

**A iniciação científica como motor da formação de pesquisadores:
uma análise utilizando *Coarsened Exact Matching***

*Scientific initiation as a researcher training engine:
an analysis using Coarsened Exact Matching*

DOI: <https://dx.doi.org/10.26694/2764-1392.799>

Andrea Felipe Cabello¹
Geisa Novais²
André Nunes³
Celso Vila Nova⁴

Resumo: No artigo faz-se uma análise do efeito do programa de iniciação científica (ProIC) da Universidade de Brasília (UnB), considerando o desempenho dos discentes beneficiários frente aos não beneficiários no ingresso na pós-graduação e no mercado de trabalho formal. O objetivo é identificar se a participação do aluno no ProIC produz efeito no ingresso em programas de pós-graduação, no mercado de trabalho formal, em sua remuneração auferida e em sua área de atuação (pesquisa e ensino). Utilizou-se procedimento de correspondência *Coarsened Exact Matching* (CEM) e modelos de regressão (*Logit*, *Probit* e Mínimos Quadrados Ordinários). Os resultados revelam um efeito positivo para os beneficiários do ProIC no ingresso na pós-graduação, um efeito negativo para o ingresso no mercado de trabalho e em relação à remuneração média auferida pelos discentes beneficiários. Porém, ao verificar o ingresso no mercado de trabalho em atividades ligadas à pesquisa e ensino, nota-se um efeito positivo do programa.

Palavras-Chave: Iniciação Científica. Avaliação de Programas. *Coarsened Exact Matching*.

Abstract: In the article analyzes the effect of the scientific undergraduate research program (ProIC) at the University of Brasília (UnB) considering the performance of beneficiary students *vis-à-vis* non-beneficiaries in entering graduate school and formal labor market. The goal is to identify whether the student's participation in ProIC produces any significant differentiation for entry into graduate programs and, also, if participation in ProIC produces a positive effect when entering the formal labor market, in their remuneration earned and in their area of activity (research and teaching). The matching procedure *Coarsened Exact Matching* (CEM) and regression models (*Logit*, *Probit* and Ordinary Least Squares) were used. The results reveal a positive effect for the beneficiaries of ProIC in entering graduate school, a negative effect for entering the labor market and in relation to the average remuneration earned by the beneficiary students. However, when verifying the entry into the job market in activities related to research and teaching, a positive effect of the program is observed.

Keywords: Scientific Research Program. Program Evaluation. *Coarsened Exact Matching*.

Artigo recebido em: 3/4/2021. Aceito em: 13/9/2021.

¹ Doutora em Economia pela Universidade de Brasília (UnB). Professora na graduação e na pós-graduação do Departamento de Economia da UnB. E-mail: andreafc@gmail.com; ORCID: <https://orcid.org/0000-0003-1489-0676>

² Mestra em Economia pela UnB. Decanato de Planejamento, Orçamento e Avaliação da UnB. E-mail: geisaestatistica@gmail.com; ORCID: <https://orcid.org/0000-0003-3821-5210>

³ Doutor em Economia pela UnB. Professor de Economia na UnB (*Campus* de Planaltina). E-mail: andrenunes@unb.br; ORCID: <https://orcid.org/0000-0001-9928-6245>

⁴ Doutor em Economia pela UnB. Professor de Economia na UnB (*Campus* de Planaltina). E-mail: celso.vilanova@gmail.com; ORCID: <https://orcid.org/0000-0003-3623-9251>

Introdução

A Iniciação Científica (IC) se insere no ensino superior por meio da vinculação entre o ensino e a pesquisa, na qual discentes de graduação têm acesso a projetos de pesquisa de docentes da instituição de ensino e podem atuar na produção científica, uma vez que a IC tem como finalidade “despertar a vocação científica e incentivar talentos potenciais entre estudantes de graduação universitária, mediante participação em projeto de pesquisa, orientados por pesquisador qualificado” (BRASIL, 2006, *np*).

Estudos sinalizam (FAVA-DE-MORAES e FAVA, 2000; MASSI e QUEIROZ, 2010; CGEE, 2017; PINHO, 2017; VILLAS BÔAS, 2003) as contribuições da IC para os discentes que dela participam, tais como: melhor desempenho na graduação, desenvolvimento pessoal em relação às habilidades que são despertadas, maior integração entre a graduação e a pós-graduação, formação de quadros para a pesquisa, entre outras. Pinho (2017) destaca a importância da inserção dos alunos de graduação em programas de iniciação, visto que a atividade de pesquisa possibilita a realização das práticas de aprendizagens além do que é proposto na grade curricular do curso e amplia a formação tanto acadêmica quanto profissional. Já Massi e Queiroz (2010) levantaram as principais contribuições da atividade de IC ao analisar trabalhos publicados no Brasil sobre a temática: discentes beneficiários de IC apresentam melhores coeficientes de rendimento em seus cursos de graduação; desenvolvimento pessoal e profissional; compreensão do “fazer ciência”; maior socialização profissional e favorecimento do bom desenvolvimento das atividades de IC; encaminhamento do aluno para a pós-graduação.

Em 1992, a Universidade de Brasília (UnB) passou a participar do Programa Institucional de Bolsas de Iniciação Científica, do CNPq, tendo implementado as 100 bolsas que foram concedidas inicialmente (UnB, 1993). Todos os anos são publicados três editais específicos destinados aos discentes e docentes do ensino de graduação, a saber: Programa de Iniciação Científica (PIBIC), Programa de Iniciação Científica em Desenvolvimento Tecnológico e Inovação (PIBITI) e Programa de Iniciação Científica nas Ações Afirmativas (PIBIC-AF).

Os editais estabelecem regras relativas à participação dos potenciais beneficiários e docentes orientadores do Programa, aos procedimentos de inscrição, às etapas e critérios de seleção de candidaturas qualificadas, às diretrizes de fomento à concessão e compromissos dos docentes participantes e discentes beneficiários ao evento de iniciação científica e ao cronograma de atividades do Programa (UnB, 2011). Uma das finalidades do Programa é fazer com que o discente da graduação tenha um primeiro contato com o universo da pesquisa científica e, assim, estimular que esse ingresse em um programa de pós-graduação e contribuir para a formação de recursos humanos para o mercado de trabalho, com viés para a pesquisa.

Na literatura, encontram-se diversos trabalhos que abordam as contribuições da IC para discentes de Instituições de Ensino Superior (IES), ora em cursos específicos ora para a instituição como um todo. Porém, nota-se uma limitação de estudos que trabalhem a temática de forma quantitativa (avaliação com índice econômico e avaliação de impacto) e no âmbito da UnB. Nesse sentido, esse trabalho busca contribuir com essa literatura ao utilizar ferramentas quantitativas de avaliação, por meio de métodos poucos explorados nessa temática e uma avaliação sistemática do programa de iniciação científica a fim de auxiliar a instituição em possíveis tomadas de decisões e aperfeiçoamentos da execução da política. Esses objetivos se justificam pelo escopo proposto pelo programa de iniciação científica. Assim, alguns dos impactos esperados por um programa desse tipo são que beneficiários tenham maior probabilidade de se engajar em atividades de pesquisa, de forma acadêmica ou profissional.

Destarte, o objetivo do trabalho é avaliar os resultados dos egressos do Programa de Iniciação Científica da UnB, considerando os diferenciais dos egressos do programa e dos que não fizeram o programa, em cinco pontos: i) ingresso no mestrado; ii) ingresso no doutorado; iii) ingresso no mercado de trabalho formal; iv) diferenciais de rendimento; v) estar empregado no setor educacional e/ou de pesquisa. Utilizou-se procedimento de correspondência *Coarsened Exact Matching* (CEM) e modelos de regressão (*Logit*, *Probit* e Mínimos Quadrados Ordinários – MQO). Os resultados, para quatro dos cinco pontos, foram apresentados por meio de modelos *logit* e *probit* para amostras completas e amostras pareadas.

Já para a diferenciação da remuneração, os resultados foram apresentados por meio de modelos MQO para amostra completa e amostra pareada. Os resultados revelaram um efeito positivo para os beneficiários do ProIC no ingresso na pós-graduação, com coeficientes iguais a 1,074 para o modelo *logit* de ingresso no mestrado e 1,154 para o modelo *logit* de ingresso no doutorado, ambos para a amostra pareada. Quanto ao ingresso no mercado de trabalho, os resultados apontam um efeito negativo, com coeficiente igual a -0,201 para o modelo *logit* para a amostra pareada. Assim como o ingresso no mercado de trabalho, os resultados indicam que os discentes beneficiários do ProIC recebem, em média, R\$ 398 a menos que os discentes não beneficiários do programa. Por fim, verifica-se que o programa apresenta efeito positivo sobre o ingresso no mercado de trabalho em atividades ligadas à pesquisa e ao ensino. Assim, os resultados sugerem que o ProIC cumpre seu propósito de incentivar uma trajetória acadêmica, seja no ingresso na pós-graduação para uma maior capacitação e atuação não necessariamente na academia, seja para atuação em áreas de educação ou técnicas.

1 Avaliação de impacto

Define-se como avaliação de impacto a análise quantitativa do efeito causal de uma determinada política ou programa, ou seja, estimam-se as mudanças que são diretamente atribuíveis àquele conjunto de medidas sobre um resultado de interesse. Para isto, as avaliações de impacto utilizam como base a construção de bons grupos de comparação (contra factual), ou seja, grupos que não sofreram a intervenção e ao serem comparados com o grupo de tratamento (beneficiários do programa/política) são similares nas demais características. O maior desafio em uma avaliação de impacto é a seleção de contra factuais válidos, para que as comparações produzam boas estimativas. Um dos requisitos para que as análises sejam fidedignas ao mensurar o efeito de determinado programa é que as características de ambos os grupos, tratados e contra factuais, sejam em média iguais.

O método de avaliação é definido de acordo com os critérios de elegibilidade utilizados para a seleção dos beneficiários. Nesse sentido, quando o programa a ser avaliado não utilizou para a seleção dos beneficiários um índice de elegibilidade, uma seleção aleatória ou não tem regras de seleção claras, pode-se utilizar o método de pareamento (*matching*) para que o problema de viés de seleção seja minimizado.

