

<https://doi.org/10.18222/eae.v32.6744>

IMPACTOS DA INICIAÇÃO CIENTÍFICA SOBRE O DESEMPENHO ESTUDANTIL NO ENADE 2015-2017

 FELIPE CÉSAR MARQUES^I

 MARINA SILVA DA CUNHA^{II}

I Universidade Estadual de Maringá, Maringá-PR, Brasil; felipe311.marques@gmail.com

II Universidade Estadual de Maringá, Maringá-PR, Brasil; mscunha@uem.br

RESUMO

O objetivo deste trabalho é avaliar o impacto das bolsas de iniciação científica sobre o desempenho dos alunos bolsistas do ensino presencial nas universidades federais e estaduais do Brasil. Para tanto, utiliza-se como medida de resultado a nota dos alunos concluintes no Exame Nacional de Desempenho dos Estudantes de 2015 a 2017. Devido ao problema do viés de autosseleção, foi empregada a técnica de pareamento por escore de propensão. Os resultados mostram um impacto positivo, significativo e robusto em favor dos alunos bolsistas, da ordem de 0,21 a 0,31 desvios-padrão. Tendo alunos bolsistas de outras modalidades em um segundo grupo de controle, os resultados, em geral, ainda indicam o melhor desempenho dos alunos de IC, embora as diferenças sejam reduzidas.

PALAVRAS-CHAVE AVALIAÇÃO DE PROGRAMAS • INICIAÇÃO CIENTÍFICA • EDUCAÇÃO SUPERIOR.

IMPACTOS DE LA INICIACIÓN CIENTÍFICA SOBRE EL DESEMPEÑO ESTUDANTIL EN EL ENAD 2015-2017

RESUMEN

El objetivo de este trabajo es evaluar el impacto de las becas de iniciación científica sobre el desempeño de los estudiantes becarios de la educación presencial en las universidades federales y provinciales de Brasil. Para ello se utiliza como medida de resultado la nota de los alumnos que concluyeron el Exame Nacional de Desempenho dos Estudantes, del 2015 al 2017. Debido al problema del sesgo de autoselección, se utilizó la técnica de apareamiento por escore de propensión. Los resultados muestran un impacto positivo, significativo y robusto en favor de los estudiantes becarios, del orden del 0,21 al 0,31 desviaciones estándar. Con la presencia de estudiantes becarios de otras modalidades en un segundo grupo de control, en general los resultados todavía indican el mejor desempeño de los estudiantes de IC, aunque las diferencias sean reducidas.

PALABRAS CLAVE EVALUACIÓN DEL PROGRAMAS • INICIACIÓN CIENTÍFICA • EDUCACIÓN SUPERIOR.

IMPACTS OF UNDERGRADUATE RESEARCH SCHOLARSHIPS ON STUDENT PERFORMANCE ON THE ENADE 2015-2017

ABSTRACT

The aim of this paper is to assess the impact of undergraduate research scholarships on the performance of scholarship students enrolled in face-to-face courses at Brazilian federal and state universities. The senior students' test scores on the specific component test of the *Exame Nacional de Desempenho dos Estudantes* for the years 2015 to 2017 were used as an outcome measure. To deal with the self-selection bias, the propensity score matching technique was applied. The results show a positive, significant and robust impact in favor of scholarship students, on the order of 0.21 to 0.31 standard deviations, depending on the matching algorithm and knowledge area considered. Assessing the performance of other students with different kinds of scholarships in a second control group, the results, in general, still indicate the better performance of undergraduate students with research scholarships, although the differences are smaller.

KEYWORDS PROGRAM EVALUATION • UNDERGRADUATE SCIENTIFIC RESEARCH • HIGHER EDUCATION.

INTRODUÇÃO

Segundo Massi e Queiroz (2015, p. 7), a iniciação científica (IC) é definida como “o desenvolvimento de um projeto de pesquisa elaborado e realizado sob orientação de um docente da universidade, executado com ou sem bolsa para os alunos”. Partindo do conceito da palavra “iniciação”, as autoras propõem definir alternativamente a IC como o “processo no qual é fornecido o conjunto de conhecimentos indispensáveis para iniciar o jovem nos ritos, técnicas e tradições da ciência” (MASSI; QUEIROZ, 2010, p. 174).

Os estudos, em geral, apontam para o êxito do Programa Institucional de Bolsas de Iniciação Científica (Pibic). Para Massi e Queiroz (2010), as pesquisas realizadas demonstram os efeitos positivos da IC no desenvolvimento pessoal do aluno em áreas como desempenho na graduação, transição e sucesso na pós-graduação, socialização, autonomia, responsabilidade e maturidade, bem como os resultados positivos para as universidades, por meio do fortalecimento da pesquisa nas instituições.

Não obstante a disseminação da IC e seu êxito em estabelecer uma ponte entre a graduação e a pós-graduação, alguns autores ressaltam a pequena quantidade de estudos realizados sobre essa temática no país (MASSI; QUEIROZ, 2015). Apesar do crescimento recente do número de pesquisas, verificado por Nogueira e Canaan (2009) e Canaan (2012), os autores têm indicado a necessidade de diversificação dos procedimentos metodológicos na área, tendo em vista a “notável semelhança entre as questões de pesquisa e as metodologias adotadas para coleta e análise dos dados em vários trabalhos” (MASSI; QUEIROZ, 2010, p. 193). Os trabalhos, em geral, têm-se focado em questionários e entrevistas com alunos bolsistas de IC, ex-bolsistas e professores orientadores.

Poucos estudos possuem como proposta uma abordagem quantitativa de inferência estatística, para além da estatística descritiva. De fato, os únicos trabalhos levantados que se propõem a esse tipo de análise foram os de Canaan e Nogueira (2015) e do Centro de Gestão e Estudos Estratégicos – CGEE (2017). Interpretando as bolsas de IC como bem educacional escasso, Canaan e Nogueira (2015) analisam, por meio do modelo de regressão logística, quais aspectos da herança cultural dos alunos favorecem o acesso às bolsas.

Já o estudo do CGEE (2017) busca avaliar, a partir do pareamento por escore de propensão combinado com técnicas de regressão logística e de mínimos quadrados ordinários, quais os impactos da participação no Pibic em diferentes aspectos profissionais dos alunos graduados, como a atuação na pós-graduação, inserção no mercado de trabalho e remuneração, embora o estudo não explore a relação antecedente entre IC e o sucesso ainda durante a graduação.

O presente estudo busca, portanto, contribuir com a literatura acerca da IC a partir da abordagem quantitativa pouco explorada, investigando o impacto das

bolsas de IC sobre o desempenho acadêmico dos alunos bolsistas, por meio da técnica de pareamento por escore de propensão. Para tanto, considera-se a nota de componente específico dos alunos concluintes no Exame Nacional de Desempenho dos Estudantes (Enade) de 2015 a 2017, contemplando assim 65 cursos de graduação em licenciatura e bacharelado presenciais, ofertados em 150 instituições de ensino superior federais e estaduais. Além da análise global, envolvendo todos os alunos, as avaliações são realizadas também de acordo com as oito grandes áreas do conhecimento definidas pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Complementar à avaliação principal, que considera alunos bolsistas de IC como grupo de tratamento e alunos que não receberam nenhum tipo de bolsa durante a graduação como grupo de controle, uma análise adicional é conduzida, estabelecendo como novo grupo de controle alunos bolsistas de outras modalidades além da IC, no sentido de buscar a obtenção de robustez nos resultados e a possibilidade de separação dos efeitos financeiros da bolsa sobre os impactos do programa de IC.

Quanto ao método empregado, o pareamento por escore de propensão tem sido utilizado na literatura para avaliar o impacto de diferentes políticas voltadas para o ensino superior, por meio da avaliação de desempenho no Enade, como o Programa Universidade para Todos (Prouni) (LÉPINE, 2016; DUTRA, 2016), Ciência sem Fronteiras (CONCEIÇÃO; FRANÇA; JACINTO, 2016) e as políticas de cotas (PEREIRA; BITTENCOURT; SILVA JUNIOR, 2013), o que justifica sua escolha como ferramenta de análise na presente pesquisa.

Este estudo divide-se em quatro seções, além desta introdução. A seguir, apresenta-se o histórico, os objetivos e os resultados do Programa Institucional de Bolsas de Iniciação Científica. Posteriormente, é abordada a metodologia, mostrando o problema do viés de autosseleção subjacente ao estudo das avaliações de impacto, juntamente com a proposta do pareamento por escore de propensão para lidar com este viés, além de detalhes a respeito do tratamento dos dados empregado no trabalho. Os resultados obtidos são analisados para o caso geral e desagregados segundo as grandes áreas do conhecimento. Também são comparados os resultados com os da análise complementar, relativa ao segundo grupo de controle. Por fim, apresentam-se as considerações finais do estudo.

