a média de sete perguntas, as quais refletem se o professor consegue manter a sala disciplinada, se dá atenção a todos os alunos igualmente, se é arbitrário ao mandar alunos embora da sala de aula, se esclarece as dúvidas e se acredita que todos podem aprender. Ambas foram construídas para cada ano da amostra.

A distorção idade-série, segundo o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), acontece quando o aluno está mais de dois anos atrasado ou adiantado na escola, tendo em vista a idade “certa” em que ele deveria estar em determinada série. Nas variáveis *Distorcido* e *Distorcido14*, que capturam este efeito, mais de 90% dos alunos que se encaixam nesta definição estão atrasados; pouquíssimos estão adiantados.

Por fim, duas variáveis refletem a percepção do aluno sobre o próprio esforço e aprendizagem. A variável esforço foi construída a partir da resposta à seguinte afirmação: “Eu capricho na hora de fazer meus trabalhos”, para 2013, e “Eu faço meus trabalhos”, para 2014. Apesar da ligeira diferença entre as afirmações nos dois anos, as duas foram consideradas como medida de esforço. Os alunos que concordaram total ou parcialmente com a afirmativa foram categorizados com valor 1 e os demais com valor 0. Já a variável *Aprendizagem* foi construída a partir da resposta à afirmação “Aprendo a matéria que o professor ensina”. Estudantes que reportaram concordância total ou parcial com a afirmativa tiveram suas respostas codificadas com valor 1 e os demais com valor 0.

5 ESTRATÉGIA EMPÍRICA

Para investigar o efeito do prêmio Aprender para Valer sobre as expectativas educacionais dos alunos cearenses, utilizou-se a metodologia de regressão descontínua com scores múltiplos.

Segundo Jacob et al. (2012), a utilização do método de regressão descontínua é válida desde que a variável score não seja determinada pela variável correspondente ao tratamento; o *threshold* seja exógeno e a condição de o indivíduo ter acesso ao tratamento dependa apenas da sua posição em relação aos valores da variável score, ou seja, que não tenha como o indivíduo ser incluído no tratamento de outra forma; e que não haja outra política concomitante que afete diretamente as unidades observadas.

Os padrões de desempenho são definidos de acordo com metas educacionais estabelecidas pelo SPAECE. Não existe meio de o aluno manipular o resultado, ou seja, garantir que receberá o prêmio mesmo que sua nota esteja aquém do ponto de corte determinado. Quanto à existência de políticas concomitantes, algumas escolas participam de outro programa, às vezes mais de um, e há também premiações para as escolas de acordo com o desempenho dos alunos no SPAECE (Prêmio Escola Nota 10). Contudo, o fato de as escolas serem alvo de várias políticas não invalida o uso de regressão descontínua; apenas torna necessário certa cautela na análise dos resultados.

Para ter direito ao prêmio Aprender para Valer, não basta o aluno apresentar um padrão de desempenho adequado em uma das provas, mas em ambas. Por este motivo, é necessário generalizar o método para regressão descontínua com scores múltiplos. No presente trabalho, há duas variáveis score e dois cutoffs distintos – a nota de corte de português (a partir da qual o desempenho do aluno é considerado adequado) é 325, e a de matemática é 350.

Segundo Cattaneo, Idrobo e Titiunik (2020), para lidar com mais de uma variável score, assume-se que cada unidade da variável score é um vetor denotado por \mathbf{X}_i . Quando há duas variáveis score, o valor da unidade i é