O pareamento é um dos métodos que pode ser utilizado no contexto de quase todas as regras de seleção dos beneficiários, contanto que haja um conjunto que não tenha participado do programa/política. Esse método busca identificar uma comparação pareada para cada unidade de tratamento por meio das características observáveis, sendo que as unidades de comparação diferem-se por não terem sofrido a intervenção, ou seja, os resultados das unidades de comparação são utilizados para obter qual seria o resultado da unidade tratada na ausência do programa. Neste estudo, a avaliação de impacto seguiu a linha retrospectiva, na qual as análises ocorrem após a implementação do programa e os grupos de tratamento e de comparação são gerados *a posteriori*.

A literatura relevante de estudos de avaliação de impacto para nosso trabalho se divide em quatro grupos. Nosso foco é em estudos que utilizam métodos de pareamento, seja de *propensity score matching* ou de *coarsened exact matching*, por exemplo. O primeiro grupo, mais amplo, corresponde aos estudos em educação em geral.

A avaliação de impacto de programas por meio de pareamento foi utilizada, por exemplo, na educação de nível fundamental. França e Gonçalves (2010) utilizam os dados do SAEB/2003 para avaliar o desempenho entre escolas públicas e privadas na fase inicial do ensino fundamental. Eles observam impactos relevantes de raça, salários dos professores, nível socioeconômico dos discentes

e especificidades regionais. Felício e Vasconcellos (2007) também utilizaram a mesma base de dados do SAEB/2003 para investigar o efeito da frequência sobre desempenho escolar em matemática, observando efeitos positivos entre as duas variáveis. Lombardi Filho (2015) também avalia a rede pública do ensino fundamental, utilizando dados da Prova Brasil.

O segundo grupo de estudos relevantes considera a avaliação de impacto de programas sobre a educação superior. Nesse grupo, podemos mencionar os trabalhos que se referem à assistência estudantil, como o estudo de Brand (2017), por exemplo, que analisa os efeitos do PNAES na Universidade Federal do Rio Grande (FURG), encontrando impacto positivo sobre o desempenho acadêmico dos alunos beneficiados no caso dos benefícios de transporte, alimentação mais transporte e alimentação mais moradia. Freitas *et al.* (2017) também avaliam o efeito do PNAES na mesma instituição, com impactos positivos sobre a redução da evasão e melhoria do desempenho acadêmico. Resultados semelhantes foram encontrados por Brand (2017) para o mesmo programa e instituição. Já Conceição *et al.* (2018) analisaram o efeito do programa Ciência sem Fronteiras (CsF) utilizando microdados do ENADE, concluindo que o programa tem um impacto positivo e significativo nas medidas de desempenho dos discentes presentes na nota do ENADE. Nikaj e Lund (2019) utilizam o método de *coarsened exact matching* para analisar o impacto de um programa de mentoria sobre a carreira de jovens cientistas em um programa com similaridades a um de iniciação científica. O programa fornece financiamento a jovens pesquisadores e, segundo a análise feita pelos autores, ser beneficiário aumenta a chance de receber financiamento individualmente (e de financiamentos maiores) em sua carreira como pesquisador profissional.

No terceiro grupo, podemos mencionar os estudos que analisam a avaliação de impacto de programas de educação superior sobre o mercado de trabalho. Nesse caso, muitos trabalhos se focaram no FIES. Rocha *et al.* (2016) fazem uma análise similar as de impactos dos salários também utilizando a base de dados da Relação Anual de Informações Sociais (RAIS). No entanto, os autores avaliam o impacto do Programa de Financiamento Estudantil (Fies) sobre o salário do trabalhador formal. Segundo o mesmo estudo, a participação no programa depende da idade, da cor, da região e do setor de trabalho do indivíduo, enquanto o programa tem efeito positivo. Após o pareamento, os dados mostram o efeito médio do tratamento sobre o salário mensal de quem utilizou o Fies, com um salário médio aproximadamente 20% maior do que aqueles que não concluíram o ensino superior. Não foram encontradas, no entanto, diferenças entre os formados que utilizaram e não utilizaram o programa – ou seja, o programa apenas tem efeitos como indutor de conclusão.

Sousa (2017) também estudou o impacto do Fies sobre o mercado de trabalho utilizando métodos de pareamento e de diferenças em diferenças, com foco na taxa de ocupação. Nessa análise, a participação se relaciona com sexo, região, idade e grau de instrução, e os resultados indicam que o programa teve um impacto negativo e significativo sobre a taxa de ocupação dos beneficiários. Ainda sobre o Fies, Becker e Mendonça (2019) analisam o efeito do programa sobre o tempo de conclusão dos beneficiários – observando um aumento nesse tempo, principalmente após um período de “afrouxamento nas regras de concessão e pagamento do empréstimo”. Ainda sobre salários, Teixeira e Balbinotto Netto (2016), utilizando métodos semelhantes, não encontraram impactos do programa de Seguro-Desemprego sobre o salário de reinserção dos seus beneficiados.

No quarto grupo, há trabalhos que fazem avaliação de impacto de programas sobre a produtividade científica. Kannebley Júnior *et al.* (2013) analisaram o impacto de projetos de pesquisa financiados pelos Fundos Setoriais sobre a produtividade científica dos pesquisadores universitários envolvidos nesses projetos, utilizando dados da Plataforma Lattes do CNPq – os autores observaram um aumento de 5% a 6% na produção acadêmica. Ayroza *et al.* (2019) analisaram os efeitos do Programa de Produtividade em Pesquisa (PPP) da Universidade Federal do Tocantins (UFT), utilizando os currículos *Lattes*, modelo de *propensity score matching* e Diferenças em Diferenças. Como resultado, a evolução da produtividade do grupo de controle foi maior que a do tratado, o que fez os autores chamarem a atenção para os limites dos incentivos financeiros desse tipo de programa. De Castro Garcia *et al.* (2017) analisaram os efeitos de colaborações entre universidades e empresas sobre a produtividade acadêmica de grupos de pesquisa no Brasil, observando efeitos positivos principalmente no caso de interação contínua. Sohn (2014) analisou o efeito de P&D de indústrias de biotecnologia local sobre a produtividade acadêmica com o método de *coarsened exact matching* e observou efeitos positivos sobre a produção nos assuntos correlacionados.

Em relação às avaliações de IC, nos anos mil novecentos e noventa, duas avaliações foram produzidas, Marcuschi (1996) e Aragón (1999), e ambas assinalaram os resultados relevantes obtidos pelo programa. Recentemente, mais um estudo de avaliação do programa foi lançado: “A formação de novos quadros para CT&I: avaliação do Programa Institucional de Bolsas de Iniciação Científica (PIBIC)” (CGEE, 2017). Apontou, dentre os resultados relevantes para aqueles alunos beneficiários do programa, a redução do tempo de titulação do mestrado e a considerável chance de completar tanto o mestrado quanto o doutorado, quando comparados aos alunos que não foram beneficiários do programa.

2 Metodologia

A base de dados do ProIC contempla informações dos discentes que participaram do programa no período de agosto de 2004 a julho de 2018 e edital, tipo de participação, área de conhecimento e orientador.

A base de dados de currículos, extraída da Plataforma Lattes/CNPq, contempla informações dos discentes que graduaram na UnB. As informações utilizadas para as análises da presente pesquisa foram aquelas relacionadas à formação acadêmica/titulação, atuação profissional, produções e orientações. Ressalta-se que as informações são autodeclaradas e não possuem obrigatoriedade quanto ao preenchimento e à atualização.

A RAIS fornece, a cada ano, informações sobre o trabalhador vinculado ao mercado de trabalho formal no Brasil. Neste contexto, a base disponibilizada para o estudo tem como ano-base 2017. Para os casos com mais de um vínculo empregatício, considerou-se como critério de seleção aquele entendido como o vínculo principal, isto é, com maior carga horária contratada, maior tempo de serviço no emprego e maior remuneração média nominal.

A base de graduados da UnB, extraída do Sistema de Informação Acadêmica da Graduação (SIGRA), abrange informações dos indivíduos que concluíram alguma graduação na universidade até 2017, tais como: ano de ingresso e ano de conclusão, idade ao concluir a graduação, área do curso, Índice de Rendimento Acadêmico (IRA), curso, grande área de conhecimento do curso segundo a Classificação Internacional Padronizada da Educação (*International Standard Classification of Education*, ISCED), entre outras. O recorte temporal definido para o estudo, período de 2004 a 2017, levou em consideração o período contemplado por duas bases de dados do conjunto de informações. Nesse sentido, selecionou-se, para compor a base de graduados da UnB, os discentes que se graduaram entre 2004 e 2017.

Por fim, a base da pós-graduação provê dados de todos os discentes que ingressaram na modalidade *stricto sensu* da UnB até o primeiro semestre de 2019, informando o ano de ingresso, o ano de saída, o curso e seu nível (mestrado ou doutorado).

As variáveis utilizadas para as regressões logísticas e lineares estão descritas no Quadro 1. As variáveis desse estudo foram escolhidas considerando as recomendações da literatura e as variáveis utilizada no estudo do CGEE (2017), por exemplo, que avaliou o impacto do PIBIC para a Universidade Estadual Paulista (Unesp).