HISTÓRICO E OBJETIVOS DA INICIAÇÃO CIENTÍFICA NO BRASIL

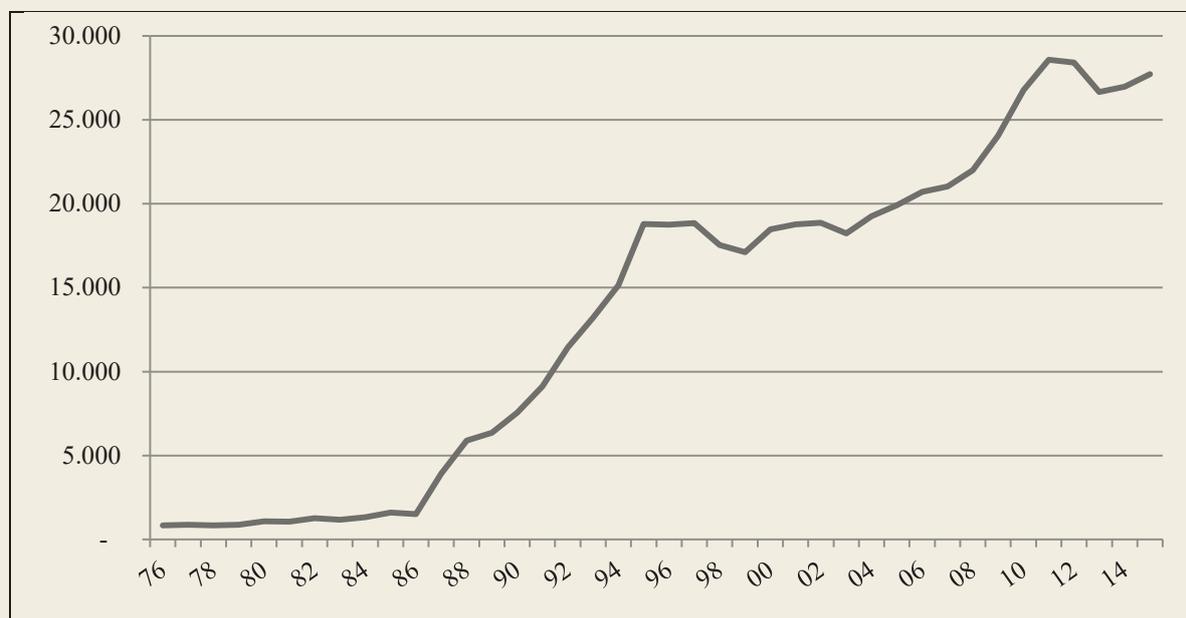
A pesquisa desenvolvida em universidades com o auxílio de estudantes de graduação já existia de forma incipiente no Brasil desde a década de 1940 (BARIANI, 1998). Um importante marco no desenvolvimento dessas atividades foi a criação do CNPq, em

1951, estabelecendo a IC enquanto objeto de política nacional, que passou a ser organizada, promovida e financiada pelo governo a partir de 1963 (MASSI; QUEIROZ, 2010; 2015).

Em 1968 ocorreu a Reforma Universitária, que teve, entre suas consequências, a consolidação e a ampliação dessas atividades, a partir do reconhecimento da indissociabilidade do ensino-pesquisa na educação superior (MARTINS, 2009; MALDONADO, 1998).

A evolução do número de bolsas de IC ofertadas no país no período de 1976 a 2015 pode ser observada no Gráfico 1. Em 1976 havia apenas 845 bolsas no país e, até meados dos anos 1980, esse número apresentou crescimento modesto, passando para 1.600 bolsas em 1985. Porém, na segunda metade dessa década, houve um aumento mais expressivo, atingindo 7.548 bolsas em 1990. Destaca-se nesse período a criação, pelo CNPq, em 1988, do Programa Institucional de Bolsas de Iniciação Científica, um instrumento de fomento à IC, a partir do qual as bolsas passaram a ser concedidas diretamente às instituições de ensino superior e aos institutos de pesquisa, que se tornaram responsáveis pelo seu gerenciamento e concessão (MASSI; QUEIROZ, 2010).

GRÁFICO 1 - Número de bolsas de iniciação científica ofertadas pelo CNPq, 1976-2015



Fonte: CNPq (s.d.).

Graças ao Pibic, a primeira metade da década de 1990 foi um período de forte expansão na concessão de bolsas de IC, com um crescimento de 149%, alcançando 18.790 bolsas em 1995. Essa fase foi considerada por Martins e Martins (1999) como o “período da IC”. Ademais, para Marcuschi (1996), o Pibic representa a novidade mais importante do ensino superior dessa época.

No período seguinte, observou-se estagnação no número de bolsas de IC, não superando 20 mil bolsas até 2005. Todavia, verifica-se novamente um crescimento contínuo das bolsas de iniciação científica de 2003 até 2011, quando se atingiu o maior nível nos anos analisados, com 28.580 bolsas concedidas. Porém, no final do período, o total de bolsas voltou a decrescer, somando 27.717 em 2015. Outros números ajudam a confirmar a importância assumida pela IC na esfera governamental. Segundo dados do CNPq (s.d.), 30% do total de bolsas disponibilizadas pela instituição no país foi destinado à IC em 2015. Já em termos monetários, as bolsas de IC correspondem a 10% do valor total investido em bolsas no mesmo ano.

O sucesso do Pibic levou as fundações estaduais de amparo à pesquisa a adotarem programas similares de incentivo à iniciação científica, como são os casos da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig), de Minas Gerais, da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp), de São Paulo, e da Fundação Araucária, do Paraná (NOGUEIRA; CANAAN, 2009).

Atualmente, as normas e o funcionamento da IC e do Pibic são regidos por meio da Resolução Normativa n. 17/2006 (RN-017/2006), do CNPq. Segundo a resolução, a finalidade da IC é “despertar vocação científica e incentivar talentos potenciais entre estudantes de graduação universitária, mediante participação em projeto de pesquisa, orientados por pesquisador qualificado” (CNPQ, 2006, não paginado).

Ainda segundo a RN-017/2006 (CNPQ, 2006), o Pibic possui três objetivos gerais: contribuir para a formação de recursos humanos para a pesquisa; contribuir para a formação científica de recursos humanos que se dedicarão a qualquer atividade profissional; e contribuir para reduzir o tempo médio de permanência dos alunos na pós-graduação.

Diferentes estudos, predominantemente qualitativos, vêm atestando a importância e o sucesso do Programa em atingir seus objetivos. Para Marcuschi (1996), a IC representa a possibilidade de colocar o aluno em contato desde cedo com a pesquisa e, dessa forma, engajá-lo. A IC permite a ampliação da experiência acadêmica de jovens alunos, proporcionando a eles mais chances de seguirem carreiras que envolvam a pesquisa (CGEE, 2017).

Segundo o estudo de Aragón, Martins e Velloso (1999), um bolsista Pibic possuía 6 vezes mais chances de ingressar no mestrado comparado aos não bolsistas. Além disso, o prazo médio de transição entre a graduação e o ingresso no mestrado era de 1,2 ano para ex-bolsistas e de 6,8 anos entre os não bolsistas. Cabrero (2007) constatou que, entre os alunos participantes do Pibic na Universidade Federal de São Carlos (UFSCar), 51% haviam se titulado mestres ou doutores. Estimativas mais recentes realizadas pelo CGEE (2017) mostram que os ex-alunos Pibic da Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp) têm 2,2 vezes mais chances de completarem o mestrado, enquanto a chance de concluir o doutorado é 1,5 vez maior.

Participantes da IC apresentam melhor rendimento na graduação (BREGLIA, 2002; PIRES, 2002; BRIDI, 2004) e menor tendência à evasão (VILLAS BOAS, 2003; AGUIAR, 1997). Costa *et al.* (1999) e Mazzafera e Suguimoto (2016) constataram também que, uma vez na pós-graduação, alunos que participaram da iniciação científica produzem mais trabalhos científicos do que seus pares.

Pesquisas revelam ainda o desenvolvimento pessoal adquirido com a experiência da IC, relacionado ao desenvolvimento do pensamento crítico, autonomia, responsabilidade e maturidade (CABRERO, 2007; CALAZANS, 1999).

Um questionário aplicado pelo CNPq aos participantes do Pibic no período 2013-2014, abordando os benefícios trazidos pela iniciação científica, revela que as experiências vivenciadas pelos alunos são, de modo geral, bastante satisfatórias. Enquanto apenas 1,5% dos respondentes consideraram que a IC não contribuiu para a sua formação, 75,3% afirmaram que a IC despertou seu interesse pela pesquisa e 78,1% mencionaram que o programa aprofundou seus conhecimentos na área estudada. Ainda neste questionário, 68% dos bolsistas indicaram que pretendem cursar a pós-graduação.

Por fim, Massi e Queiroz (2015) destacam também a importância da existência de um programa institucional de IC de alcance nacional, configurando-se um diferencial brasileiro, uma vez que, na maior parte dos países americanos e europeus, tal atividade está atrelada à iniciativa individual dos docentes.

METODOLOGIA

Avaliação de impacto

O maior desafio no estudo das avaliações de impacto é encontrar o contrafactual – determinar o que ocorreria com o beneficiário do programa avaliado caso o programa não tivesse ocorrido. No presente estudo, trata-se de determinar o resultado no Enade dos alunos bolsistas de IC, caso eles optassem por não participar da iniciação científica. Devido à ausência de dados reais que pudessem mostrar, simultaneamente, o desempenho dos mesmos alunos com e sem a bolsa de IC, torna-se necessário criar um grupo de comparação similar ao grupo dos tratados, denominação atribuída aos indivíduos beneficiados pelo programa (KHANDKER; KOOLWAL; SAMAD, 2009).

Os questionários do Enade nas provas de 2015 a 2017, nos quais se baseia a presente pesquisa, perguntam aos alunos somente se eles foram beneficiados com bolsas de IC durante a graduação, não havendo informações a respeito da realização da IC sem o recebimento de bolsa. Dessa forma, diante da limitação dos dados disponíveis, uma análise semelhante não pode ser feita em relação a alunos de IC não bolsistas.

Considerando D_i uma variável binária que atribui valor 1 ao i -ésimo indivíduo caso ele seja beneficiário do programa, a variável de resultado é definida por $Y_i(D_i)$. O efeito do tratamento sobre o indivíduo i é dado então pela equação (1):

$$\tau_i = Y_i(1) - Y_i(0) \quad (1)$$

Surge assim o problema, pois, para cada indivíduo, apenas um dos possíveis resultados é observado, impossibilitando a estimativa de τ_i (o efeito do tratamento) para cada indivíduo. As avaliações devem se concentrar, então, nos efeitos médios do tratamento sobre a população (CALIENDO; KOPEINIG, 2008).