$$\mathbf{X}_i = (X_{1i}, X_{2i}),$$

e a condição de tratamento é

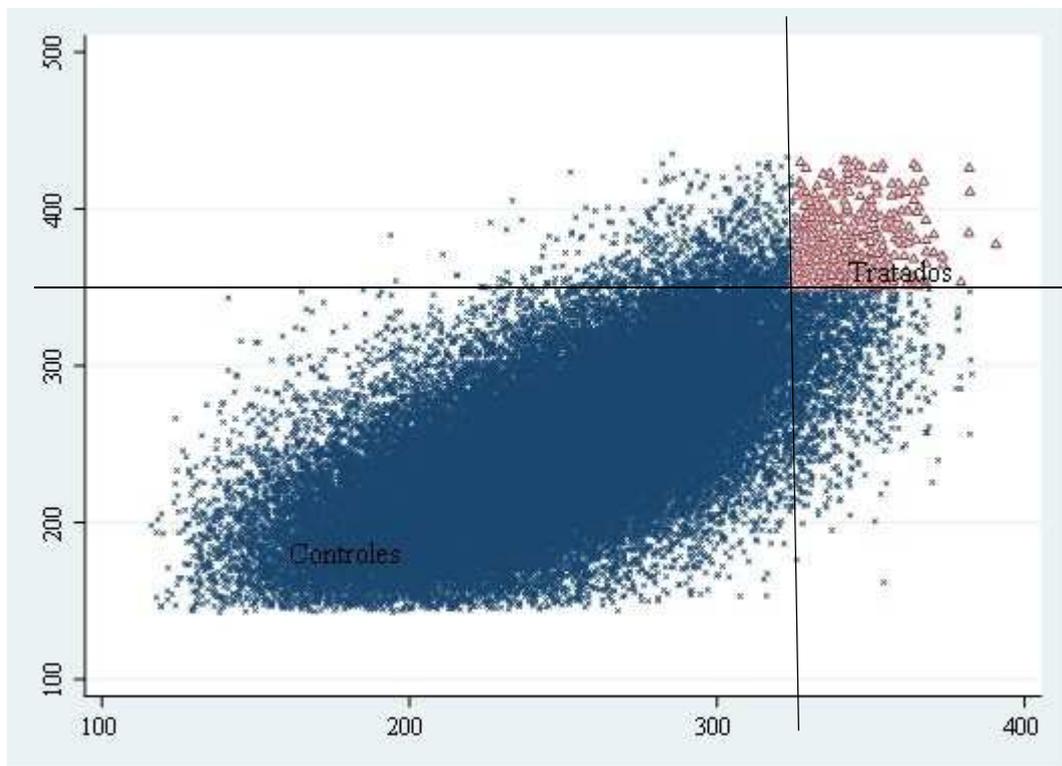
$$T_i = \mathbf{1}(X_{1i} > b_1) \cdot \mathbf{1}(X_{2i} > b_2),$$

onde b_1 e b_2 são os pontos de corte em casa uma das dimensões. Por simplicidade, assume-se que as funções de resultado potencial sejam $Y_i(1)$ e $Y_i(0)$.

O parâmetro de interesse muda, porque, ao contrário do caso da regressão descontínua simples, onde havia apenas um ponto onde a probabilidade de tratamento mudava,

agora há um conjunto de pontos onde o tratamento muda de forma descontínua. Por exemplo, um estudante que tivesse obtido uma pontuação de 450 em língua portuguesa e 349 em matemática perderia o tratamento por muito pouco; de igual modo, um estudante com uma pontuação de 350 em matemática e 324 em português também estaria muito próximo de ser um tratado, mas acaba caindo no grupo de controle. Assim, no caso de scores múltiplos a descontinuidade ocorre ao longo de uma fronteira de pontos, como ilustrado no gráfico abaixo.

Gráfico 1 - Áreas de tratamento e controle



Fonte: SPAECE 2013.

Nota: O eixo y corresponde à variável proficiência em matemática em 2013; o eixo x corresponde à variável proficiência em língua portuguesa em 2013. O grupo de tratamento é formado dos alunos que ganharam o prêmio em 2013. As linhas pretas marcam a nota de corte de cada prova.

O parâmetro de interesse no design com scores múltiplos é uma generalização do método de regressão descontínua sharp, onde o efeito médio do tratamento é calculado em alguns pontos ao longo da fronteira, ou seja, nos pontos onde a chance de tratamento muda de 0 para 1:

$$E[Y_i(1) - Y_i(0)|X_i = b], b \in \mathcal{B},$$

onde \mathcal{B} denota a fronteira que determina as áreas de tratamento e controle. Neste caso,

$$\mathcal{B} = \{(x_1, x_2): x_1 = 325 \text{ e } x_2 = 350\}.$$

A análise do caso com scores múltiplos é bem simples. O resultado de identificação é análogo ao caso com uma única variável score:

$$\tau_{SRD}(b) = \lim_{x \rightarrow b; x \in \mathcal{B}_t} E[Y_i | X_i = x] - \lim_{x \rightarrow b; x \in \mathcal{B}_c} E[Y_i | X_i = x], b \in \mathcal{B},$$

Onde \mathcal{B}_t e \mathcal{B}_c correspondem às áreas de tratamento e controle, respectivamente. Logo, conceitualmente, a metodologia de regressão descontínua com scores múltiplos gera uma família ou curva de efeito tratamento, uma para cada ponto da fronteira $b \in \mathcal{B}$.

Para a estimação com as covariadas, levou-se em conta o Teorema de Frisch Waugh. Segundo este Teorema, na regressão linear múltipla do vetor Y em dois conjuntos de variáveis, X_1 e X_2 , o subvetor β_2 é o conjunto de coeficientes obtidos quando os resíduos da regressão de Y em X_1 são regredidos no conjunto de resíduos obtidos quando cada coluna de X_2 é regredida em X_1 .

6 RESULTADOS

6.1 Estatísticas descritivas

Todas as variáveis com suas definições estão no Quadro 1 em apêndice. A tabela 1 abaixo traz as estatísticas descritivas de cada variável. Vale ressaltar que as estatísticas abaixo consideraram os discentes para os quais havia informação para a expectativa em 2014 de cursar o Ensino Superior, visto que estes seriam utilizados na estimação e as médias poderiam diferir em relação ao total da amostra.