Quadro 1 – Descrição das Variáveis

Variável	Descrição
Vínculo empregatício	Variável binária: assume valor 1 se o discente apresentou vínculo laboral ocorrido ou em curso no ano de 2017, valor 0 caso contrário.
Vínculo empregatício nas seções “Educação” ou “Atividades profissionais, científicas e técnicas”	Variável binária: assume valor 1 se o discente que apresentou vínculo laboral estava empregado nas seções “Educação” ou “Atividades profissionais, científicas e técnicas” de acordo com a Classificação Nacional de Atividades Econômicas (CNAE), valor 0 caso contrário.
Remuneração	Remuneração média do trabalhador (valor nominal).
ProIC	Variável binária: assume valor 1 se beneficiário ProIC, valor 0 caso contrário.
Ingresso Mestrado	Variável binária: assume valor 1 se o discente ingressou no mestrado, valor 0 caso contrário.
Ingresso Doutorado	Variável binária: assume valor 1 se o discente ingressou no doutorado, valor 0 caso contrário.
Salário Médio RAIS 2017	Salário médio dos egressos por curso (valor nominal).
Licenciatura	Variável binária: assume valor 1 se o curso do discente era licenciatura, valor 0 caso contrário.
Cotas – PPI ⁵	Variável binária: assume valor 1 se o discente ingressou na UnB por meio das cotas para pretos, pardos ou indígenas, valor 0 caso contrário.
Cotas – Baixa Renda	Variável binária: assume valor 1 se o discente ingressou na UnB por meio das cotas para baixa renda, valor 0 caso contrário.
Índice de Rendimento Acadêmico (ajustado)	Variável discreta: construída a partir do Índice de Rendimento Acadêmico do Aluno e estratificado por níveis.
Idade ao sair - Até 24 anos	Variável binária: assume valor 1 se o discente tinha até 24 anos ao concluir a graduação, valor 0 caso contrário.
Idade ao sair – De 25 a 29 anos	Variável binária: assume valor 1 se o discente tinha entre 25 e 29 anos ao concluir a graduação, valor 0 caso contrário.
Idade ao sair – De 30 a 34 anos	Variável binária: assume valor 1 se o discente tinha entre 30 e 34 anos ao concluir a graduação, valor 0 caso contrário.
Idade ao sair – Mais que 35 anos	Variável binária: assume valor 1 se o discente tinha mais que 35 anos ao concluir a graduação, valor 0 caso contrário.
Sexo - Feminino	Variável binária: assume valor 1 se feminino, valor 0 caso contrário.
Grande Área do Conhecimento: Ciências Exatas	Variável binária: assume valor 1 se a área de conhecimento do curso do aluno for exatas, valor 0 caso contrário.
Grande Área do Conhecimento: Saúde e Vida	Variável binária: assume valor 1 se a área de conhecimento do curso do aluno for saúde e vida, valor 0 caso contrário.
Ano de Saída: 2004-2017	Variável binária: serão criadas 14 variáveis, cada uma referente a um ano do intervalo. Para cada uma dessas assume valor 1 se a variável é igual ao ano de saída do aluno na UnB, valor 0 caso contrário.

Fontes: Base de graduados UnB 2004-2017. Base ProIC 2004-2017. Base *Lattes*. Base pós-graduação UnB. Elaboração própria.

O número de discentes graduados pela UnB entre 2004 e 2017 que apresentam CPF válido⁶ foi de 42.512. Desses, 26,72% participaram do ProIC durante sua graduação. A Tabela 1 apresenta o perfil dos beneficiários e não beneficiários para o período considerado nesse estudo.

⁵ A variável Cotas (PPI) é o somatório das seguintes cotas: escola pública alta renda (PPI), escola pública baixa renda (PPI), indígenas e negros.

⁶ Os alunos sem CPF válido foram descartados da análise.

Tabela 1 – Perfil de beneficiários e não beneficiários do programa ProIC da UnB (2004-2017)

Variáveis	Beneficiário ProIC		Não beneficiário ProIC		Total	
	n	(%)	n	(%)	n	(%)
Licenciatura	1.714	15,09	7.290	23,40	9.004	21,18
Bacharelado	9.646	84,91	23.862	76,60	33.508	78,82
Feminino	6.744	59,37	15.904	51,05	22.648	53,27
Masculino	4.616	40,63	15.248	48,95	19.864	46,73
Cotas ⁷						
<i>Escola Pública Alta Renda – Não PPI</i>	24	0,21	39	0,13	63	0,15
<i>Escola Pública Alta Renda - PPI</i>	31	0,27	57	0,18	88	0,21
<i>Escola Pública Baixa Renda – Não PPI</i>	10	0,09	9	0,03	19	0,04
<i>Escola Pública Baixa Renda – PPI</i>	15	0,13	24	0,08	39	0,09
<i>Índigena</i>	13	0,11	16	0,05	29	0,07
<i>Negro</i>	1.073	9,45	2.528	8,12	3.601	8,47
<i>Universal</i>	10.194	89,74	28.479	91,42	38.673	90,97
Idade ao concluir a graduação						
<i>Até 24 anos</i>	8.496	74,79	19.492	62,57	27.988	65,84
<i>De 25 a 29 anos</i>	2.554	22,48	8.726	28,01	11.280	26,53
<i>De 30 a 34 anos</i>	218	1,92	1.550	4,98	1.768	4,16
<i>Mais que 35 anos</i>	92	0,81	1.384	4,44	1.476	3,47
IRA (ajustado)						
<i>Ira 5</i>	1	0,01	5	0,02	6	0,01
<i>Ira 10</i>	-	-	5	0,02	5	0,01
<i>Ira 15</i>	2	0,02	36	0,12	38	0,09
<i>Ira 20</i>	19	0,17	440	1,41	459	1,08
<i>Ira 25</i>	183	1,61	2.336	7,50	2.519	5,93
<i>Ira 30</i>	1.504	13,24	6.065	19,47	7.569	17,80
<i>Ira 35</i>	3.510	30,90	9.579	30,75	13.089	30,79
<i>Ira 40</i>	4.359	38,37	9.185	29,48	13.544	31,86
<i>Ira 45</i>	1.782	15,69	3.501	11,24	5.283	12,43
Grande Área do Conhecimento						
<i>Ciências Exatas</i>	3.610	31,78	7.989	25,65	11.599	27,28
<i>Ciências Humanas</i>	4.095	36,05	18.392	59,04	22.487	52,90
<i>Saúde e Vida</i>	3.655	32,17	4.771	15,32	8.426	19,82
Ano de Saída						
2004	51	0,45	1.512	4,85	1.563	3,68
2005	265	2,33	1.727	5,54	1.992	4,69
2006	428	3,77	2.013	6,46	2.441	5,74
2007	622	5,48	2.327	7,47	2.949	6,94
2008	641	5,64	2.321	7,45	2.962	6,97

7 Cota Escola Pública Alta Renda = renda familiar bruta superior a 1 ½ salário mínimo *per capita*; Cota Escola Pública Baixa Renda = renda familiar bruta igual ou inferior a 1 ½ salário mínimo *per capita*.

Variáveis	Beneficiário ProIC		Não beneficiário ProIC		Total	
	n	(%)	n	(%)	n	(%)
2009	713	6,28	2.355	7,56	3.068	7,22
2010	695	6,12	2.213	7,10	2.908	6,84
2011	682	6,00	2.255	7,24	2.937	6,91
2012	832	7,32	1.972	6,33	2.804	6,60
2013	957	8,42	2.161	6,94	3.118	7,33
2014	1.157	10,18	2.265	7,27	3.422	8,05
2015	1.279	11,26	2.454	7,88	3.733	8,78
2016	1.471	12,95	2.719	8,73	4.190	9,86
2017	1.567	13,79	2.858	9,17	4.425	10,41

Fontes: Base de graduados UnB 2004-2017. Base ProIC 2004-2017. Base *Lattes*. Base pós-graduação UnB. Elaboração própria.

A Tabela 2 apresenta os dados das variáveis utilizadas nos modelos de regressão.

Tabela 2 – Perfil de beneficiários e não beneficiários do programa ProIC da UnB após a graduação – ingresso na Pós-Graduação e no Mercado de Trabalho

Variáveis	Beneficiário ProIC		Não beneficiário ProIC		Total	
	n	(%)	n	(%)	n	(%)
Ingresso Mestrado						
<i>Sim</i>	4.203	37,00	4.970	15,95	9.173	21,58
<i>Não</i>	7.157	63,00	26.182	84,05	33.339	78,42
Ingresso Doutorado						
<i>Sim</i>	1.348	11,87	1.215	3,90	2.563	6,03
<i>Não</i>	10.012	88,13	29.937	96,10	39.949	93,97
Vínculo empregatício em 2017						
<i>Sim</i>	5.554	56,71	19.211	67,90	24.765	65,02
<i>Não</i>	4.239	43,29	9.083	32,10	13.322	34,98
Empregado nas seções “Educação” ou “Atividades profissionais, científicas e técnicas” da CNAE						
<i>Sim</i>	1.120	20,17	2.669	13,89	3.789	15,30
<i>Não</i>	4.434	79,83	16.542	86,11	20.976	84,70

Fontes: Base de graduados UnB 2004-2017. Base ProIC 2004-2017. Base *Lattes*. Base pós-graduação UnB. Base RAIS 2017. Elaboração própria.

Observa-se que um pouco mais de 20% dos discentes graduados na UnB, entre 2004 e 2017, ingressaram em um mestrado. Porém, o percentual dos discentes beneficiários do ProIC é superior ao de não beneficiários. De modo semelhante, o ingresso no doutorado tem maior representatividade no grupo de beneficiários do ProIC. Já em relação ao emprego formal, pouco mais de 65% dos discentes que concluíram a graduação até 2016 estavam empregados em 2017, sendo que a distribuição percentual de empregados e não empregados entre os beneficiários do ProIC são próximas, enquanto

que no grupo de não beneficiários a maioria dos discentes (67,9%) estavam empregados. Quanto aos que estavam empregados em 2017, aproximadamente 15% possuíam vínculo nas seções de educação ou atividades profissionais, científicas e técnicas da CNAE, sendo que a representatividade dessas atividades econômicas é maior para o grupo de beneficiários do ProIC (20,1%).

2.1 *Coarsened Exact Matching* (CEM)

Para verificar se os discentes beneficiários do ProIC diferem em seus ingressos na pós-graduação e no mercado de trabalho formal em relação aos discentes não beneficiários, utilizou-se procedimento de correspondência. Verificou-se a necessidade de correspondência no estudo ao compararmos um grupo pequeno (beneficiários ProIC, $N = 11.360$) com um grupo maior (discentes não beneficiários ProIC, $N > 31.000$). Nesse sentido, supõe-se que gêmeos estatísticos possam ser encontrados para cada discente beneficiário do ProIC nas variáveis: sexo, habilitação da graduação (bacharelado ou licenciatura), grande área do conhecimento do curso, IRA, cotas e idade.