Tomando a população como objeto de análise, o parâmetro τ pode ser calculado de duas formas distintas, que são diferenciadas pelo enfoque recebido entre os grupos populacionais avaliados. A primeira delas é definida como efeito médio do tratamento ou *average treatment effect* (ATE):

$$\tau_{ATE} = E [Y(1) - Y(0)] \quad (2)$$

O ATE é calculado como a esperança da diferença da variável de resultado antes e após o tratamento. De acordo com Caliendo e Kopeinig (2008), este parâmetro mostra o valor esperado do resultado, caso os indivíduos da população recebessem o tratamento aleatoriamente. No contexto deste trabalho, a equação (2) mostra qual seria a melhoria esperada por um aluno na prova do Enade que recebeu aleatoriamente uma bolsa de iniciação científica durante a graduação.

A segunda maneira de se calcular τ tem como foco somente a parcela da população beneficiada pelo programa, recebendo o nome de efeito médio do tratamento sobre os tratados ou *average treatment effect on treated* (ATT). O ATT é calculado como a diferença dos valores esperados de $Y(0)$ e $Y(1)$ entre os indivíduos que, de fato, receberam o tratamento, como mostra a equação (3). O ATT corresponde, então, ao efeito esperado da IC para os alunos contemplados com a bolsa.

$$\tau_{ATT} = E [Y(1) | D=1] - E [Y(0) | D=1] \quad (3)$$

Considerando que a RN-017/2006 estabelece como finalidade da IC “despertar vocação científica e incentivar talentos potenciais entre estudantes de graduação universitária”, enquanto o manual do usuário Pibic (1998) define a IC como “um instrumento que permite introduzir os estudantes de graduação, potencialmente mais promissores, na pesquisa científica”, fica claro que as bolsas de IC não são distribuídas aleatoriamente, mas possuem um público-alvo definido, de tal modo que o parâmetro de interesse no restante do estudo será o ATT.

Nesse caso, portanto, o contrafactual, não observado, é o segundo termo do lado direito da equação (3). Utilizando, como alternativa, o valor observado do resultado dos não tratados $E [Y(0) | D=0]$, tem-se:

$$E [Y(1) | D=1] - E [Y(0) | D=0] = \tau_{\text{ATT}} + \{E [Y(0) | D=1] - E [Y(0) | D=0]\} \quad (4)$$

O parâmetro ATT será corretamente identificado somente se o segundo termo do lado direito na equação (4), denominado viés de autosseleção, igualar-se a zero. Faz-se necessário, portanto, estabelecer duas hipóteses para solucionar o problema do viés de autosseleção (CALIENDO; KOPEINIG, 2008).

A hipótese 1 – independência condicional: $[Y(0), Y(1)] \perp T | X$, onde \perp denota independência – afirma que, dado um conjunto de covariadas X , não afetadas pelo tratamento, a distribuição dos resultados potenciais de tratados e não tratados é independente do tratamento. Essa hipótese assume que, controlando as características observáveis dos indivíduos, qualquer diferença verificada nos resultados pode ser atribuída à participação no programa (LI; PALMA; XU, 2017). Embora não seja uma hipótese diretamente testável, a existência de um banco de dados completo contribui para a sua sustentação, ao permitir o controle de um grande número de variáveis que podem afetar a participação no programa (KHANDKER; KOOLWAL; SAMAD, 2009).

Uma hipótese mais fraca, suficiente para o cálculo do efeito médio sobre os tratados, estabelece que, dado o conjunto de covariadas, apenas o resultado potencial do grupo dos não tratados deve ser independente da variável de tratamento: $Y(0) \perp T | X$.

A hipótese 2 – suporte comum ou sobreposição: $0 < P(D = 1 | X) < 1$ – garante que indivíduos com as mesmas características observáveis tenham, simultaneamente, probabilidade positiva de participar e não participar do programa, garantindo a existência de observações comparáveis a todos os indivíduos (HECKMAN; LALONDE; SMITH, 1999). Novamente, para o cálculo da ATT, pode-se assumir a hipótese da sobreposição fraca: $P(D = 1 | X) < 1$.

Pareamento por escore de propensão

Rosenbaum e Rubin (1983) mostram que, satisfazendo as hipóteses de independência condicional e sobreposição, tratados e não tratados com características semelhantes podem ser pareados por meio da probabilidade de pertencer ao grupo de tratamento (escore de propensão, $P(X)$). Assim, o ATT é calculado por meio da diferença média da comparação dos resultados entre os indivíduos tratados e não tratados, como mostra a equação (5). Esta técnica é conhecida como pareamento por escore de propensão ou *propensity score matching* (PSM).

$$\tau_{\text{ATT}} = E_{P(X) | D=1} \{E[(Y(1) | D=1, P(X)) - E[(Y(0) | D=0, P(X))]\} \quad (5)$$

Para a execução do PSM, deve-se primeiramente calcular o escore de propensão de cada indivíduo. Para tanto, pode-se recorrer às técnicas econométricas tradicionais de resposta binária, como os modelos *logit* ou *probit*. Devem ser incluídas no modelo tanto variáveis que influenciem a participação no programa quanto a variável de resultado (KHANDKER; KOOLWAL; SAMAD, 2009).

O próximo passo envolve selecionar o método de pareamento a ser utilizado. Os algoritmos de pareamento diferenciam-se em relação não somente à forma como é definida a vizinhança de comparação dos tratados, mas também aos pesos que devem ser atribuídos a cada um desses vizinhos (CALIENDO; KOPEINIG, 2018).

O algoritmo de pareamento mais objetivo é o de *Nearest Neighbour matching* (NN). Nele, cada indivíduo do grupo de tratamento é comparado ao indivíduo do grupo de controle com o escore de propensão mais próximo. Uma variação desse algoritmo consiste em comparar o indivíduo tratado com os n indivíduos com os escores de propensão mais próximos do grupo de controle. No presente trabalho, considera-se o *Nearest Neighbour* com $n = 1$ e $n = 5$. Nesses casos, os erros-padrão do ATT serão calculados por meio de *bootstrap*, com 200 replicações.

Outros algoritmos adotados no trabalho são o *radius matching*, em que os indivíduos do grupo de tratamento são comparados com todos os indivíduos do grupo de controle cuja proximidade do escore de propensão esteja dentro de um valor predefinido, neste trabalho com uma tolerância de 0,05; e o *kernel matching*, que cria, para cada indivíduo tratado, um contrafactual que engloba todo o grupo de tratamento, ponderando-o segundo uma função de *kernel*. Considerando a maior demanda computacional exigida por esses métodos e o grande número de alunos avaliados, o *bootstrap* é substituído pela aproximação de erros-padrão proposta por Lechner (2001). Salienta-se que Lechner (2002) encontra poucas diferenças entre essas aproximações e os erros-padrão calculados via *bootstrap*.

Tratamento dos dados

Utilizam-se neste estudo os microdados do Enade de 2015, 2016 e 2017. Considerando-se que as provas para cada curso são aplicadas trienalmente, o estudo engloba todos os cursos de graduação do tipo bacharelado e licenciatura atualmente avaliados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Nessas três avaliações, a prova não foi aplicada aos alunos ingressantes. Consequentemente, apenas alunos concluintes fazem parte da amostra do estudo.

Os cursos de formação de tecnólogo não são avaliados. Também não são consideradas as instituições de ensino superior municipais e privadas, bem como

curso de graduação a distância, devido à baixa oferta de bolsas de IC nessas modalidades. São avaliados, portanto, cursos presenciais de instituições públicas federais e estaduais.

Foram descartados os dados de alunos com provas em branco ou com nota zero nas provas de formação geral ou de conhecimentos específicos, evitando assim a comparação de alunos cujo esforço na prova foi muito baixo. Também foram desconsiderados os dados de alunos que não responderam a todas as questões necessárias dos questionários socioeconômico e de percepção da prova. Por fim, excluíram-se os dados das instituições em que nenhum dos alunos considerados foram beneficiados com bolsas de iniciação científica.

O grupo de tratamento é formado pelos alunos concluintes que, durante a graduação, foram beneficiados com bolsas de IC. Esta informação é obtida por meio do questionário socioeconômico aplicado aos estudantes. Infelizmente, não é possível identificar no grupo de tratamento alunos que tenham feito iniciação científica sem o recebimento das bolsas.

São estabelecidos dois grupos de controle. No primeiro, consideram-se os alunos concluintes das instituições federais e estaduais de ensino superior que não receberam nenhum tipo de bolsa acadêmica durante a graduação. O segundo grupo de controle é formado por alunos que, ao longo da graduação, foram beneficiados com bolsas acadêmicas de outras modalidades além da IC (extensão, monitoria/tutoria e programa de educação tutorial – PET), possibilitando, dessa forma, testar a robustez do método empregado e a desvinculação da bolsa de IC sobre o impacto da iniciação científica, como argumentado por Pires (2015).

A variável de resultado refere-se à nota de componente específico obtida pelo aluno concluinte na prova do Enade. A descrição detalhada das variáveis de controle empregadas é apresentada no Quadro 1. Já as variáveis binárias de referência, omitidas do modelo, são: Sudeste; pais com, no máximo, ensino fundamental I completo; alunos que moram sozinhos ou em outros tipos de moradia; renda familiar de até 1,5 salário mínimo; alunos que não trabalham; e curso da área de exatas e da terra.

Para que diferentes cursos pudessem ser avaliados conjuntamente, as notas dos alunos foram padronizadas, de tal forma que os resultados são avaliados em termos do número de desvios-padrão em relação aos demais alunos do mesmo curso. A avaliação de impacto das bolsas de IC sobre os bolsistas (ATT) é feita por meio do PSM, utilizando os algoritmos de pareamento discutidos anteriormente. Todos os cálculos do trabalho foram realizados com o *software* Stata 15.