Tabela 1 - Estatísticas descritivas

Variável	Observações	Média	Desvio-Padrão	Mínimo	Máximo
Superior	33.142	3,385	0,791	1	4
Superior14	37.655	3,244	0,988	1	4
Ganhou13	37.655	0,015	0,121	0	1
Proficiência LP 13	35.446	251,2	44,01	116,0	391,1
Proficiência MT 13	35.453	252,8	51,33	142,7	433,1
Proficiência LP 14	32.990	255,4	46,27	122,8	385,5
Proficiência MT 14	32.957	258,7	52,44	142,6	452,9
Sexo	37.655	0,512	0,500	0	1
Branco	37.655	0,166	0,372	0	1
Negro	37.655	0,114	0,317	0	1
Distorcido	37.655	0,114	0,318	0	1
Distorcido14	37.655	0,118	0,323	0	1
Mãe sem ginásio	37.655	0,404	0,491	0	1
Mãe sem ginásio 14	37.655	0,470	0,499	0	1
Bolsa família	33.240	0,717	0,450	0	1
Bolsa família 14	37.342	0,673	0,469	0	1
Pais leem	37.655	0,542	0,498	0	1
Pais leem14	37.655	0,438	0,496	0	1
Metodologia	37.655	0,750	0,338	0	1
Metodologia 14	37.655	0,728	0,355	0	1
Conduta	37.655	0,642	0,303	0	1
Conduta 14	37.655	0,722	0,200	0	1
Sentimento escola	37.655	0,765	0,232	0	1
Sentimento escola 14	37.655	0,786	0,201	0	1
Nota média escola LP 13	37.655	249,6	17,35	208,6	309,3
Nota média escola MT 13	37.655	250,9	21,72	210,4	347,3
Nota média escola LP 14	37.654	254,8	18,25	202,2	325,2
Nota média escola MT 14	37.654	258,2	22,92	188,9	373,0
Esforço	37.655	0,796	0,403	0	1
Esforço 14	37.655	0,914	0,281	0	1
Aprendizagem	37.655	0,766	0,424	0	1
Aprendizagem 14	37.655	0,882	0,322	0	1

Fonte: SPAECE 2013, 2014.

Note que a média das expectativas de cursar o Ensino Superior é bastante alta, acima de 3, numa escala que vai de 1 a 4. Ou seja, muitos jovens acreditavam que chegariam ao Ensino Superior. Porém, a média das proficiências é relativamente baixa, tanto em Língua Portuguesa quanto em Matemática, ficando próxima à nota de corte que separa os padrões de desempenho muito crítico e crítico (esta nota é 250 para Matemática e 225 para Língua Portuguesa). Cerca de 1,5% dos alunos tinham direito a prêmio Aprender para Valer em 2013.

Em relação aos atributos demográficos, 51,2% da amostra são mulheres, 16,6% se autodeclararam brancos e 11,4% negros. Em 2013, 40% das mães dos alunos não tinham Ensino Médio; em 2014 esse percentual era 47%. Em 2013, cerca de 71% dos estudantes afirmaram que eles ou alguém em sua casa recebia Bolsa família; o percentual em 2014 era de 67,3%.

Em 2013, mais da metade dos alunos afirmou que seus pais costumavam ler em casa; em 2014, 43,8% reportou que os pais tinham esse hábito. As variáveis *Metodologia*, *Conduta* e *Sentimento escola* são médias, e, de modo geral, os alunos parecem ter uma boa visão sobre os professores e o ambiente escolar.

A proficiência média das escolas também está perto dos padrões de desempenho muito crítico e crítico. Quanto à distorção idade-série, em 2013 haviam 11,4% dos estudantes nesta situação; em 2014 eram 11,8%. Apesar disso, em 2013, cerca de 79,6% dos alunos reportaram se esforçar nos deveres da escola e 76,6% disse que aprende a matéria ensinada pelo professor. Em 2014, os percentuais eram de 91,4% e 88,2%, respectivamente.

Seguindo a intuição dada pelo Gráfico 1, os dados compreendem quatro grupos de alunos: os que alcançaram desempenho adequado em ambas as provas, os que tiveram desempenho adequado em Língua Portuguesa mas não em Matemática, os que tiveram desempenho adequado em Matemática, mas não em Língua Portuguesa, e os que não alcançaram a nota de corte mínima em nenhuma das provas. É questionável que estes grupos de estudantes sejam similares em suas expectativas e características demográficas, socioeconômicas ou nos hábitos de estudo. Assim, faz-se necessário analisar algumas descritivas para cada um desses grupos. Primeiro, avaliando as expectativas em 2014 dado o desempenho em 2013, é possível observar as proporções de alunos para cada grupo na Tabela 2 abaixo.