O método utilizado foi *Coarsened Exact Matching* (CEM) (IACUS *et al.*, 2008; BLACKWELL *et al.*, 2010; IACUS *et al.* 2012), uma vez que permitiu abordar a amostra com combinações observacionais equivalentes, dada sua capacidade de reduzir o desequilíbrio, dependência, erro de estimação, viés, variância, erro quadrático médio e outros critérios. O método CEM é simples e requer menos suposições.

O método CEM baseia-se em um algoritmo que fornece um estratificado das variáveis incluídas no modelo por meio de recodificação, ou seja, cada uma das variáveis é temporariamente categorizada, até que ocorra “correspondência exata” nas variáveis que foram recodificadas. Nesse sentido, o algoritmo cria um conjunto de estratos de tamanhos variados, diga $s \in S$, em que cada estrato possui unidades com valores coesos nas características observáveis X . Unidades em estratos que contêm pelo menos uma unidade tratada e uma unidade de controle são retidas e as unidades sem correspondência nos estratos restantes são removidas da amostra.

Os seguintes passos são adotados pelo algoritmo CEM:

- i) são feitas cópias das covariáveis X , denotadas por X^* ;
- ii) as cópias X^* são engrossadas de acordo com as definições de corte (número de categorias) do usuário ou, caso elas não sejam especificadas, com aquelas definidas automaticamente no algoritmo;
- iii) para cada observação de X^* , é criado um estrato, e cada observação é colocada em um dos estratos;
- iv) os estratos são atribuídos aos dados originais, X , e são eliminadas todas as observações cujos estratos não contêm, pelo menos, uma unidade de controle e uma tratada.

As unidades tratadas são denotadas por T^S no estrato s e o número de unidades tratadas por estratos por $m_T^S = T^S$, de forma similar, as unidades controle são denotados por C^S e $m_C^S = C^S$. Para cada unidade i correspondida no estrato s , o método CEM atribui os seguintes pesos:

$$W_i = 1, i \in T^S \text{ ou } \frac{m_C m_T^S}{m_T m_C^S}, i \in C^S$$

onde:

$m_C = \cup_{s \in S} m_C^S$ são as unidades correspondidas para controle;

$m_T = \cup_{s \in S} m_T^S$ são as unidades correspondidas para tratamento.

Os pesos gerados pelo algoritmo servem para calcular o diferencial no tamanho dos estratos. Nesse sentido, posteriormente ao pareamento, o efeito médio do tratamento pode ser obtido mediante uma regressão linear da variável resposta pela variável de tratamento, ponderada pelos pesos gerados pelo método CEM.

A mensuração do efeito do aluno beneficiário do ProIC na UnB exerce sobre a probabilidade de ingresso na pós-graduação e no mercado de trabalho formal ocorrerá por modelos *logit* e *probit*. Já o efeito do programa de iniciação científica sobre a variável resposta renda será avaliado por meio de um modelo de mínimos quadrados ordinários.

Para analisar a influência do ProIC nas duas situações diferentes de análise propostas nesse artigo (pós-graduação e mercado de trabalho), será necessário fazer dois pareamentos: no primeiro serão considerados todos os discentes que concluíram a graduação entre 2004 e 2017; no segundo, serão considerados os discentes que concluíram a graduação até 2016 (ou seja, entre 2004 e 2016), pois é necessário considerar um espaço de tempo para que os discentes consigam ingressar no mercado de trabalho (CGEE,2017).

As Tabelas 3 e 4 apresentam o desbalanceamento das covariadas nas amostras antes dos pareamentos. De acordo com Blackwell *et al.* (2010), a estatística L1 indica o desbalanceamento em relação à distribuição conjunta, considerando todas as interações das covariadas, sendo que o equilíbrio global perfeito é indicado pela estatística L1=0 e valores maiores indicam maior desequilíbrio entre os grupos analisados, com o máximo de L1=1, o que indica uma separação completa entre os grupos. Ela serve como um ponto de comparação entre as soluções de pareamento, assim, permitindo mensurar o aumento no balanceamento obtido. As medidas de L1, para cada variável, têm o mesmo papel. As colunas de média, mínimo, 25%, 50%, 75% e máximo indicam a diferença entre médias e quantis das distribuições dos dois grupos para cada variável.

Tabela 3 – Estatística L1 – amostra para pareamento 1 referente a pós-graduação

Variável	L1	Média	Mínimo	25%	50%	75%	Máximo
Amostra para Pareamento 1 – Pós-Graduação							
Licenciatura	0,083	-0,083	0	0	0	0	0
Grande Área – Ciências Exatas	0,061	0,061	0	0	0	0	0
Grande Área – Saúde e Vida	0,169	0,169	0	0	0	1	0
Índice de Rendimento Acadêmico (ajustado)	0,135	2,002	0	5	5	0	0
Cotas – Baixa Renda	0,001	0,001	0	0	0	0	0
Cotas – Pretos, Pardos ou Indígenas (PPI)	0,015	0,015	0	0	0	0	0
Sexo – Feminino	0,083	0,083	0	0	0	0	0
Idade ao Ingressar	0,100	-12,091	4	0	0	-1	-14
Idade ao Concluir	0,122	-12,948	2	-1	-1	-1	-1
Distância L1 Multivariada	0,428						

Fontes: Base de graduados UnB 2004-2017. Base ProIC 2004-2017. Base *Lattes*. Base pós-graduação UnB. Elaboração própria.

Tabela 4 – Estatística L1 – amostra para pareamento 2 referente ao mercado de trabalho

Variável	L1	Média	Mínimo	25%	50%	75%	Máximo
Amostra para Pareamento 2 - Mercado de trabalho							
Licenciatura	0,085	-0,085	0	0	0	0	0
Grande Área – Ciências Exatas	0,061	0,061	0	0	0	0	0
Grande Área – Saúde e Vida	0,169	0,169	0	0	0	1	0
Índice de Rendimento Acadêmico (ajustado)	0,144	21,009	0	5	5	0	0
Cotas – Baixa Renda	0,000	0,000	0	0	0	0	0
Cotas – Pretos, Pardos ou Indígenas (PPI)	0,017	0,017	0	0	0	0	0
Sexo - Feminino	0,084	0,084	0	0	0	0	0
Idade ao Ingressar	0,103	-12,447	2	0	0	-1	-14
Idade ao Concluir	0,128	-13,523	2	-1	-1	-2	-15
Distância L1 Multivariada	0,436						

Fontes: Base de graduados UnB 2004-2017. Base ProIC 2004-2017. Base *Lattes*. Base pós-graduação UnB. Elaboração própria.

No geral, as diferenças são pequenas, indicando pequeno desbalanceamento em ambos os casos. As Tabelas 5 e 6 apresentam os resultados após os pareamentos pelo método CEM.

Tabela 5 – Estatística L1 após os pareamentos – pós-graduação

Variável	L1	Média	Mínimo	25%	50%	75%	Máximo
Amostra para Pareamento 1 - Pós-Graduação							
Licenciatura	0,000	0,000	0	0	0	0	0
Grande Área – Ciências Exatas ⁸	0,000	0,000	0	0	0	0	0
Grande Área – Saúde e Vida	0,000	0,000	0	0	0	0	0
Índice de Rendimento Acadêmico (ajustado) ⁹	0,000	0,000	0	0	0	0	0
Cotas – Baixa Renda	0,000	0,000	0	0	0	0	0
Cotas – Pretos, Pardos e Indígenas (PPI)	0,000	0,000	0	0	0	0	0
Sexo – Feminino	0,000	0,000	0	0	0	0	0
Idade ao Ingressar	0,028	-0,025	0	0	0	0	0
Idade ao Sair	0,020	-0,038	1	0	0	0	0
Distância L1 Multivariada	0,217						
Observações	Não Beneficiário ProIC			Beneficiário ProIC			
Pareadas	28.293			11.231			
Não Pareadas	2.859			129			

Fontes: Base de graduados UnB 2004-2017. Base ProIC 2004-2017. Base Lattes. Base pós-graduação UnB. Elaboração própria.

Tabela 6 – Estatística L1 após os pareamentos – mercado de trabalho

Variável	L1	Média	Mínimo	25%	50%	75%	Máximo
Amostra para Pareamento 2 - Mercado de trabalho							
Licenciatura	0,000	0,000	0	0	0	0	0
Grande Área – Ciências Exatas	0,000	0,000	0	0	0	0	0
Grande Área – Saúde e Vida	0,000	0,000	0	0	0	0	0
Índice de Rendimento Acadêmico (ajustado)	0,000	0,000	0	0	0	0	0
Cotas – Baixa Renda	0,000	0,000	0	0	0	0	0
Cotas – Pretos, Pardos e Indígenas (PPI)	0,000	0,000	0	0	0	0	0
Sexo – Feminino	0,000	0,000	0	0	0	0	0
Idade ao Ingressar	0,029	-0,041	0	0	0	0	0
Idade ao Sair	0,030	-0,060	1	0	0	0	1
Distância L1 Multivariada	0,225						
Observações	Não Beneficiário ProIC			Beneficiário ProIC			
Pareadas	25.551			9.683			
Não Pareadas	2.743			110			

Fontes: Base de graduados UnB 2004-2017. Base ProIC 2004-2017. Base Lattes. Base pós-graduação UnB. Elaboração própria.

Após o pareamento, feito pelo método CEM, a medida que fornece o desequilíbrio do modelo final passou de 0,428 para 0,217, indicando um bom balanceamento entre as covariadas. Segundo Blackwell *et al.* (2010), ocorre um bom pareamento quando a estatística L1, depois do pareamento, é inferior ao valor que apresentava antes. Já Firestone (2015) considera, em seu estudo, uma estatística L1 aceitável após o pareamento próximo a 0,2. Verifica-se que apenas as variáveis de idade não

8 As variáveis de área de conhecimento foram construídas com base na classificação *ISCED* (de 8 categorias) e agrupadas em 3 para facilitar a análise.