QUADRO 1 - Descrição das variáveis

VARIÁVEL	DESCRIÇÃO
Nota de CE	Nota do componente específico do aluno no Enade padronizada
Nota de FG	Nota de formação geral do aluno no Enade padronizada
Idade	Idade do aluno no ano de realização da prova
Federal	Variável binária igual a 1 se o aluno estudou em uma instituição federal
Norte	Variável binária igual a 1 se o curso é na região Norte
Nordeste	Variável binária igual a 1 se o curso é na região Nordeste
Sul	Variável binária igual a 1 se o curso é na região Sul
Centro-Oeste	Variável binária igual a 1 se o curso é na região Centro-Oeste
Mulher	Variável binária igual a 1 se é do sexo feminino
Achou fácil	Variável binária igual a 1 se achou a prova fácil/muito fácil
Minoria étnica	Variável binária igual a 1 se o aluno se declara negro, pardo ou indígena
Pais ens. 9º fund.	Variável binária igual a 1 se nenhum dos pais tem além do ens. fund. II
Pais ens. méd.	Variável binária igual a 1 se nenhum dos pais tem além do ensino médio
Pais ens. sup.	Variável binária igual a 1 se ao menos um dos pais tem ensino superior
Mora com pais	Variável binária igual a 1 se o aluno mora com os pais e/ou parentes
Mora com cônjuge	Variável binária igual a 1 se o aluno mora com o cônjuge e/ou filho(s)
1,5 a 3 SM	Variável binária igual a 1 se a renda familiar do aluno está entre 1,5 e 3 SM
3 a 4,5 SM	Variável binária igual a 1 se a renda familiar do aluno está entre 3 e 4,5 SM
4,5 a 6 SM	Variável binária igual a 1 se a renda familiar do aluno está entre 4,5 e 6 SM
6 a 10 SM	Variável binária igual a 1 se a renda familiar do aluno está entre 6 e 10 SM
10 a 30 SM	Variável binária igual a 1 se a renda familiar do aluno está entre 10 e 30 SM
Mais de 30 SM	Variável binária igual a 1 se a renda familiar do aluno é superior a 30 SM
Trab. menos 40h	Variável binária igual a 1 se o aluno trabalha menos de 40 horas semanais
Trab. mais 40h	Variável binária igual a 1 se o aluno trabalha 40 horas semanais ou mais
Pol. afirmativas	Variável binária igual a 1 se o aluno ingressou via políticas de ação afirmativa ou inclusão social
Ens. med. pub.	Variável binária igual a 1 se estudou o ensino médio em escola pública
Leu 6 livros	Variável binária igual a 1 se leu seis ou mais livros no ano
Estuda mais de 8h	Variável binária igual a 1 se o aluno se dedicou mais de 8h semanais aos estudos
Boa relação prof.-aluno	Em uma escala crescente de 1 até 6, o aluno respondeu 5 ou 6 quanto às relações professor-aluno ao longo do curso
Domínio do prof.	Em uma escala crescente de 1 até 6, o aluno respondeu 5 ou 6 quanto ao domínio dos professores do curso
Boa infraestrutura	Em uma escala crescente de 1 até 6, o aluno respondeu 5 ou 6 quanto à infraestrutura das salas de aula da instituição
Biológicas	Variável binária igual a 1 se o curso é na área de biológicas
Engenharias	Variável binária igual a 1 se o curso é na área de engenharias
Saúde	Variável binária igual a 1 se o curso é na área de saúde
Agrárias	Variável binária igual a 1 se o curso é na área de agrárias
Sociais aplicadas	Variável binária igual a 1 se o curso é na área de sociais aplicadas
Humanas	Variável binária igual a 1 se o curso é na área de humanas
Ling., letras e artes	Variável binária igual a 1 se o curso é na área de linguística, letras e artes

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

RESULTADOS

Análise preliminar

No total, são avaliados 129.450 alunos de 65 cursos diferentes nos anos de 2015 a 2017, ofertados em 150 instituições, sendo 46 estaduais e 104 federais, selecionadas de acordo com os critérios definidos na seção anterior. Do total de alunos, 30,2% foram contemplados com uma bolsa de IC durante a graduação. A divisão dos alunos em tratados e não tratados, segundo a área de conhecimento avaliada, é apresentada na Tabela 1. A lista completa dos cursos avaliados encontra-se no Apêndice.

TABELA 1 - Alunos tratados (bolsistas de IC) e não tratados (não bolsistas), segundo a área de conhecimento avaliada no Enade de 2015 a 2017

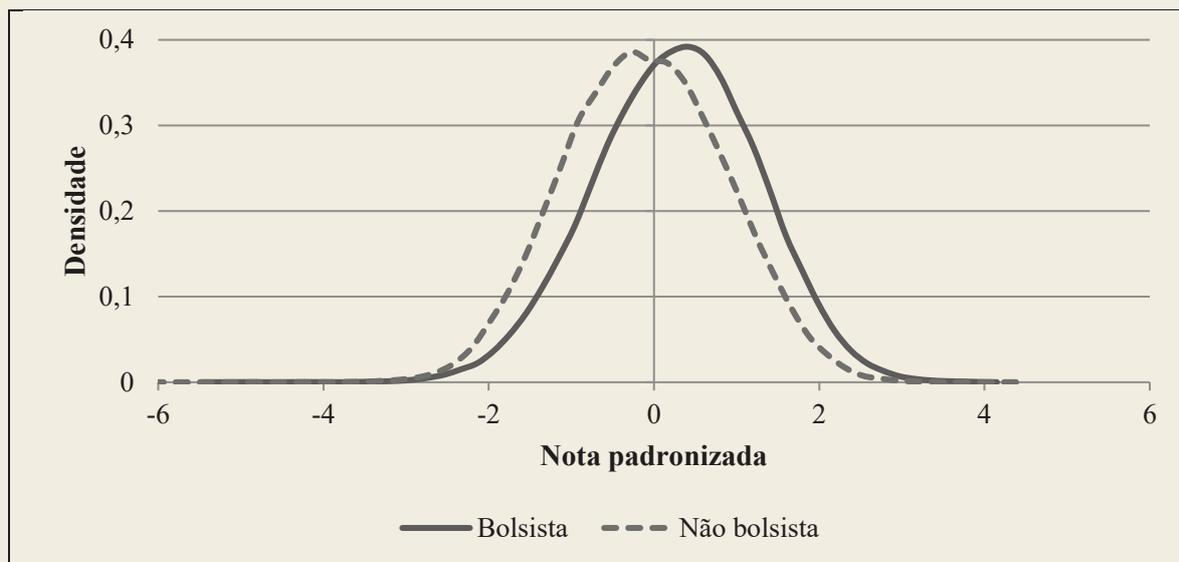
ÁREA DE CONHECIMENTO	TOTAL	BOLSISTAS DE IC		% BOLSISTAS DE IC
		SIM	NÃO	
Exatas e da terra	11.489	4.904	6.585	42,7
Biológicas	6.111	3.278	2.833	53,6
Engenharias	21.009	8.496	12.513	40,4
Saúde	15.120	5.989	9.131	39,6
Agrárias	11.400	5.729	5.671	50,3
Sociais aplicadas	32.236	3.540	28.696	11,0
Humanas	23.604	5.368	18.236	22,7
Linguística, letras e artes	8.481	1.829	6.652	21,6
TOTAL	129.450	39.133	90.317	30,2

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

Nota: A coluna "Não" refere-se tanto aos alunos que fizeram IC sem bolsa quanto aos alunos que não fizeram IC ao longo da graduação.

Observa-se, em termos relativos, uma proporção bem maior de bolsistas nas áreas biológicas e agrárias, com percentuais acima de 50%, enquanto a proporção de bolsistas na área de ciências sociais aplicadas é de apenas 11%. Porém, os dados não revelam o tempo médio de permanência com a bolsa por parte dos alunos, nem a quantidade de alunos que participaram de projetos de IC sem bolsa, que, consequentemente, fazem parte do grupo dos não tratados.

Comparando o desempenho médio de alunos que foram bolsistas de IC com os não bolsistas, verifica-se resultado superior dos bolsistas. A nota média padronizada do grupo de tratamento é 0,27, enquanto a do grupo de controle corresponde a -0,12; além da média, o resultado entre alunos bolsistas é superior ao longo de toda a distribuição das notas, como se verifica no Gráfico 2.

GRÁFICO 2 – Distribuição das notas do Enade de 2015 a 2017, tratados e não tratados

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

No entanto, além das notas, as características pessoais de tratados e não tratados divergem substancialmente. Todas as variáveis levantadas são estatisticamente diferentes entre os grupos de tratamento e controle de acordo com o teste *t* para comparar médias, sendo que apenas duas delas não apresentaram diferenças significativas ao nível de 1%. Os detalhes da estatística descritiva e o resultado desses testes encontram-se no Apêndice.

Alunos bolsistas de IC, por exemplo, apresentam nota média de formação geral 0,26 desvios-padrão superior em relação aos não bolsistas, têm probabilidade 10% menor de terem frequentado o ensino médio em escolas públicas e estão mais concentrados nas universidades do Sul e do Sudeste do país. Alunos bolsistas têm ainda média de idade quase três anos menor, possuem probabilidade aproximadamente 9% menor de pertencerem a minorias étnicas e registram chance 10% maior de possuir ao menos um dos pais com ensino superior completo.

Levando em conta essas diferenças e o seu potencial de influenciar o desempenho dos alunos no Enade, a comparação direta das notas médias entre tratados e não tratados é impossibilitada, devido à forte sugestão da existência do viés de autosseleção, sendo necessário, portanto, o uso de métodos que sejam capazes de atenuar tais diferenças, como o pareamento por escores de propensão.