Tabela 2: Proporções de alunos segundo o desempenho em 2013 e a expectativa em 2014

Padrão de desempenho	Expectativa de cursar o Ensino Superior				Total discentes
	Não	Provavelmente não	Provavelmente sim	Sim	
Adequado em ambas as provas	9,75%	2,48%	13,65%	74,11%	564
Adequado em Língua Portuguesa	9,54%	6,98%	24,87%	58,60%	3.164
Adequado em Matemática	10,05%	7,26%	26,26%	56,42%	2.974
Não adequado em nenhuma prova	10,65%	8,38%	28,72%	52,25%	30.953

Fonte: SPAECE 2013, 2014.

Nota: N = 37.655.

Note que, de modo geral, a diferença é maior entre o grupo de alunos com desempenho adequado em ambas as provas em relação ao que não teve desempenho adequado em nenhuma, do que entre os que tiveram desempenho adequado em pelo menos uma avaliação em relação àqueles que não o tiveram em nenhuma. Perceba também que, em relação aos outros grupos, o de alunos com desempenho adequado em ambas as provas é menos inclinado à dúvida, sendo o grupo com maior percentual de estudantes que afirmou ter certeza de que cursaria o Ensino Superior, cerca de 74% do total.

O grupo de discentes que não obteve desempenho adequado em nenhuma prova tem os maiores percentuais de alunos que consideram as hipóteses “Provavelmente não” e “Provavelmente sim”, indicando maior dúvida, além de ter o maior percentual que afirmou ter certeza de que não chegaria ao Ensino Superior.

Agora, a Tabela 3 abaixo mostra as descritivas (proporções) para cada grupo em relação aos atributos demográficos e socioeconômicos, bem como ao esforço e aprendizagem, além de considerar a distorção idade-série.

Tabela 3: Proporções de cada grupo de alunos, segundo o desempenho em ambas as provas do SPAECE de 2013.

Padrão de desempenho	Sexo	Branco	Negro	Mãe sem ginásio	Bolsa família	Distorção idade-série	Esforço	Aprendizagem
Adequado em ambas as provas	47,32	19,55	9,00	26,44	52,71	2,11	83,54	84,25
Adequado em Língua Portuguesa	54,51	17,33	11,55	38,05	68,54	10,91	78,88	75,49
Adequado em Matemática	45,73	18,09	11,00	38,71	69,76	11,02	79,55	77,25
Não adequado em nenhuma prova	50,54	16,18	11,37	41,06	72,23	13,88	78,29	75,52

Fonte: SPAECE 2013.

Nota: N=37.655.

Perceba que, novamente, é maior a diferença entre o grupo de alunos que teve desempenho adequado em ambas as provas em relação àqueles que não alcançaram o padrão adequado em nenhuma, do que entre os que tiveram desempenho adequado em pelo menos uma das provas comparado com aqueles que não o tiveram em nenhuma.

Analisando a Tabela 3 mais profundamente, nota-se que as médias dos grupos que tiveram desempenho adequado em apenas uma prova são bem próximas, exceto no atributo sexo, pois há mais mulheres com desempenho adequado em Língua Portuguesa, e mais homens com desempenho adequado em Matemática. Percebe-se também que a proporção de negros para cada grupo de alunos quase não se altera, mas que a proporção de brancos entre os que tiveram desempenho adequado nas duas avaliações é maior do que o total da amostra (19,5% em comparação com 16,6% - vide Tabela 1).

Concernente aos atributos socioeconômicos, veja que, entre os alunos que obtiveram desempenho adequado nas duas provas, cerca da metade recebia Bolsa Família, enquanto no grupo dos que não alcançaram desempenho adequado em nenhuma prova essa proporção cresce para 72,2%. E quanto à escolaridade da mãe, entre os alunos que atingiram a nota mínima em Língua Portuguesa e Matemática, 26,6% delas não tinham o ginásio; entre aqueles que não atingiram essa nota em nenhuma prova, o percentual é de 41,0%.

Dos alunos que tiveram desempenho adequado nas duas provas, apenas 2,1% se encaixavam na distorção idade-série, enquanto que entre os que não tiveram desempenho

adequado em nenhuma prova a proporção era de 13,8%, que é maior do que a média geral vista na Tabela 1.

Quanto à autopercepção de esforço, 83,5% dos alunos que alcançaram desempenho adequado nas duas avaliações afirmaram que se esforçavam nas tarefas de casa, e este número é maior do que a média geral da amostra, que era de 79,6%, vista na Tabela 1. E na questão da aprendizagem, quase 85% dos que atingiram o desempenho adequado nas duas provas afirmaram aprender a matéria que o professor ensina, enquanto nos demais grupos esse número se aproxima da média total da amostra, cerca de 76%.