9 Para este estudo o rendimento acadêmico varia entre 0 e 50. Os níveis utilizados são: **5** = [0,10[; **10** = [10,15[; **15** = [15,20[; **20** = [20,25[; **25** = [25,30[; **30** = [30,35[; **35** = [35,40[; **40** = [40,45[e **45** = [45,50[.

atingiram estatística L1 igual a zero, entretanto ficaram menores que antes do pareamento e próximo a zero, o que atende a literatura a respeito do método. Nesse sentido, após o pareamento, a amostra para as análises do ingresso na pós-graduação passou a ter 39.4524 discentes, sendo 28.293 discentes que não participaram do ProIC e 11.231 discentes beneficiários.

Ressalta-se que 129 discentes (1,13%) que participaram do ProIC e 2.859 (9,18%) não beneficiários do ProIC não encontram contra factuais e, assim, foram descartados. Já na Tabela 6, a medida de distância multivariada L1 para amostra que será utilizada para analisar o ingresso no mercado de trabalho, reduziu de 0,436 para 0,225, indicando bom pareamento. De forma similar à amostra anterior, a estatística L1 para as covariadas foi igual a zero para quase todas, sendo a exceção as covariadas de idade. Assim, após o pareamento, foi possível parear 25.551 discentes que não participaram do ProIC e 9.683 discentes que participaram do ProIC, sendo que 110 discentes (1,12%) beneficiários do ProIC e 2.743 (9,69%) discentes não beneficiários do ProIC não foram pareados.

Segundo Firestone (2015), se os grupos de tratamento e controle são equilibrados em suas covariáveis, os indivíduos de cada grupo não devem ser mensurados de forma diferente entre si em fatores que influenciem a probabilidade de exposição. Dessa maneira, se for encontrada uma diferença entre os grupos expostos e não expostos no resultado de interesse, o único motivo dessa diferença deve ser o programa do qual o grupo exposto fazia parte.

Assim, por meio dos resultados finais das comparações entre os grupos obtidos mediante a aplicação do método CEM, verifica-se que o pareamento reduziu a distância multivariada L1 de ambas as amostras, garantindo uma qualidade satisfatória das amostras após o pareamento para que sejam feitas as análises posteriores.

3 Resultados

Em relação ao ingresso na pós-graduação, foi analisada a probabilidade do discente ter ingressado em algum programa de mestrado, ou seja, a variável resposta é binária (Ingresso Mestrado = 1 ou Não Ingresso Mestrado = 0). Logo a seguir, a Tabela 7 apresenta os resultados dos modelos *logit* e *probit* para a amostra completa e a amostra pareada pelo método CEM.

Tabela 7 – Modelos *Logit* e *Probit*, amostra completa e pareada, variável explicada – Ingresso no Mestrado

Variáveis	<i>LOGIT</i>	<i>LOGIT</i>	<i>PROBIT</i>	<i>PROBIT</i>
	AMOSTRA COMPLETA	AMOSTRA PAREADA	AMOSTRA COMPLETA	AMOSTRA PAREADA
Ingresso Mestrado				
ProIC	1,171*** (0,03)	1,074*** (0,03)	0,685*** (0,02)	0,636*** (0,02)
Salário médio RAIS 2017	-0,000*** (0,00)	-0,000*** (0,00)	-0,000*** (0,00)	-0,000*** (0,00)
Licenciatura	-0,094** (0,04)	-0,076 (0,04)	-0,055** (0,02)	-0,046 (0,03)
Saúde e Vida	-0,081* (0,04)	-0,040 (0,04)	-0,043* (0,02)	-0,016 (0,02)
Ciências Exatas	0,792*** (0,03)	0,832*** (0,04)	0,461*** (0,02)	0,495*** (0,02)
Sexo – Feminino	-0,216*** (0,03)	-0,241*** (0,03)	-0,123*** (0,02)	-0,141*** (0,02)
IRA (ajustado)	0,056*** (0,00)	0,058*** (0,00)	0,031*** (0,00)	0,033*** (0,00)
Cotas - PPI	-0,078 (0,05)	0,011 (0,06)	-0,043 (0,03)	0,012 (0,04)
Cotas - Baixa Renda	0,176 (0,37)	0,428 (1,07)	0,095 (0,21)	0,209 (0,58)
Idade ao sair – Até 24 anos	0,895*** (0,10)	0,832*** (0,19)	0,480*** (0,05)	0,469*** (0,10)
Idade ao sair – De 25 a 29 anos	0,583*** (0,10)	0,507** (0,19)	0,303*** (0,05)	0,281** (0,10)
Idade ao sair – De 30 a 34 anos	0,504*** (0,12)	0,314 (0,21)	0,262*** (0,06)	0,171 (0,12)
Ano Saída - 2005	-0,370*** (0,08)	-0,301** (0,10)	-0,218*** (0,05)	-0,180** (0,06)
Ano Saída - 2006	-0,531*** (0,07)	-0,459*** (0,09)	-0,310*** (0,04)	-0,276*** (0,06)
Ano Saída - 2007	-0,768*** (0,07)	-0,698*** (0,09)	-0,448*** (0,04)	-0,417*** (0,05)
Ano Saída - 2008	-0,893*** (0,07)	-0,850*** (0,09)	-0,516*** (0,04)	-0,503*** (0,06)
Ano Saída - 2009	-0,855*** (0,07)	-0,803*** (0,09)	-0,498*** (0,04)	-0,478*** (0,06)
Ano Saída - 2010	-0,754*** (0,07)	-0,737*** (0,09)	-0,440*** (0,04)	-0,444*** (0,05)
Ano Saída - 2011	-0,844*** (0,07)	-0,772*** (0,09)	-0,491*** (0,04)	-0,460*** (0,06)
Ano Saída - 2012	-0,980*** (0,07)	-0,937*** (0,09)	-0,569*** (0,04)	-0,557*** (0,06)

Variáveis	LOGIT	LOGIT	PROBIT	PROBIT
	AMOSTRA COMPLETA	AMOSTRA PAREADA	AMOSTRA COMPLETA	AMOSTRA PAREADA
Ano Saída - 2013	-0,931*** (0,07)	-0,854*** (0,09)	-0,545*** (0,04)	-0,512*** (0,05)
Ano Saída - 2014	-1,214*** (0,07)	-1,185*** (0,09)	-0,706*** (0,04)	-0,704*** (0,05)
Ano Saída - 2015	-1,290*** (0,07)	-1,231*** (0,09)	-0,751*** (0,04)	-0,733*** (0,05)
Ano Saída - 2016	-1,462*** (0,07)	-1,441*** (0,09)	-0,843*** (0,04)	-0,850*** (0,05)
Ano Saída - 2017	-1,650*** (0,07)	-1,621*** (0,09)	-0,949*** (0,04)	-0,952*** (0,05)
Constante	-3,435*** (0,16)	-3,315*** (0,25)	-1,945*** (0,09)	-1,921*** (0,14)
N	42.512	39.524	42.512	39.524
R²	0,1069	0,092	0,1069	0,092

Fonte: elaboração própria.

Nota 1: desvio padrão entre parênteses.

Nota 2: níveis de significância: *p<0.1, **p<0.05, *** p<0.01.

Observa-se que ter sido ou não beneficiário do ProIC/UnB é estatisticamente significativo a 1% em todos os modelos referentes ao ingresso no mestrado com efeito positivo. Ao analisar o módulo do coeficiente, tanto no modelo *logit* quanto no *probit*, verifica-se que o efeito é um pouco menor na amostra pareada, mas igualmente significativo como na amostra completa.

Por meio da variável “salário médio RAIS 2017”, buscou-se verificar se o salário médio dos discentes da UnB, por curso, influenciaria um estudante de ingressar em uma pós-graduação *Stricto Sensu*. Nota-se, no entanto, que, apesar da variável ser estatisticamente significativa a 1%, o seu efeito é próximo a zero, ou seja, o custo de oportunidade salarial não parece ter um efeito robusto sobre a escolha de ingresso em mestrado. Quanto à habilitação do curso de graduação, ou seja, se é um curso de licenciatura ou não, o efeito é estatisticamente significativo (negativo) apenas na amostra não pareada. Ou seja, não há evidências de que o discente ter feito uma licenciatura interfira na probabilidade de ingresso em um mestrado.

Em relação às grandes áreas de conhecimento dos cursos, a comparação foi feita em relação às Ciências Humanas, omitida do modelo. Assim, a única área que se mostrou estatisticamente significativa a 1%, nos quatro modelos, foi Ciências Exatas com um efeito positivo. Logo, discentes que concluíram um curso da área de Ciências Exatas têm maior probabilidade de ingressar em um mestrado do que um discente que concluiu algum curso da área de Ciências Humanas e esse efeito é maior quando se controla todas as variáveis em um método de pareamento. No caso de discentes da área de Saúde e Vida, o coeficiente não é estatisticamente significativo na amostra pareada.

A variável “Sexo Feminino” apresentou um efeito negativo e estatisticamente significativo a 1%. Ou seja, as mulheres têm uma probabilidade menor de ingressar em um mestrado do que os homens, e esse efeito é maior quando se considera o pareamento. Quanto às variáveis de cotas (PPI e Baixa Renda), estas não são estatisticamente significativas, e assim não demonstram evidência ao influenciar o ingresso em um mestrado. No que se refere ao rendimento acadêmico, observou-se que um maior índice de rendimento acadêmico (IRA) aumenta a probabilidade do discente ingressar em um mestrado.

Também foi analisado o efeito da idade ao concluir uma graduação, por meio das faixas etárias, no ingresso em um mestrado. Nota-se que os alunos que concluíram a graduação até os 24 anos ou entre os 25 e 29 anos de idade, respectivamente, ao nível de significância de 1% e 5%, têm maior probabilidade de ingressar em um mestrado do que os discentes que concluíram após os 35 anos. Entretanto, ao verificar os modelos para a amostra pareada, não há evidências de que os discentes que graduaram entre 30 e 34 anos de idade difiram na probabilidade de ingresso no mestrado em relação aos discentes que graduaram após os 35 anos.