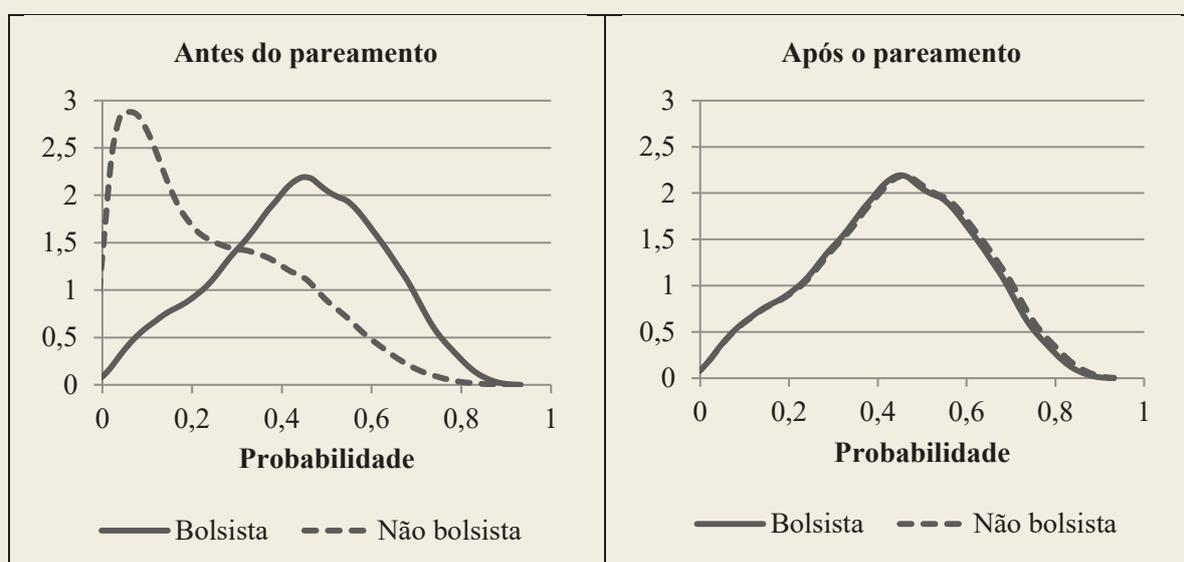
Pareamento

Para validar a aplicação do PSM, é necessário que sejam satisfeitas as hipóteses de independência condicional e de sobreposição. A primeira pode ser justificada pela qualidade do banco de dados utilizado, que conta com mais de 120.000 alunos e inclui uma vasta gama de características destes, que incluem a nota de formação geral, considerada uma *proxy* do conhecimento prévio do aluno (LÉPINE, 2016), idade,

sexo, etnia, região geográfica, renda, trabalho, educação dos pais, moradia, hábitos de estudo, participação em políticas afirmativas, percepção sobre a prova e sobre a instituição de ensino e a área de estudo.

A validade da hipótese de sobreposição fraca, por sua vez, pode ser atestada por meio do Gráfico 3, na situação anterior ao pareamento. Os escores de propensão estimados a partir do modelo *probit* para o recebimento das bolsas de IC para tratados e não tratados é sempre inferior à unidade ($P(D=1 | X) < 1$), garantindo a existência de observações comparáveis entre os grupos.

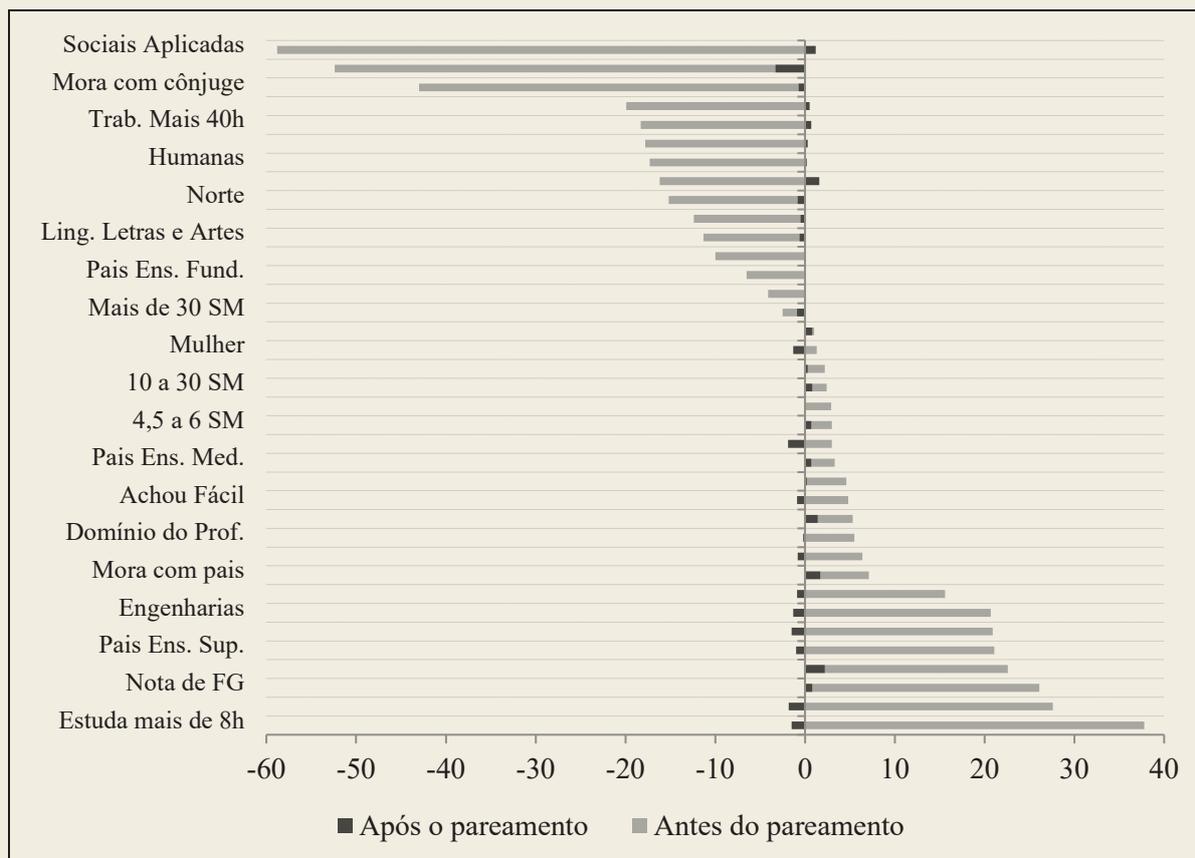
GRÁFICO 3 – Escore de propensão, bolsistas e não bolsistas, pareados e não pareados, Enade de 2015 a 2017



Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

Destaca-se que o modelo *probit* apresentou, em geral, parâmetros significativos ao nível de 1%. As únicas variáveis não significativas ao nível mínimo de 5% foram as faixas de renda de 1,5 a 4,5 e de 6 a 10 salários mínimos, além da variável binária indicando se o aluno acredita que seus professores possuíam domínio sobre o conteúdo lecionado. Dessa forma, o modelo se mostra capaz de considerar uma série de fatores relevantes na determinação da participação do aluno em projetos de IC. O resultado completo do modelo *probit* é apresentado no Apêndice.

A partir dos escores de propensão, alunos bolsistas são comparados com os não bolsistas mais próximos, em termos de probabilidade de participação no grupo de tratamento. A qualidade do pareamento realizado é apresentada no Gráfico 3, na situação posterior ao tratamento, que mostra a probabilidade de participação dos programas de IC de alunos bolsistas e seus pares. A distribuição de probabilidade é praticamente idêntica entre os grupos de tratamento e controle.

GRÁFICO 4 – Viés padronizado antes e após pareamento, Enade de 2015 a 2017

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

A qualidade do pareamento é atestada também por meio do Gráfico 4, que mostra a redução do viés padronizado das variáveis antes e após o pareamento, sugerida por Rosenbaum e Rubin (1985). O viés padronizado é definido como a diferença da média amostral entre os grupos de tratamento e controle, ponderados pela raiz quadrada da média das variâncias dos dois grupos. Após o pareamento, o viés padronizado de todas as variáveis encontra-se abaixo do limite de 5% sugerido por Caliendo e Kopeinig (2008).

As evidências demonstradas, portanto, justificam o uso do pareamento por escores de propensão como forma de obter estimativas não enviesadas dos impactos da IC sobre seus alunos bolsistas.

Efeito médio do tratamento sobre os tratados (ATT)

As estimativas para o ATT são apresentadas na Tabela 2, segundo os algoritmos de pareamento NN(1), NN(5), *radius* e *kernel*. Para os resultados desagregados por áreas, os escores de propensão foram calculados incluindo variáveis binárias para os cursos, substituindo as variáveis relativas às áreas do conhecimento. Para fins de comparação, também são apresentados resultados obtidos por meio do modelo de regressão linear, cuja variável independente é a nota do aluno no Enade, e utilizam-se as mesmas

covariadas do modelo *probit*, além de uma variável binária indicando recebimento da bolsa de IC. As estimativas dos erros-padrão estão em parênteses.

A análise global mostra resultados bastante robustos em relação ao método de pareamento utilizado: o efeito das bolsas de IC sobre os bolsistas na prova do Enade é estimado entre 0,2388 e 0,2476 desvios-padrão, sendo todas as estimativas significativas ao nível de significância de 1%. Mesmo o parâmetro calculado por meio dos mínimos quadrados ordinários apresenta um resultado semelhante aos demais.