No apêndice B é possível ver esta mesma análise para os dados de 2014. Em suma, é notório que o grupo de estudantes que tinha direito ao prêmio, por sua performance nas duas avaliações do SPAECE, tinha maiores expectativas do que os demais e características distintas dos demais; generalizando, melhores indicadores socioeconômicos e uma autopercepção de maiores esforço e aprendizagem, além de melhores indicadores na escola, se considerar a distorção idade-série. Esta última característica é muito importante, pois, já que não há informações sobre a nota do aluno na escola, é a melhor informação que a base de dados fornece sobre o desempenho deste ao longo da trajetória escolar.

6.2 Regressão descontínua com scores múltiplos

A estimação das regressões descontínuas com scores múltiplos foram feitas considerando um polinômio de 4ª ordem. A largura dos *bins* no caso sem as covariadas foi de 39,4, e, com as covariadas, 27,7. Os resultados aparecem na Tabela 4 abaixo, sendo o primeiro sem as covariadas e o segundo com as covariadas, a fim de trazer maior precisão para a estimativa. O primeiro valor do *cutoff* é para a proficiência em Língua Portuguesa, e o segundo para a proficiência em Matemática.

Tabela 4 - Resultados da estimação.

Cutoff	Coefficiente	P-valor	IC (95%)		h	N(h)
(325,350)	0,843	0,048	0,009	1,829	39,407	2324
(325,350)	0,057	0,493	-0,166	0,343	27,775	994

Fonte: SPAECE 2013,2014.

Nota: h corresponde à largura do *bin*. N(h) é o tamanho da amostra dentro do *bin*.

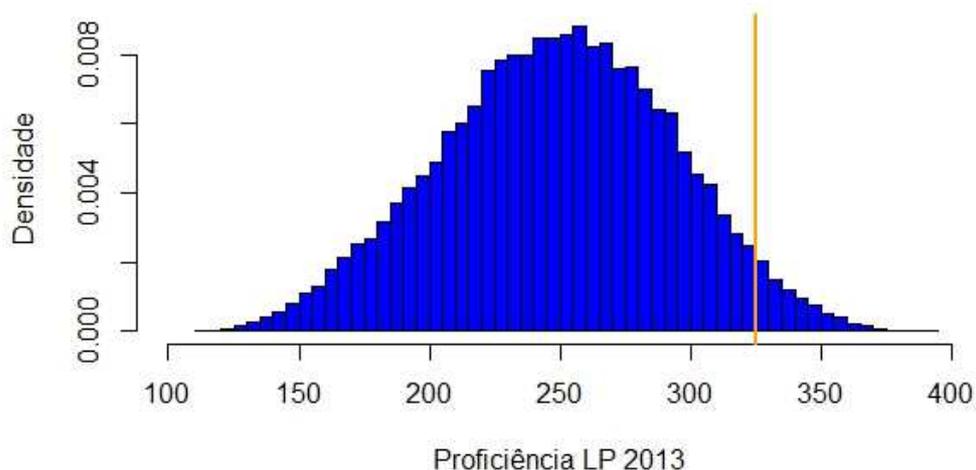
Note que, na estimação sem as covariadas, conforme esperado, o resultado é positivo e significativo: o fato de ter atingido o padrão adequado de desempenho em ambas as provas leva a um aumento médio de 0,843 nas expectativas do estudante com relação à cursar o Ensino Superior. Assim, este primeiro resultado sugere que a habilidade acadêmica

observável do aluno cearense é um dos determinantes das suas expectativas acadêmicas. Porém, com as covariadas – que aumentam a precisão da estimativa, o resultado manteve o sinal do modelo simples, mas sem significância. Ressalta-se que ainda é necessário corrigir o intervalo de confiança, fato que poderia levar o coeficiente a ser significativo.