Por fim, as variáveis *dummies* de ano de saída foram todas estatisticamente significativas a 1% e negativas quando comparadas ao ano de 2004 – ou seja, todas indicam uma menor probabilidade de ingresso no mestrado. No entanto, seu comportamento comparado não é linear. Observa-se que aqueles ex-discentes, concluintes entre 2013 e 2018, têm probabilidades semelhantes de ingresso no mestrado. Ou seja, há três padrões claros: i) um momento com probabilidades menores de ingresso para aqueles que concluíram após 2013, mas linearmente crescentes; ii) um momento com probabilidades um pouco maiores e semelhantes para aqueles que concluíram entre 2013 e 2018; iii) de novo, um momento com probabilidades linearmente crescentes e maiores para os que concluíram antes de 2008. Nossa hipótese é que, nesse momento intermediário de carreira, há uma consolidação no mercado de trabalho que tem efeitos sobre o ingresso no mestrado.

Utilizou-se modelos *logit* e *probit* para a amostra completa e a amostra pareada pelo método CEM, elaboradas para explicar o ingresso no doutorado. Os resultados dos modelos utilizados são apresentados na Tabela 8.

Tabela 8 – Modelos *Logit* e *Probit*, amostra completa e pareada, variável explicada – Ingresso no Doutorado

Variáveis	<i>LOGIT</i>	<i>LOGIT</i>	<i>PROBIT</i>	<i>PROBIT</i>
	AMOSTRA COMPLETA	AMOSTRA PAREADA	AMOSTRA COMPLETA	AMOSTRA PAREADA
Ingresso Doutorado				
ProIC	1,304*** (0,05)	1,154*** (0,05)	0,655*** (0,02)	0,593*** (0,03)
Salário médio RAIS 2017	-0,000*** (0,00)	-0,000*** (0,00)	-0,000** (0,00)	-0,000** (0,00)
Licenciatura	-0,175** (0,06)	-0,051 (0,08)	-0,091** (0,03)	-0,023 (0,04)
Saúde e Vida	0,002 (0,06)	0,087 (0,07)	0,012 (0,03)	0,062 (0,04)
Ciências Exatas	0,793*** (0,05)	0,848*** (0,06)	0,408*** (0,03)	0,443*** (0,03)
Sexo – Feminino	-0,380*** (0,05)	-0,441*** (0,06)	-0,187*** (0,02)	-0,220*** (0,03)
IRA (ajustado)	0,070*** (0,00)	0,080*** (0,01)	0,034*** (0,00)	0,040*** (0,00)
Cotas - PPI	-0,092 (0,09)	-0,014 (0,12)	-0,049 (0,04)	-0,009 (0,06)
Cotas - Baixa Renda	Omitida	Omitida	Omitida	Omitida
Idade ao sair – Até 24 anos	1,393*** (0,24)	1,007** (0,38)	0,616*** (0,10)	0,495** (0,17)
Idade ao sair – De 25 a 29 anos	1,061*** (0,25)	0,656 (0,39)	0,455*** (0,10)	0,323 (0,17)
Idade ao sair – De 30 a 34 anos	1,014*** (0,28)	0,560 (0,42)	0,435*** (0,12)	0,279 (0,19)
Ano Saída - 2005	-0,334** (0,11)	-0,286* (0,13)	-0,163** (0,06)	-0,144* (0,07)
Ano Saída - 2006	-0,709*** (0,11)	-0,601*** (0,14)	-0,346*** (0,06)	-0,305*** (0,07)
Ano Saída - 2007	-0,810*** (0,11)	-0,642*** (0,13)	-0,404*** (0,06)	-0,333*** (0,07)
Ano Saída - 2008	-0,988*** (0,11)	-0,898*** (0,13)	-0,491*** (0,06)	-0,461*** (0,07)
Ano Saída - 2009	-0,910*** (0,11)	-0,823*** (0,14)	-0,449*** (0,06)	-0,426*** (0,07)
Ano Saída - 2010	-0,972*** (0,11)	-0,917*** (0,13)	-0,473*** (0,06)	-0,467*** (0,07)
Ano Saída - 2011	-0,907*** (0,11)	-0,853*** (0,14)	-0,447*** (0,06)	-0,436*** (0,07)

Variáveis	<i>LOGIT</i>	<i>LOGIT</i>	<i>PROBIT</i>	<i>PROBIT</i>
	AMOSTRA COMPLETA	AMOSTRA PAREADA	AMOSTRA COMPLETA	AMOSTRA PAREADA
Ano Saída - 2012	-1,149*** (0,11)	-1,031*** (0,15)	-0,570*** (0,06)	-0,529*** (0,07)
Ano Saída - 2013	-1,122*** (0,11)	-0,978*** (0,14)	-0,562*** (0,06)	-0,504*** (0,07)
Ano Saída - 2014	-1,601*** (0,12)	-1,449*** (0,14)	-0,795*** (0,06)	-0,730*** (0,07)
Ano Saída - 2015	-1,822*** (0,12)	-1,695*** (0,14)	-0,901*** (0,06)	-0,857*** (0,07)
Ano Saída - 2016	-2,454*** (0,14)	-2,365*** (0,16)	-1,191*** (0,07)	-1,163*** (0,08)
Ano Saída - 2017	-4,432*** (0,29)	-4,154*** (0,37)	-1,972*** (0,11)	-1,866*** (0,14)
Constante	-5,646*** (0,33)	-5,613*** (0,49)	-2,909*** (0,15)	-2,991*** (0,23)
N	42.512	39.512	42.512	39.512
R²	0,1434	0,1257	0,1439	0,1258

Fonte: elaboração própria.

Nota 1: desvio padrão entre parênteses.

Nota 2: níveis de significância: *p<0.1, **p<0.05, *** p<0.01.

De forma semelhante ao encontrado no ingresso no mestrado, discentes que participaram do ProIC/UnB têm maior probabilidade de ingressar em um doutorado do que discentes que não participaram, considerando um nível de significância igual a 1%. Conclui-se, assim, que o programa cumpre um de seus objetivos, que é o de incentivar a entrada na pós-graduação de seus beneficiários.

Em relação às demais variáveis, muitos dos resultados também são similares como para a variável de salário médio por curso, para a variável Licenciatura, para as variáveis de área de conhecimento – Saúde e Vida não teve coeficiente estatisticamente significativo, enquanto Ciências Exatas teve efeito positivo e mais forte na amostra pareada. O mesmo ocorreu com a variável “Sexo Feminino”, ou seja, reforçando a dificuldade das mulheres em ingressarem em programas de pós-graduação. Já o IRA teve efeito positivo, também semelhante ao modelo de ingresso no mestrado e mais forte no modelo pareado. As variáveis de cotas não se mostraram estatisticamente significativas, de forma semelhante à ocorrida no modelo de ingresso no mestrado. Os coeficientes das estratificações de idade também apresentaram resultado semelhante ao modelo de ingresso no mestrado – em que, quanto mais novo é a idade do aluno, ao egressar, maior é a probabilidade de ele ingressar no doutorado.

Em relação ao ano de saída, os resultados são diferentes: eles se mantêm estatisticamente significativos e negativos, mas seu módulo tem um comportamento crescente no sentido de que, quanto mais distante do momento de formatura, maior é a probabilidade do ingresso no doutorado.

Nesse modelo, foram considerados os discentes da UnB que concluíram a graduação até o ano de 2016. Conforme estudo elaborado pelo CGEE (2017), considerou-se um período de 1 ano do fim da graduação até o ingresso no mercado de trabalho. A probabilidade de os discentes estarem empregados formalmente em 2017, de acordo com suas características, são apresentados na Tabela 9.

Tabela 9 – Modelos *Logit* e *Probit*, amostra completa e pareada, variável explicada
Ingresso no mercado de trabalho formal

Variáveis	<i>LOGIT</i>	<i>LOGIT</i>	<i>PROBIT</i>	<i>PROBIT</i>
	AMOSTRA COMPLETA	AMOSTRA PAREADA	AMOSTRA COMPLETA	AMOSTRA PAREADA
Vínculo Empregatício				
ProIC	-0,246*** (0,03)	-0,201*** (0,03)	-0,150*** (0,02)	-0,122*** (0,02)
Licenciatura	0,633*** (0,03)	0,657*** (0,04)	0,373*** (0,02)	0,386*** (0,02)
Saúde e Vida	-0,071* (0,03)	-0,093* (0,04)	-0,043* (0,02)	-0,053* (0,02)
Ciências Exatas	-0,084** (0,03)	-0,110** (0,04)	-0,049** (0,02)	-0,064** (0,02)
Sexo – Feminino	-0,100*** (0,02)	-0,036 (0,03)	-0,060*** (0,01)	-0,021 (0,02)
IRA (ajustado)	0,004 (0,00)	0,004 (0,00)	0,003* (0,00)	0,002 (0,00)
Cotas - PPI	0,234*** (0,04)	0,207*** (0,06)	0,145*** (0,02)	0,130*** (0,04)
Cotas - Baixa Renda	-0,366 (0,74)	Omitida	-0,227 (0,46)	Omitida
Idade ao sair – Até 24 anos	0,161* (0,06)	-0,075 (0,19)	0,114** (0,04)	-0,018 (0,11)
Idade ao sair – De 25 a 29 anos	0,015 (0,07)	-0,202 (0,19)	0,026 (0,04)	-0,094 (0,11)
Idade ao sair – De 30 a 34 anos	0,382*** (0,09)	0,144 (0,23)	0,239*** (0,05)	0,110 (0,13)
Ano Saída - 2005	0,058 (0,08)	-0,011 (0,11)	0,035 (0,05)	-0,005 (0,06)