Quando desagregados por grandes áreas, eles ainda apresentam robustez. As estimativas de todos os parâmetros mantêm a significância estatística. Em geral, a área em que a iniciação científica apresentou o menor impacto foi a de biológicas, cujo valor mínimo correspondeu a 0,2141 desvios-padrão, a partir do pareamento de Nearest Neighbour, seguida da área de linguística, letras e artes. Já as áreas com maiores impactos foram as de ciências agrárias e ciências exatas e da terra, com parâmetros atingindo máximas superiores a 0,3 desvios-padrão.

Uma forma alternativa de avaliar a robustez do impacto da bolsa de IC é compará-la com o efeito de outras bolsas acadêmicas ofertadas durante a graduação. Dessa forma, pode-se inferir se a iniciação científica é de fato um exercício efetivo e diferenciado de engajamento acadêmico ou se seu efeito é semelhante a outras formas de atividades universitárias extraclasse, considerando que a participação nessas atividades, independentemente da modalidade, gera efeitos positivos sobre o desempenho dos alunos.

TABELA 2 - Estimativas do efeito médio da bolsa de iniciação científica para os bolsistas de IC frente aos alunos sem nenhuma modalidade de bolsa. Enade de 2015 a 2017

ÁREA	NN(1)	NN(5)	RADIUS	KERNEL	OLS
Global	0,2405 *	0,2388 *	0,2476 *	0,2471 *	0,2309 *
	(0,0093)	(0,0075)	(0,007)	(0,007)	(0,006)
Exatas e da terra	0,3084 *	0,2960 *	0,2786 *	0,2789 *	0,2514 *
	(0,0347)	(0,0246)	(0,0243)	(0,0243)	(0,0184)
Biológicas	0,2141 *	0,2204 *	0,2346 *	0,2330 *	0,2360 *
	(0,0369)	(0,0354)	(0,0303)	(0,0303)	(0,0242)
Engenharias	0,2548 *	0,2685 *	0,2679 *	0,2680 *	0,2579 *
	(0,0229)	(0,0183)	(0,0156)	(0,0156)	(0,0134)
Saúde	0,2921 *	0,2799 *	0,2865 *	0,2858 *	0,2797 *
	(0,0268)	(0,0211)	(0,0183)	(0,0184)	(0,0162)
Agrárias	0,3041 *	0,2752 *	0,2930 *	0,2923 *	0,2865 *
	(0,0272)	(0,0239)	(0,0202)	(0,0202)	(0,0174)
Sociais aplicadas	0,2617 *	0,2540 *	0,2641 *	0,2627 *	0,2489 *
	(0,0304)	(0,0213)	(0,0189)	(0,019)	(0,0163)
Humanas	0,2901 *	0,2837 *	0,2824 *	0,2820 *	0,2751 *
	(0,0248)	(0,0204)	(0,0182)	(0,0183)	(0,0147)
Ling., letras e artes	0,2718 *	0,2342 *	0,2398 *	0,2376 *	0,2408 *
	(0,0418)	(0,0329)	(0,0294)	(0,0294)	(0,0236)

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

Nota: * $p < 0,01$. As estimativas dos erros-padrão estão em parênteses.

Cria-se, portanto, um novo grupo de controle, composto por alunos bolsistas de projetos de extensão, de monitoria/tutoria e bolsistas PET. O desempenho desses bolsistas é então comparado ao dos bolsistas de IC, por meio de um novo modelo de pareamento por escore de propensão. Os resultados desse exercício são apresentados na Tabela 3, com erros-padrão entre parênteses. Novamente, foram utilizados diferentes métodos de pareamento, e os resultados foram desagregados segundo áreas do conhecimento.

O resultado global mostra que a iniciação científica de fato traz benefícios acadêmicos aos alunos participantes, mesmo quando comparada com outras formas de atividades extraclasse. Todas as estimativas são significativas e robustas quanto à amplitude do efeito, variando entre 0,0626 e 0,0677 desvios-padrão. Apesar do efeito positivo, nota-se que o impacto é bastante reduzido quando comparado ao grupo de controle anterior, independentemente do algoritmo de pareamento.

TABELA 3 – Estimativas do efeito médio da bolsa de iniciação científica para os bolsistas de IC frente aos alunos bolsistas de outras modalidades. Enade de 2015 a 2017

ÁREA DE CONHECIMENTO	NN(1)	NN(5)	RADIUS	KERNEL	OLS
Global	0,0644 *	0,0626 *	0,0677 *	0,0671 *	0,0594 *
	(0,0103)	(0,0088)	(0,0076)	(0,0076)	(0,0067)
Exatas e da terra	0,0386	0,0691 *	0,0788 *	0,0783 *	0,0634 *
	(0,0334)	(0,024)	(0,0243)	(0,0244)	(0,0199)
Biológicas	0,0552	0,0319	0,0378	0,0373	0,0339
	(0,0421)	(0,0325)	(0,0311)	(0,0312)	(0,0266)
Engenharias	0,0426 ***	0,0649 *	0,0710 *	0,0706 *	0,0619 *
	(0,0242)	(0,0201)	(0,018)	(0,0181)	(0,0158)
Saúde	0,0578 **	0,0634 *	0,0661 *	0,0661 *	0,0646 *
	(0,0247)	(0,0204)	(0,018)	(0,018)	(0,0163)
Agrárias	0,0894 *	0,1079 *	0,1112 *	0,1109 *	0,0973 *
	(0,0312)	(0,0239)	(0,0214)	(0,0214)	(0,0193)
Sociais aplicadas	0,0462	0,0508 **	0,0644 *	0,0640 *	0,0658 *
	(0,0333)	(0,0241)	(0,0223)	(0,0223)	(0,0193)
Humanas	0,0704 *	0,0706 *	0,0724 *	0,0718 *	0,0683 *
	(0,0267)	(0,0235)	(0,0195)	(0,0196)	(0,0167)
Ling., letras e artes	0,0051	0,0326	0,0459	0,0459	0,0396
	(0,0466)	(0,037)	(0,0318)	(0,0318)	(0,0267)

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

Nota: *p<0,01; **p<0,05; ***p<0,1. As estimativas dos erros-padrão estão em parênteses.

Analisando os resultados desagregados, observa-se para a maior parte das áreas avaliadas que as bolsas de iniciação científica possuem, de fato, um efeito positivo superior às demais modalidades de bolsas, ainda que este efeito seja inferior ao verificado com o grupo de controle dos não bolsistas. As áreas de biológicas e linguística, letras e artes foram as únicas que não apresentaram diferenças estatisticamente significativas de desempenho entre bolsistas IC e os demais bolsistas em todos os métodos de pareamento empregados, justamente as áreas de menor impacto na análise anterior.

Assim, os resultados em geral sugerem que, independentemente do método de pareamento e do grupo de controle utilizado, a IC gera ganhos acadêmicos para os alunos contemplados com bolsas, endossando, a partir de uma nova abordagem quantitativa, os prognósticos de trabalhos anteriores que tratam da questão a partir de uma metodologia predominantemente teórica e descritiva (PIRES, 2002; CABERLON, 2003; AGUIAR, 1997).

Outro importante resultado deste trabalho, derivado da constatação de que as bolsas de IC possuem impacto superior às demais bolsas acadêmicas, é a possibilidade de separação do efeito da bolsa de IC sobre o programa de IC. Como destaca Pires (2015), é possível desvincular a IC da bolsa de IC. Enquanto a IC representa um

instrumento básico de formação, as bolsas de IC correspondem a instrumentos de incentivo individual aos melhores alunos (CNPq, 1998). Sendo assim, “a IC é uma atividade bem mais ampla que a sua pura e simples realização mediante o pagamento de uma bolsa” (MARCUSCHI, 1996, p. 6). Dessa forma, os impactos das bolsas de IC podem ocorrer a partir de duas vias distintas: a atividade da IC em si e os incentivos financeiros oferecidos pelas bolsas de IC.

Portanto, a superioridade do efeito das bolsas de IC, quando comparadas com outras bolsas acadêmicas, que também oferecem incentivos financeiros a seus participantes, é um indicativo de que a iniciação científica, enquanto atividade de pesquisa, constitui um instrumento de incentivo acadêmico, independente das possíveis vantagens financeiras atreladas a ela. Esses resultados reforçam os argumentos de Velloso (2001), Bridi e Pereira (2004) e Pires (2002), que consideram ser pequena a proporção de alunos que dependem financeiramente das bolsas de iniciação científica.

CONSIDERAÇÕES FINAIS

Este trabalho buscou avaliar o impacto das bolsas de IC sobre o desempenho acadêmico dos alunos concluintes de diferentes cursos de graduação, a partir de uma abordagem quantitativa, pouco empregada em estudos na área, dominada por análises baseadas em questionários e entrevistas.

Para determinar o contrafactual dos alunos bolsistas, utilizou-se o método do pareamento por escores de propensão, que busca encontrar as observações comparativamente mais próximas aos alunos que receberam bolsa durante a graduação. A comparação de resultados tem como base o desempenho dos alunos na prova de componente específico do Enade dos anos 2015 a 2017. Tal comparação é realizada a partir de diferentes grupos de controle, formados por alunos de graduação não bolsistas e por alunos bolsistas de outras modalidades acadêmicas.

Os resultados deste estudo atestam, em grande medida, as conjecturas de trabalhos anteriores a respeito da eficácia da iniciação científica em promover a educação superior. Independentemente dos algoritmos de pareamento envolvidos e da área de estudo avaliada, alunos bolsistas de IC apresentaram resultados estatisticamente superiores na prova do Enade em relação a alunos não bolsistas, indicando assim a robustez dos resultados encontrados.

Quando os alunos bolsistas de IC são comparados com alunos bolsistas de outras modalidades, os resultados, em geral, ainda demonstram a superioridade das notas dos alunos de IC, embora a dimensão dessa diferença seja reduzida. As exceções desses resultados são as áreas de biológicas e linguística, letras e artes, nas quais as diferenças de desempenho, embora positivas, não são estatisticamente significativas. A

importância desses resultados reside na possibilidade de desagregar os impactos do programa de iniciação científica dos incentivos financeiros atrelados a ele.