6.3 Validação

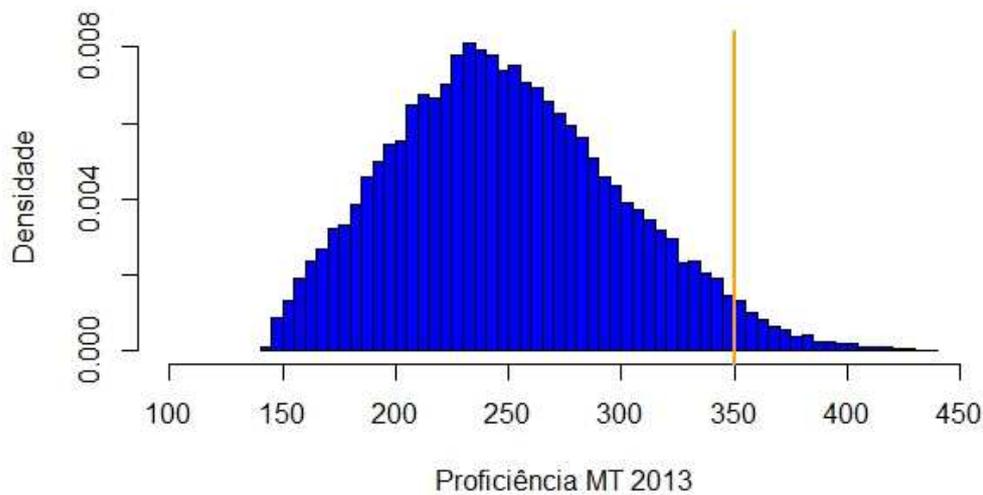
A literatura aponta alguns testes que podem ser feitos a fim de validar o uso da metodologia de regressão descontínua (Jacob et al., 2012). Primeiro, são apresentados nos gráficos 2 e 3 abaixo os histogramas das variáveis score, a fim de verificar que não há manipulação na forma como as notas foram atribuídas aos alunos. Para montar o gráfico, foram consideradas os valores mínimos e máximos atingidos em cada prova. Note que a linha laranja aponta a nota de corte mínima para obtenção do padrão de desempenho adequado cada prova. É possível observar que não há descontinuidade próximo à nota de corte, resultado favorável ao método de regressão descontínua.

Gráfico 2 - Histograma da Proficiência em Língua Portuguesa em 2013



Fonte: SPAECE 2013.

Gráfico 3 - Histograma da Proficiência em Matemática em 2013



Fonte: SPAECE 2013.

Outro teste consiste em verificar se os resultados são sensíveis a alterações no tamanho dos bins (Jacob et al., 2012; Cattaneo, Idrobo e Titiunik, 2020). Para conduzir este teste, decidiu-se repetir as estimativas considerando um *bin* para os pontos abaixo do *cutoff* e outro para os pontos acima. A Tabela 3 abaixo mostra os resultados. Note que o coeficiente se mantém significativo e a magnitude é próxima da estimativa com um único tamanho para o *bin*.

Tabela 5 - Resultados da estimação com dois *bins* distintos.

Cutoff	Coefficiente	P-valor	IC (95%)		h (abaixo)	h (acima)	N(h)
(325,350)	0,822	0,002	0,303	1,399	208,463	31,01	30236

Fonte: SPAECE 2013,2014.

Nota: h corresponde à largura do bin. N(h) é o tamanho da amostra dentro do bin.

Por fim, um teste com notas de corte placebo foi conduzido. A motivação para este teste refere ao fato de que a premissa básica do método de regressão descontínua é que as funções de regressão devem ser contínuas na ausência do tratamento. Embora não seja possível verificar esta hipótese no ponto de corte, é possível verificar se as funções são contínuas em outros pontos, pois assim não haverá dúvidas sobre a presença de outros fatores que afetem as unidades tratadas (Cattaneo, Idrobo e Titiunik, 2020). A Tabela 4 abaixo apresenta os resultados. Note que os coeficientes não são significantes, indicando que não se pode rejeitar a hipótese de que a função é contínua fora do *cutoff*.

Tabela 6 - Estimativas com *cutoffs* placebo.

Cutoff	Coefficiente	P-valor	IC (95%)		h	N(h)
(320,345)	0,998	0,183	-0,499	2,607	72,384	7323
(330,355)	-0,56	0,617	-3,375	2,002	38,281	1803

Fonte: SPAECE 2013,2014.

Nota: h corresponde à largura do *bin*. N(h) é o tamanho da amostra dentro do *bin*.

6.4 Robustez

A estimativa das proficiências em relação a expectativa se mostrou positiva e significativa, mas sua interpretação não é tão simples devido ao caráter ordinal da variável dependente. A fim de facilitar o entendimento e também aprofundar a análise, novas estimativas foram conduzidas para outra variável chamada *SuperiorSim*. Esta foi criada e codificada com valor 1 para os alunos que em 2014 disseram ter certeza de que iriam cursar o ES, e valor 0 para os demais. O resultado está apresentado na Tabela 5 abaixo.

Tabela 7 - Teste de robustez.

Variável Dependente	Cutoff	Coefficiente	P-valor	IC (95%)		h	N(h)
<i>SuperiorSim</i>	(325,350)	0,435	0,018	0,079	0,848	38,955	2279

Fonte: SPAECE 2013,2014.

Nota: h corresponde à largura do *bin*. N(h) é o tamanho da amostra dentro do *bin*.

Perceba que o coeficiente é menor do que no caso com a variável dependente *Superior14*, mas ainda é significativo e positivo. Este resultado corrobora o primeiro, de que a habilidade acadêmica tem uma relação causal positiva com a expectativa educacional.

De modo geral, as estimativas corroboraram a ideia proposta por Manski (1993) de que a habilidade acadêmica observável é um fator importante na formação das expectativas acadêmicas. É notório o efeito do prêmio Aprender para Valer, que de fato funcionou como uma fonte de informação para o aluno sobre a própria habilidade.