Variáveis	LOGIT	LOGIT	PROBIT	PROBIT
	AMOSTRA COMPLETA	AMOSTRA PAREADA	AMOSTRA COMPLETA	AMOSTRA PAREADA
Ano Saída - 2006	0,021 (0,08)	-0,027 (0,11)	0,012 (0,05)	-0,014 (0,06)
Ano Saída - 2007	-0,018 (0,08)	-0,089 (0,10)	-0,010 (0,04)	-0,050 (0,06)
Ano Saída - 2008	-0,144 (0,08)	-0,294** (0,10)	-0,081 (0,04)	-0,167** (0,06)
Ano Saída - 2009	-0,170* (0,08)	-0,334** (0,10)	-0,097* (0,04)	-0,192** (0,06)
Ano Saída - 2010	-0,339*** (0,08)	-0,518*** (0,10)	-0,197*** (0,04)	-0,300*** (0,06)
Ano Saída - 2011	-0,538*** (0,07)	-0,774*** (0,10)	-0,317*** (0,04)	-0,455*** (0,06)
Ano Saída - 2012	-0,652*** (0,07)	-0,947*** (0,10)	-0,385*** (0,04)	-0,560*** (0,06)
Ano Saída - 2013	-0,895*** (0,07)	-1,132*** (0,10)	-0,534*** (0,04)	-0,675*** (0,06)
Ano Saída - 2014	-1,109*** (0,07)	-1,416*** (0,10)	-0,668*** (0,04)	-0,852*** (0,06)
Ano Saída - 2015	-1,395*** (0,07)	-1,727*** (0,10)	-0,847*** (0,04)	-1,046*** (0,06)
Ano Saída - 2016	-1,693*** (0,07)	-2,027*** (0,10)	-1,032*** (0,04)	-1,231*** (0,05)
Constante	1,031*** (0,12)	1,457*** (0,25)	0,603*** (0,07)	0,841*** (0,14)
N	38.087	35.234	38.087	35.234
R²	0,0774	0,0932	0,0772	0,093

Fonte: elaboração própria.

Nota 1: desvio padrão entre parênteses.

Nota 2: níveis de significância: *p<0.1, **p<0.05, *** p<0.01.

O fato de o discente ter participado do ProIC/UnB mostrou-se estatisticamente significativo a 1% quanto ao ingresso no mercado de trabalho formal, mas o coeficiente dessa vez é negativo. Desta maneira, verifica-se que esse discente tem menor probabilidade de estar empregado no ano de 2017, corroborando com os resultados do estudo do CGEE (2017). Esse resultado, no entanto, é condizente com os dos modelos anteriores de ingresso na pós-graduação, uma vez que discentes que participam de atividades de iniciação científica têm maior probabilidade de ingressar em uma pós-graduação (mestrado e doutorado) e o esperado é que esses concluam seus estudos para depois ingressarem no mercado de trabalho.

A variável Licenciatura, diferente dos resultados para os modelos de ingresso na pós-graduação, revelou-se estatisticamente significativa ao avaliar o ingresso no mercado de trabalho, e com sinal positivo. Deste modo, o resultado sugere que os discentes que cursaram a habilitação em Licenciatura têm maior probabilidade de estarem empregados formalmente. Isso sugere que tais cursos têm um viés mais profissionalizante do que acadêmico, uma vez que, no caso do ingresso da pós-graduação, não foi observado tal efeito.

Em relação à área do curso, há um indicativo de que os discentes que se graduaram em cursos de Ciências Exatas têm menor probabilidade de estarem empregados do que os discentes de Humanas – deve-se lembrar que esses discentes tinham maior probabilidade de ingressar na pós-graduação (em situação semelhante ao que ocorre com os beneficiários do próprio programa ProIC/UnB, sugerindo um viés mais acadêmico para essa área). Em relação aos discentes que se graduaram em cursos de Saúde e Vida, neste modelo, o coeficiente também é estatisticamente significativo, todavia, negativo, indicando menor probabilidade de estarem empregados do que os discentes de Humanas.

A variável “Sexo Feminino” não apresentou significância estatística quando verificada a amostra pareada, isto é, não há evidência suficiente que aponte diferença na probabilidade de homens e mulheres estarem empregados formalmente em 2017.

Em relação às cotas, nota-se que estas foram o único modelo em que Cotas PPI apresentaram estatística significativa a 1% na amostra completa e pareada. A este quadro, ressalta-se a importância em averiguar tal resultado em estudos futuros, que controlem por vínculo empregatício (Consolidação das Leis Trabalhistas e Serviço Público). Quanto à idade, ao concluir a graduação, não há evidências de que essa influencie no ingresso no mercado de trabalho quando se controla as demais variáveis pelo método de pareamento. As variáveis *dummies*, de ano de saída, foram estatisticamente significativas a partir do ano de 2008 quando comparadas ao ano de 2004, com efeito negativo e módulo do coeficiente crescente. Ou seja, a probabilidade de ingressar no mercado de trabalho formal, em relação aos discentes que concluíram em 2004, diminui à medida que os anos vão passando.

Prosseguindo com as análises no âmbito do mercado de trabalho formal, buscou-se verificar a remuneração média mensal dos discentes graduados até 2016 que possuíam vínculo empregatício formal em 2017. A Tabela 10 mostra os resultados dos modelos de regressão linear, para amostra completa e amostra pareada.

**Tabela 10 – Modelo MQO Amostra completa e pareada, variável explicada
Remuneração média mensal dos discentes graduados**

Variáveis	MQO	
	AMOSTRA COMPLETA	AMOSTRA PAREADA
ProIC	-542,081*** (94,82)	-398,528*** (90,77)
Licenciatura	-3125,550*** (88,21)	-3331,034*** (104,50)
Saúde e Vida	-952,719*** (104,32)	-1041,549*** (98,74)
Ciências Exatas	1169,014*** (94,15)	1108,623*** (99,65)
Sexo – Feminino	-2025,196*** (79,72)	-2122,530*** (82,11)
IRA (ajustado)	143,925*** (7,35)	165,570*** (8,82)
Cotas - PPI	-612,520*** (136,90)	-515,118*** (137,50)
Cotas - Baixa Renda	799164 (3373,13)	Omitida
Idade ao sair – Até 24 anos	-1936,428*** (202,55)	-1663,422** (508,91)
Idade ao sair – De 25 a 29 anos	-2316,078*** (209,75)	-2001,056*** (513,92)
Idade ao sair – De 30 a 34 anos	-1991,505*** (258,50)	-1732,722** (570,69)
Ano Saída - 2005	93398 (221,80)	43189 (208,58)
Ano Saída - 2006	-911,947*** (213,58)	-1111,417*** (203,33)
Ano Saída - 2007	-1637,958*** (207,01)	-1945,587*** (198,37)
Ano Saída - 2008	-2204,200*** (208,06)	-2797,679*** (200,11)
Ano Saída - 2009	-2727,417*** (207,34)	-3171,891*** (199,17)
Ano Saída - 2010	-3882,929*** (211,18)	-4216,734*** (201,75)
Ano Saída - 2011	-4525,812*** (213,03)	-5065,971*** (207,21)
Ano Saída - 2012	-5681,515*** (217,06)	-6385,969*** (211,13)
Ano Saída - 2013	-7032,564*** (216,60)	-7677,815*** (211,39)
Ano Saída - 2014	-7810,949*** (217,42)	-8448,104*** (211,93)
Ano Saída - 2015	-8398,686*** (219,84)	-9138,267*** (216,92)
Ano Saída - 2016	-8916,867*** (222,56)	-9871,991*** (219,87)
Constante	11745,741*** (375,41)	11162,994*** (636,57)
N	24.765	22.527

Fonte: elaboração própria.

Nota 1: desvio padrão entre parênteses.

Nota 2: níveis de significância: *p<0.1, **p<0.05, *** p<0.01.

Assim como o modelo anterior revelava menor probabilidade de ingresso no mercado formal para os beneficiários do programa de iniciação científica, os resultados apontam que discentes participantes do programa também têm menor rendimento médio mensal do que os não participantes do programa. Porém, esse efeito é menor quando feito o pareamento.

Quanto às características dos cursos, os resultados sugerem que os discentes licenciados auferem menor remuneração em relação aos bacharéis (apesar de maior probabilidade de ingresso no mercado formal vista no modelo anterior – sugerindo que, apesar de um perfil mais profissionalizante, são voltados para ocupações pouco valorizadas do ponto vista remuneratório).

Quando comparadas as áreas de Ciências Exatas e Saúde e Vida com Ciências Humanas, os discentes que se graduaram em cursos da área de Exatas recebem remunerações mais significativas – enquanto àqueles da área de Saúde e Vida recebem menores remunerações em relação aos graduados na área de Humanas. Isso pode ser explicado com base nos resultados anteriores: graduados na área de Exatas possuem menor chance de ingresso no mercado de trabalho formal, porém, maior possibilidade de admissão na pós-graduação. Logo, quando de fato entram no mercado, presume-se que o façam com maior capacitação que seus pares.

Quanto aos graduados na área de Saúde e Vida, apresentavam menor contingência de inserção no mercado formal sem uma viabilidade de ingresso na pós-graduação que a compensasse. Deve-se chamar a atenção, no entanto, que muitas carreiras da área de saúde seguem padrões de vínculo trabalhistas que não são percebidos pela RAIS, o que pode provocar viés de análise.

As variáveis “Sexo Feminino” e Cotas PPI são estatisticamente significantes a 1% e com efeito negativo. Para a variável sexo, percebe-se, pelo modelo anterior, que não há diferença para ingressar no mercado de trabalho formal entre homens e mulheres, mas homens, em média, auferem melhores remunerações. Para a variável Cotas PPI, essas apresentam consideráveis expectativas de admissão no mercado de trabalho, mas recebem menos (apesar de o módulo do coeficiente ser relativamente pequeno quando comparado ao da variável “Sexo Feminino”). A variável de rendimento acadêmico, mais uma vez, mostrou-se estatisticamente significativa a 1% e com coeficiente positivo, isto é, esses discentes se privam de rendimentos mais elevados, reforçando o fato de que os alunos com melhor rendimento na graduação estão à frente dos demais.

Em relação à idade, ao concluir a graduação, todas as faixas etárias são estatisticamente significativas, e com efeito negativo ao serem comparadas com os discentes que graduaram com mais de 35 anos de idade. Quanto às *dummies* de ano de saída, foram estatisticamente significativas a partir do ano de 2006, quando comparadas a 2004, com efeito negativo e módulo do coeficiente crescente. Ou seja, quanto mais recente o ano de conclusão, menor a remuneração dos discentes.