Dessa forma, o presente estudo reforça, a partir de uma nova abordagem, o argumento generalizado encontrado na literatura acerca da necessidade de manutenção e expansão dos programas de incentivo à iniciação científica, tendo em vista os impactos positivos apresentados por eles em relação à formação acadêmica dos estudantes contemplados.

A limitação deste trabalho consiste na impossibilidade de comparação direta entre alunos bolsistas da IC e alunos participantes da IC não contemplados com a bolsa. Assim, embora este estudo tenha demonstrado, por meio do segundo grupo de controle, que o desempenho apresentado pelos bolsistas não se deve exclusivamente à ajuda financeira recebida, não foi viável mensurar a dimensão na qual esse apoio financeiro também possa impactar os bolsistas. Diante de uma possível melhoria na qualidade e disponibilização dos dados, trabalhos futuros podem investigar essa lacuna deixada em aberto no estudo.

REFERÊNCIAS

- AGUIAR, L. C. C. *O perfil da iniciação científica no Instituto de Biofísica Carlos Chagas Filho e no Departamento de Bioquímica Médica da Universidade Federal do Rio de Janeiro*. 1997. Dissertação (Mestrado em Química Biológica) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1997.
- ARAGÓN, V. A.; MARTINS, C. B.; VELLOSO, J. R. *O Programa Institucional de Bolsas de Iniciação Científica (Pibic) e sua relação com a formação de cientistas*. Brasília: Nesub, 1999.
- BARIANI, I. C. D. *Estilos cognitivos de universitários e iniciação científica*. 1998. Tese (Doutorado em Educação) – Universidade Estadual de Campinas, Campinas, SP, 1998.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Enade 2015*. Brasília: Inep, 2016.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Enade 2016*. Brasília: Inep, 2017.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Microdados do Enade 2017*. Brasília: Inep, 2018.
- BREGLIA, V. L. A. *A formação na graduação: contribuições, impactos e repercussões do Pibic*. 2002. Tese (Doutorado em Educação) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2002.
- BRIDI, J. C. A. *A iniciação científica na formação do universitário*. 2004. Dissertação (Mestrado em Educação) – Universidade Estadual de Campinas, Campinas, SP, 2004.

- BRIDI, J. C. A.; PEREIRA, E. M. A. O impacto da iniciação científica na formação universitária. *Olhar de Professor*, Ponta Grossa, PR, v. 7, n. 2, p. 77-88, 2004.
- CABERLON, V. I. *Pesquisa e graduação na Furg: em busca de compreensões sob distintos horizontes*. 2003. Tese (Doutorado em Educação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.
- CABRERO, R. C. *Formação de pesquisadores na UFSCar e na área de educação especial: impactos do programa de iniciação científica do CNPq*. 2007. Tese (Doutorado em Educação Especial) – Universidade Federal de São Carlos, São Carlos, SP, 2007.
- CALAZANS, J. Articulação teoria/prática: uma ação formadora. In: CALAZANS, J. (org.). *Iniciação científica: construindo o pensamento crítico*. São Paulo: Cortez, 1999. p. 57-78.
- CALIENDO, M.; KOPEINIG, S. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, v. 22, n. 1, p. 31-72, 2008.
- CANAAN, M. G. *Quem se torna bolsista de iniciação científica na UFMG? Uma análise de fatores que influenciam no acesso à bolsa*. 2012. Dissertação (Mestrado em Educação) – Universidade Federal de Minas Gerais, Belo Horizonte, 2012.
- CANAAN, M. G.; NOGUEIRA, M. A. Bens em disputa no campo universitário: o efeito de fatores socioeconômicos e culturais no acesso à bolsa de iniciação científica. In: MASSI, L.; QUEIROZ, S. L. (org.). *Iniciação científica: aspectos históricos, organizacionais e formativos da atividade no ensino superior brasileiro*. São Paulo: Editora Unesp, 2015. p. 65-88.
- CENTRO DE GESTÃO E ESTUDOS ESTRATÉGICOS – CGEE. *A formação de novos quadros para CT&I: avaliação do Programa Institucional de Bolsas de Iniciação Científica (Pibic)*. Brasília: CGEE, 2017.
- CONCEIÇÃO, O. C.; FRANÇA, M. T. A.; JACINTO, P. A. O programa Ciência Sem Fronteiras contribui na formação dos estudantes? Uma análise a partir do Enade 2013. In: ENCONTRO NACIONAL DE ECONOMIA, 44., 2016, Foz do Iguaçu, PR. *Anais [...]*. Foz do Iguaçu, PR: Anpec, 2016.
- CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO – CNPq. *Manual do usuário do Pibic/CNPq*. Brasília, 1998.
- CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO – CNPq. *Resolução Normativa 017/2006: bolsas por quota no país*. Brasília, 2006.
- CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO – CNPq. Séries históricas até 2015. Brasília, s.d. Disponível em: <http://cnpq.br/series-historicas>. Acesso em: 24 jan. 2019.
- COSTA, D.; SOUZA, D. G. de; GIL, M. S. A.; JAMAMI, M.; CORREIA, M. A.; AGUILLERA, F. Iniciação científica e pós-graduação: perfil do pós-graduando relacionado à sua iniciação científica. *Educação Brasileira*, Brasília, n. 43, p. 95-109, 1999.
- DUTRA, K. T. Efeito do Prouni no desempenho acadêmico: uma estimação utilizando Propensity Score Matching. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 20., 2016, Foz do Iguaçu, PR. *Anais [...]*. Foz do Iguaçu, PR: Abep, 2016.

HECKMAN, J. J.; LALONDE, R. J.; SMITH, J. A. The economics and econometrics of active labor market programs. In: ASHENFELTER, O.; CARD, D. (ed.). *Handbook of labor economics*. Amsterdam: Elsevier, 1999. v. 3A, p. 1865-2097.

KHANDKER, S.; KOOLWAL, G. B.; SAMAD, H. *Handbook on impact evaluation: quantitative methods and practices*. Washington, D.C.: The World Bank, 2009.

LECHNER, M. Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In: LECHNER, M.; PFEIFFER, F. (ed.). *Econometric evaluation of labour market policies*. Heidelberg: Physica, 2001. p. 43-58.

LECHNER, M. Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, v. 165, n. 1, p. 59-82, 2002.

LÉPINE, A. G. *Essays in the economics of education in Brazil*. 2016. Tese (Doutorado em Ciências) – Universidade de São Paulo, São Paulo, 2016.

LI, Y.; PALMA, M. A.; XU, Z. P. Impacts of playing after school on academic performance: a propensity score matching approach. *Education Economics*, v. 25, n. 6, p. 575-589, 2017.

MALDONADO, L. A. *Iniciação científica na graduação em nutrição: autonomia do pensar e do fazer na visão dos pesquisadores/orientadores*. 1998. Dissertação (Mestrado em Educação) – Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 1998.

MARCUSCHI, L. *Avaliação do Programa Institucional de Bolsas de Iniciação Científica (Pibic) do CNPq e propostas de ação*. Relatório versão final. Recife: UFPE, 1996.

MARTINS, C. B. A reforma universitária de 1968 e a abertura para o ensino superior privado no Brasil. *Educação & Sociedade*, Campinas, SP, v. 30, n. 106, p. 15-35, 2009.

MARTINS, R. C. R.; MARTINS, C. B. Programas de melhoria e inovação no ensino de graduação. *Estudos e Debates: Uma Política de Ensino Superior*, Brasília, v. 20, p. 189-221, mar. 1999.

MASSI, L.; QUEIROZ, S. L. Estudos sobre iniciação científica no Brasil: uma revisão. *Cadernos de Pesquisa*, São Paulo, v. 40, n. 139, p. 173-197, jan./abr. 2010.

MASSI, L.; QUEIROZ, S. L. A perspectiva brasileira da iniciação científica: desenvolvimento e abrangência dos programas nacionais e pesquisas acadêmicas sobre a temática. In: MASSI, L.; QUEIROZ, S. L. (org.). *Iniciação científica: aspectos históricos, organizacionais e formativos da atividade no ensino superior brasileiro*. São Paulo: Editora Unesp, 2015. p. 37-56.

MAZZAFERA, B. L.; SUGUIMOTO, H. Programa de Iniciação Científica: influência no aumento quantitativo e qualitativo da produção científica. *Revista on-line de Política e Gestão Educacional*, Araraquara, SP, v. 20, n. 1, p. 38-48, 2016.

NOGUEIRA, M. A.; CANAAN, M. G. Os “iniciados”: os bolsistas de iniciação científica e suas trajetórias acadêmicas. *TOMO: Revista do Programa de Pós-Graduação em Sociologia*, São Cristóvão, SE, n. 15, p. 41-70, 2009.

PEREIRA, J. I. R.; BITTENCOURT, M. V. L.; SILVA JUNIOR, W. S. Análise do impacto da implantação das cotas na nota Enade 2008. In: ENCONTRO NACIONAL DE ECONOMIA, 41., 2013, Foz do Iguaçu, PR. *Anais [...]*. Foz do Iguaçu, PR: Anpec, 2013.

PIRES, R. C. M. *A contribuição da iniciação científica na formação do aluno de graduação numa universidade estadual*. 2002. Dissertação (Mestrado em Educação) – Universidade Federal da Bahia, Salvador, 2002.

PIRES, R. C. M. O trabalho do professor-pesquisador e o Pibic/CNPq. In: MASSI, L.; QUEIROZ, S. L. (org.). *Iniciação científica: aspectos históricos, organizacionais e formativos da atividade no ensino superior brasileiro*. São Paulo: Editora Unesp, 2015. p. 89-108.