7 CONSIDERAÇÕES FINAIS

A presente pesquisa buscou avaliar o efeito do prêmio Aprender para Valer sobre a expectativa de o estudante cursar o Ensino Superior. Considerou-se que o prêmio poderia servir como uma medida de habilidade acadêmica observável, dando ao discente uma ideia do seu potencial acadêmico.

Para isto, foi utilizada a base de dados do SPAECE para os anos de 2013 e 2014, analisando alunos do Ensino Médio. Foi aplicado o método de regressão descontínua com scores múltiplos, onde o ponto de corte considerado estava de acordo com os critérios de elegibilidade do prêmio.

A análise descritiva mostrou que as expectativas de alunos cearenses de chegar ao Ensino Superior são bastante altas, em média, apesar de a performance na avaliação do SPAECE não ser tão alta e a maioria dos alunos se encaixar em uma padrão de desempenho crítico ou muito crítico. Foi visto também que uma porcentagem considerável se encaixa na definição do INEP de distorção idade-série, ou seja, está atrasado em relação à idade adequada para a série em que se encontra. Além disto, verificou-se que as características dos alunos que tiveram uma performance adequada nas duas provas do SPAECE são distintas dos demais, fato que levanta certa cautela quanto ao efeito do prêmio sobre as expectativas.

A análise econométrica confirmou a hipótese inicial do trabalho – o Prêmio Aprender para Valer funcionou como uma forma de informar ao aluno a sua capacidade educacional, e isso teve impacto positivo sobre a expectativa dele de cursar o Ensino Superior. Porém, ao considerar as covariadas na estimação, o resultado não foi significativo.

O trabalho tem limitações, sendo a principal a forma de medir as expectativas e a habilidade dos discentes. Esta não foi possível contornar por restrição da base de dados. Além disto, os resultados se aplicam somente aos discentes do Ceará para os anos e séries consideradas. Mesmo assim, os resultados tem aplicações importantes, principalmente em estudos que considerem decisões de investimento em capital humano, como apontado por Manski (1993).

Apesar das altas taxas de retorno da educação no Brasil, especialmente para o Ensino Superior, o percentual de brasileiros matriculados neste nível de ensino ainda é bem inferior ao de outros países da Organização para Cooperação e Desenvolvimento Econômico (OCDE) (Filho e Pessôa, 2008; OCDE, 2019). Este estudo apenas iniciou o processo de compreensão sobre um dos aspectos que podem estar afetando esses números, a saber, as

expectativas acadêmicas dos estudantes. Ainda há muito que investigar sobre como estas são formadas para que seja possível propor políticas que promovam educação superior para todos.

REFERÊNCIAS

- BARBOSA FILHO, Fernando de Holanda; PESSÔA, Samuel. Retorno da educação no Brasil. IPEA. 2008.
- BENEVIDES, Alesandra de Araújo. Avaliação do desempenho educacional no Ceará. Tese defendida no Centro de Pós-Graduação em Economia da Universidade Federal do Ceará. 2016.
- BERNARD, Tanguy et al. The future in mind: Aspirations and forward-looking behaviour in rural Ethiopia. London: Centre for Economic Policy Research, 2014.
- BROWN, Sarah; ORTIZ-NUÑEZ, Aurora; TAYLOR, Karl. What will I be when I grow up? An analysis of childhood expectations and career outcomes. *Economics of Education Review*, v. 30, n. 3, p. 493-506, 2011.
- CAMPOS VÁZQUEZ, Raymundo M.; MEDINA CORTINA, Eduardo M. Identidad social y estereotipos por color de piel. Aspiraciones y desempeño en jóvenes mexicanos. *El trimestre económico*, v. 85, n. 337, p. 53-79, 2018.
- CATTANEO, Matias D.; IDROBO, Nicolás; TITIUNIK, Rocío. A Practical Introduction to Regression Discontinuity Designs: Extensions. Cambridge University Press, 2020.
- SIEGLE, Del et al. Exploring the relationship of college freshmen honors students' effort and ability attribution, interest, and implicit theory of intelligence with perceived ability. *Gifted Child Quarterly*, v. 54, n. 2, p. 92-101, 2010.
- HAVEMAN, Robert; WOLFE, Barbara. The determinants of children's attainments: A review of methods and findings. *Journal of economic literature*, v. 33, n. 4, p. 1829-1878, 1995.
- JACOB, Brian A.; WILDER, Tamara. Educational expectations and attainment. National Bureau of Economic Research, 2010.
- JACOB, Robin et al. A Practical Guide to Regression Discontinuity. MDRC, 2012.
- JENSEN, Robert. The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, v. 125, n. 2, p. 515-548, 2010.
- MANSKI, Charles F. Adolescent econometricians: How do youth infer the returns to schooling?. In: *Studies of supply and demand in higher education*. University of Chicago Press, 1993. p. 43-60.
- MANSKI, Charles F. Measuring expectations. *Econometrica*, v. 72, n. 5, p. 1329-1376, 2004.
- NGUYEN, Trang. Information, role models and perceived returns to education: Experimental evidence from Madagascar. Unpublished manuscript, v. 6, 2008.
- OYSERMAN, Daphna. Not just any path: Implications of identity-based motivation for disparities in school outcomes. *Economics of Education Review*, v. 33, p. 179-190, 2013.

PASQUIER-DOUMER, Laure; BRANDON, Fiorella Risso. Aspiration failure: a poverty trap for indigenous children in Peru?. *World Development*, v. 72, p. 208-223, 2015.

APÊNDICE A – DESCRIÇÃO DAS VARIÁVEIS

Quadro 1 - Descrição das variáveis (continua)

Nome da variável	Descrição	Valores possíveis
Superior (14)	Expectativa do aluno de chegar ao Ensino Superior em 2013 e 2014, respectivamente	1 – Não 2 – Provavelmente não 3 – Provavelmente sim 4 - Sim
Ganhou (13)	Dummy que identifica se o aluno recebeu o prêmio em 2013	0- Não 1- Sim
Proficiência LP (14)	Proficiência do aluno em Língua Portuguesa em 2013 e 2014	Contínua (0 a 500)
Proficiência MT (14)	Proficiência do aluno em Matemática em 2013 e 2014	Contínua (0 a 500)
Sexo	Sexo do aluno	0 – Masculino 1 - Feminino
Branco	Dummy que indica se o aluno se identificou como branco	0 – Não branco 1 – Branco
Negro	Dummy que indica se o aluno se identificou como negro	0 – Não negro 1 - Negro
Distorção idade-série (14)	Avalia se o aluno está em uma série adiantada ou atrasada, de acordo com definição do INEP ² em 2013 e 2014	0 – Não há distorção idade-série 1 – Há distorção idade-série
Mãe sem ginásio (14)	Dummy que indica se a mãe do aluno não possui o Ensino Fundamental completo em 2013 e 2014	0 – Possui pelo menos o Ensino Fundamental completo 1 – Não completou o Ensino Fundamental
Leitura pais (14)	Se o aluno costuma ver os pais lendo em casa em 2013 e 2014	0 – Não 1 - Sim
Bolsa Família (14)	Se o aluno ou alguém em sua casa recebe Bolsa Família em 2013 e 2014	0 – Não 1 - Sim
Sentimento escola (14)	Como o aluno se sente em relação à escola em 2013 e 2014	De 0 a 1
Localização escola (14)	Localização da escola em 2013 e 2014	0 – urbana 1 - rural
Nota média escola LP (14)	Média da proficiência em Língua Portuguesa dos alunos de cada escola em 2013 e 2014	Contínua (0 a 500)
Nota média escola MT (14)	Média da proficiência em Matemática dos alunos de cada escola em 2013 e 2014	Contínua (0 a 500)

² Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.

Quadro 1 - Descrição das variáveis (conclusão)

Nome da variável	Descrição	Valores possíveis
Metodologia professor (14)	Como o aluno vê a metodologia do professor em sala de aula em 2013 e 2014	De 0 a 1
Conduta professor (14)	Como o aluno vê a conduta do professor em sala de aula em 2013 e 2014	De 0 a 1
Esforço (14)	Autopercepção do esforço do aluno nos trabalhos da escola em 2013 e 2014	0 – Aluno acha que não se esforça 1 – Aluno acha que se esforça
Aprendizagem (14)	Autopercepção da aprendizagem do aluno em 2013 e 2014	0 – Aluno acha que não aprende a matéria que o professor ensina 1 – Aluno acha que aprende a matéria que o professor ensina

Fonte: SPAECE 2013, 2014.

APÊNDICE B – DESCRITIVAS DOS DISCENTES EM 2014 DE ACORDO COM DESEMPENHO NO SPAECE 2014

Tabela 8 - Proporções de cada grupo de alunos, segundo o desempenho em ambas as provas do SPAECE de 2014.

Padrão de desempenho	Sexo	Branco	Negro	Mãe sem ginásio	Bolsa família	Distorção idade-série	Esforço	Aprendizagem
Adequado em ambas as provas	44,98	21,96	8,74	31,17	52,14	2,72	80,64	81,46
Adequado em Língua Portuguesa	65,37	19,94	9,63	31,44	55,96	2,98	77,35	74,65
Adequado em Matemática	34,95	18,56	8,81	33,39	59,15	4,21	76,91	76,68
Não adequado em nenhuma prova	51,48	16,40	11,61	39,90	68,50	12,73	74,91	72,24

Fonte: SPAECE 2014.

Nota: N = 37.836