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

ROSENBAUM, P. R.; RUBIN, D. B. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, v. 39, n. 1, p. 33-38, 1985.

VELLOSO, J. *Ciclo virtuoso da ciência brasileira e o papel da iniciação científica*. Brasília: Núcleo de Estudos sobre o Ensino Superior, 2001.

VILLAS BOAS, G. K. Currículo, iniciação científica e evasão de estudantes de ciências sociais. *Tempo Social*, São Paulo, v. 15, n. 1, p. 45-62, 2003.

APÊNDICE

APÊNDICE A - Áreas e cursos avaliados

ÁREAS	ANOS	CURSOS	ÁREAS	ANOS	CURSOS
Exatas e da terra	2017	Ciência da Computação (Bacharelado)	Sociais aplicadas	2015	Administração
		Ciência da Computação (Licenciatura)			Administração Pública
		Engenharia da Computação			Ciências Contábeis
		Física (Bacharelado)			Ciências Econômicas
		Física (Licenciatura)			Direito
		Matemática (Bacharelado)			Jornalismo
		Matemática (Licenciatura)			Publicidade e Propaganda
		Química (Bacharelado)			Relações Internacionais
		Química (Licenciatura)			Secretariado Executivo
		Sistemas de Informação			Turismo
Biológicas	2016	Biomedicina		2016	Serviço Social
	2017	Ciências Biológicas (Bacharelado)		2017	Arquitetura e Urbanismo
		Ciências Biológicas (Licenciatura)		2015	Psicologia
Engenharias	2017	Engenharia	Humanas	2017	Ciências Sociais (Bacharelado)
		Engenharia Civil			Ciências Sociais (Licenciatura)
		Engenharia de Controle e Automação			Filosofia (Bacharelado)
		Engenharia de Produção			Filosofia (Licenciatura)
		Engenharia Elétrica			Geografia (Bacharelado)
		Engenharia Mecânica			Geografia (Licenciatura)
		Engenharia Química			História (Bacharelado)
Saúde	2016	Educação Física (Bacharelado)		2015	Design
		Enfermagem			Artes Visuais (Licenciatura)
		Farmácia			Letras – Inglês
		Fisioterapia			Letras – Português (Bacharelado)
		Fonoaudiologia			Letras – Português (Licenciatura)
		Medicina			Letras – Português e Espanhol (Licenciatura)
		Nutrição			Letras – Português e Inglês (Licenciatura)
	Odontologia	Música (Licenciatura)			
	2017	Educação Física (Licenciatura)			
	Agrárias	2016	Agronomia		
Medicina Veterinária					
Zootecnia					
2017		Engenharia Ambiental			
		Engenharia de Alimentos			
		Engenharia Florestal			

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

APÊNDICE B – Estatísticas das variáveis de caracterização dos alunos no Enade 2015 a 2017, segundo o grupo

VARIÁVEL	GLOBAL		TRATADOS		NÃO TRATADOS		TESTE t
	MÉDIA	DESVIO-PADRÃO	MÉDIA	DESVIO-PADRÃO	MÉDIA	DESVIO-PADRÃO	
Nota de CE	0,00	1,00	0,27	0,99	-0,12	0,98	64,01 *
Nota de FG	0,00	1,00	0,18	0,97	-0,08	1,00	42,84 *
Idade	26,49	6,42	24,42	3,78	27,39	7,08	-78,42 *
Federal	66,83%	0,471	73,56%	0,441	63,92%	0,480	33,99 *
Norte	10,39%	0,305	7,28%	0,260	11,73%	0,322	-24,17 *
Nordeste	31,13%	0,463	25,97%	0,438	33,36%	0,472	-26,44 *
Sul	16,35%	0,370	18,02%	0,384	15,62%	0,363	10,70 *
Centro-Oeste	9,76%	0,297	7,75%	0,267	10,63%	0,308	-16,09 *
Mulher	55,77%	0,497	56,21%	0,496	55,58%	0,497	2,08 **
Achou fácil	7,95%	0,271	8,88%	0,284	7,55%	0,264	8,12 *
Minoria étnica	45,35%	0,498	39,21%	0,488	48,01%	0,500	-29,32 *
Pais ens. 9º fund.	9,78%	0,297	8,46%	0,278	10,35%	0,305	-10,56 *
Pais ens. med.	31,23%	0,463	32,32%	0,468	30,76%	0,462	5,55 *
Pais ens. sup.	42,54%	0,494	49,80%	0,500	39,39%	0,489	34,95 *
Mora com pais	57,91%	0,494	60,35%	0,489	56,85%	0,495	11,75 *
Mora com cônjuge	18,56%	0,389	7,94%	0,270	23,16%	0,422	-65,78 *
1,5 a 3 SM	23,63%	0,425	22,43%	0,417	24,15%	0,428	-6,69 *
3 a 4,5 SM	17,47%	0,380	18,06%	0,385	17,22%	0,378	3,66 *
4,5 a 6 SM	12,09%	0,326	12,77%	0,334	11,79%	0,323	4,98 *
6 a 10 SM	14,13%	0,348	15,26%	0,360	13,64%	0,343	7,69 *
10 a 30 SM	11,29%	0,316	11,81%	0,323	11,06%	0,314	3,95 *
Mais de 30 SM	1,95%	0,138	1,71%	0,130	2,05%	0,142	-4,05 *
Trab. menos 40h	25,58%	0,436	21,87%	0,413	27,18%	0,445	-20,14 *
Trab. mais 40h	10,44%	0,306	6,74%	0,251	12,04%	0,325	-28,77 *
Pol. afirmativas	25,07%	0,433	25,96%	0,438	24,69%	0,431	4,85 *
Ens. med. pub.	58,03%	0,494	51,16%	0,500	61,01%	0,488	-33,10 *
Leu 6 livros	20,38%	0,403	21,22%	0,409	20,01%	0,400	4,94 *
Estuda mais de 8h	28,03%	0,449	40,06%	0,490	22,82%	0,420	64,41 *
Boa rel. prof.-aluno	56,61%	0,496	56,97%	0,495	56,46%	0,496	1,72 ***
Domínio do prof.	73,01%	0,444	74,71%	0,435	72,28%	0,448	9,07 *
Boa infraestrutura	52,13%	0,500	53,98%	0,498	51,32%	0,500	8,79 *
Biológicas	4,72%	0,212	8,38%	0,277	3,14%	0,174	41,09 *
Engenharias	16,23%	0,369	21,71%	0,412	13,85%	0,345	35,38 *
Saúde	11,68%	0,321	15,30%	0,360	10,11%	0,301	26,80 *
Agrárias	8,81%	0,283	14,64%	0,354	6,28%	0,243	49,20 *
Sociais aplicadas	24,90%	0,432	9,05%	0,287	31,77%	0,466	-89,48 *
Humanas	18,23%	0,386	13,72%	0,344	20,19%	0,401	-27,79 *
Ling., letras e artes	6,55%	0,247	4,67%	0,211	7,37%	0,261	-18,00 *

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

Nota: *p<0,01; **p<0,05; ***p<0,1.

APÊNDICE C - Modelo probit

VARIÁVEL	COEFICIENTE	ERRO-PADRÃO
Idade	-0,0471 *	0,001
Federal	0,2043 *	0,0091
Norte	-0,2868 *	0,0159
Nordeste	-0,1779 *	0,0108
Sul	-0,0869 *	0,012
Centro-Oeste	-0,3340 *	0,0151
Mulher	0,1094 *	0,0086
Achou fácil	0,1455 *	0,0145
Minoria étnica	-0,0210 **	0,0092
Pais ens. fund.	0,0932 *	0,0176
Pais ens. med.	0,14 *	0,0143
Pais ens. sup.	0,1771 *	0,0153
Mora com pais	-0,1509 *	0,0096
Mora com cônjuge	-0,3732 *	0,0152
1,5 a 3 SM	-0,0113	0,0127
3 a 4,5 SM	-0,0016	0,0142
4,5 a 6 SM	-0,0351 **	0,0162
6 a 10 SM	-0,0206	0,0162
10 a 30 SM	-0,0535 *	0,0179
Mais de 30 SM	-0,1748 *	0,0322
Trab. menos 40h	0,0396 *	0,0112
Trab. mais 40h	-0,2538 *	0,017
Pol. afirmativas	0,1158 *	0,0106
Ens. méd. pub	-0,1010 *	0,0106
Leu 6 livros	0,1168 *	0,0102
Estuda mais de 8h	0,3461 *	0,0088
Boa rel. prof.-aluno	0,1062 *	0,009
Domínio do prof.	0,0088	0,01
Boa infraestrutura	-0,0489 *	0,0085
Nota de FG	0,1273 *	0,0041
Biológicas	0,221 *	0,0212
Engenharias	-0,2594 *	0,0157
Saúde	-0,2312 *	0,0168
Agrárias	0,0651 *	0,0176
Sociais aplicadas	-1,1575 *	0,0163
Humanas	-0,3897 *	0,0164
Ling. letras e artes	-0,5105 *	0,021
Constante	0,9013 *	0,0359
Observações	129.450	
LR chi2	27.277,21	
Pseudo R2	0,1719	

Fonte: Elaboração dos autores a partir de dados do Inep (BRASIL, 2016, 2017, 2018).

Nota: *p<0,01; **p<0,05.

COMO CITAR ESTE ARTIGO:

MARQUES, Felipe César; CUNHA, Marina Silva da. Impactos da iniciação científica sobre o desempenho estudantil no Enade 2015-2017. *Estudos em Avaliação Educacional*, São Paulo, v. 32, e06744, 2021. DOI: <https://doi.org/10.18222/ae.v32.6744>

Recebido em: 2 AGOSTO 2019

Aprovado para publicação em: 1 JUNHO 2021



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.