Por fim, este estudo se preocupou em analisar a viabilidade de se estar empregado em áreas de “Educação” e “Atividades profissionais, científicas e técnicas”, com o objetivo de avaliar o efeito do programa sobre a inserção de discentes em atividades ligadas à pesquisa e ao ensino, conforme visto na Tabela 11.

Tabela 11 – Modelo *Probit* e *Logit*, amostra completa e pareada, variável explicada Empregado nas seções “Educação” ou “Atividades profissionais, científicas e técnicas” da CNAE

Variáveis	LOGIT	LOGIT	PROBIT	PROBIT
	AMOSTRA COMPLETA	AMOSTRA PAREADA	AMOSTRA COMPLETA	AMOSTRA PAREADA
Vínculo “Educação” ou “Atividades profissionais, científicas e técnicas”				
ProIC	0,434*** (0,04)	0,414*** (0,05)	0,246*** (0,02)	0,234*** (0,03)
Licenciatura	0,364*** (0,04)	0,340*** (0,05)	0,198*** (0,02)	0,188*** (0,03)
Saúde e Vida	-0,552*** (0,06)	-0,530*** (0,07)	-0,290*** (0,03)	-0,281*** (0,04)
Ciências Exatas	-0,103* (0,05)	-0,129* (0,06)	-0,055* (0,02)	-0,070* (0,03)
Sexo – Feminino	0,201*** (0,04)	0,112* (0,05)	0,111*** (0,02)	0,061* (0,03)
IRA (ajustado)	0,001 (0,00)	0,008 (0,01)	0,000 (0,00)	0,004 (0,00)
Cotas PPI	-0,195** (0,07)	-0,209* (0,09)	-0,111** (0,04)	-0,118* (0,05)
Cotas baixa renda	Omitida	Omitida	Omitida	Omitida
Idade ao sair - Até 24 anos	1,160*** (0,14)	1,189*** (0,31)	0,595*** (0,07)	0,622*** (0,15)
Idade ao sair - De 25 a 29 anos	0,992*** (0,14)	1,055*** (0,31)	0,506*** (0,07)	0,549*** (0,15)
Idade ao sair - De 30 a 34 anos	0,642*** (0,16)	0,967** (0,38)	0,327*** (0,08)	0,512** (0,19)
Ano Saída - 2005	-0,257* (0,11)	-0,219 (0,15)	-0,148* (0,06)	-0,124 (0,08)
Ano Saída - 2006	-0,281* (0,11)	-0,288* (0,14)	-0,164** (0,06)	-0,165* (0,08)
Ano Saída - 2007	-0,329** (0,11)	-0,276* (0,14)	-0,191*** (0,06)	-0,159* (0,07)
Ano Saída - 2008	-0,435*** (0,11)	-0,279 (0,14)	-0,236*** (0,06)	-0,148 (0,08)
Ano Saída - 2009	-0,334** (0,11)	-0,324* (0,14)	-0,192*** (0,06)	-0,182* (0,08)

Variáveis	LOGIT	LOGIT	PROBIT	PROBIT
	AMOSTRA COMPLETA	AMOSTRA PAREADA	AMOSTRA COMPLETA	AMOSTRA PAREADA
Ano Saída - 2010	-0,153 (0,11)	-0,090 (0,14)	-0,092 (0,06)	-0,054 (0,07)
Ano Saída - 2011	-0,007 (0,11)	-0,044 (0,14)	-0,016 (0,06)	-0,033 (0,07)
Ano Saída - 2012	0,242* (0,10)	0,284* (0,13)	0,124* (0,06)	0,153* (0,07)
Ano Saída - 2013	0,247* (0,10)	0,194 (0,14)	0,124* (0,06)	0,101 (0,08)
Ano Saída - 2014	0,408*** (0,10)	0,326* (0,13)	0,218*** (0,06)	0,174* (0,07)
Ano Saída - 2015	0,412*** (0,10)	0,384** (0,13)	0,217*** (0,06)	0,205** (0,07)
Ano Saída - 2016	0,506*** (0,10)	0,395** (0,13)	0,269*** (0,06)	0,209** (0,07)
Constante	-2,996*** (0,21)	-3,216*** (0,39)	-1,683*** (0,11)	-1,822*** (0,20)
N	24.762	22.810	24.762	22.810
R²	0,0377	0,0316	0,0372	0,0314

Fonte: elaboração própria.

Nota 1: desvio padrão entre parênteses.

Nota 2: níveis de significância: *p<0.1, **p<0.05, *** p<0.01.

Em relação à variável do programa de iniciação científica, ela é estatisticamente significativa e com coeficiente positivo, indicando maior probabilidade dos beneficiários do programa de atuarem em áreas de “Educação” e “Atividades profissionais, científicas e técnicas”. O mesmo ocorre com a variável Licenciatura – o que pode ser explicado por egressos trabalhando em ocupações ligadas ao ensino fundamental e médio. Quanto às variáveis relacionadas à área do conhecimento, estas têm efeito estatisticamente significativo e negativo, tanto para as Ciências Exatas quanto para a Saúde e Vida, ao serem comparadas com as Ciências Humanas. Isso implica que essas duas áreas de conhecimento atuam em áreas menos ligadas à pesquisa e de caráter mais diretamente relacionadas à atividade produtiva.

A variável “Sexo Feminino” tem efeito estatisticamente significativo e positivo, indicando um viés desse grupo para a atuação na área de educação. A variável de desempenho acadêmico não se mostrou estatisticamente significativa, não mostrando evidências de efeito sobre a atuação nessas áreas.

Somente a variável de Cotas PPI tem efeito estatisticamente significativo e negativo, indicando menor probabilidade de atuação em áreas de “Educação” e “Atividades profissionais, científicas e técnicas”. Em relação à idade, os coeficientes são estatisticamente significativos, positivos e se reduzem à medida que o discente é mais velho, indicando que quanto mais velho é o aluno, menor é a chance de ele atuar nas áreas de “Educação” e nas “Atividades profissionais, científicas e técnicas”. Isso leva ao entendimento de que essas sejam áreas de início de carreira somente para alguns profissionais.

As variáveis de ano de saída têm comportamento diverso, em que apenas algumas são estatisticamente significativas, indicando também menor possibilidade de atuação nas áreas de “Educação” e na “Atividades profissionais, científicas e técnicas” para aqueles que se formaram há menos tempo.

Conclui-se, deste modo, que, pelo fato de o programa de IC ser voltado para uma atuação acadêmica, pressupõe-se estar relacionado a vários fatores. Em primeiro lugar, com o próprio desenho do programa, que busca formar pesquisadores, criando assim um efeito sinalização em seus egressos para a academia e não para o mercado de trabalho de modo a prejudicar tais egressos quando esses tentam atuar em outras atividades não acadêmicas de pesquisa. Nesse sentido, o transbordamento dos efeitos do programa para outras profissões, no mercado de trabalho, não parece ocorrer de forma ampla, isto é, sem considerações em relação à área de atuação, uma vez que há diferenças bastante interessantes em relação à área de atuação.

Em segundo lugar, o viés acadêmico do programa pode estar relacionado a restrições do próprio mercado de trabalho brasileiro, em que setores de baixa produtividade podem ter dificuldade de absorção de mão de obra altamente qualificada, em um fenômeno não dissociado de uma fuga de cérebros – só que, neste caso, não para o exterior (e talvez para o setor público), e sim voltada para a academia, que oferece estabilidade, já que muitos dos cargos são estatutários, algo que merece investigação, mas foge do escopo desse trabalho. Essa talvez seja menos grave se os laços da academia com empresas fossem fortes e permitissem parcerias ativas e desenvolvimentos contínuos de tecnologias.

Considerações finais

O objetivo no presente trabalho foi avaliar os resultados do Programa de IC da UnB, comparando o desempenho de seus participantes com não participantes em cinco pontos: i) ingresso no mestrado; ii) ingresso no doutorado; iii) ingresso no mercado de trabalho formal; iv) diferenciais de rendimento; v) estar empregado no setor de educação ou não. Utilizou-se procedimento de correspondência *Coarsened Exact Matching* (CEM) e modelos de regressão (*Logit*, *Probit* e Mínimos Quadrados Ordinários – MQO). Os resultados, para quatro dos cinco pontos, foram apresentados por meio de modelos *logit* e *probit* para amostras completas e amostras pareadas. Já para a diferenciação da remuneração, os resultados foram apresentados por meio de modelos MQO para amostra completa e amostra pareada.

Os resultados mostraram efeito positivo para os beneficiários do ProIC no ingresso na pós-graduação, tanto no mestrado quanto no doutorado, além de efeito negativo na inserção no mercado de trabalho, inclusive com uma remuneração inferior aos não participantes. Além disso, ainda há o ingresso no mercado de trabalho em atividades ligadas à pesquisa e ensino.

Desta maneira, os resultados sugerem que o ProIC cumpre seu propósito de incentivar uma trajetória acadêmica, podendo inclusive atuar como mecanismo de sinalização nesse sentido. A atuação de seus egressos de forma concentrada na academia pode sugerir uma possível incapacidade do mercado de trabalho brasileiro de absorver esses quadros que poderiam ser remediados por uma maior aproximação das universidades com as empresas.

Referências

ARAGÓN, Virgílio (Coord.). **O Programa Institucional de Bolsas de Iniciação Científica (PIBIC) e sua relação com a formação de cientistas (Relatório Final)**. Brasília: NESUB/UnB, 1999.

AYROZA, Igor Feitosa Lacorte; RODRIGUES, Waldecy; ROCHA, Hainnan Souza; FILHO, Manoel Xavier Pedroza. Avaliação do programa de produtividade em pesquisa da Universidade Federal do Tocantins: um olhar sob a ótica da economia comportamental. **Rev. Econ. Contemp.**, v. 23, n. 1, p. 1-24, jan./abr. 2019.

BECKER, Kalinca Lea; MENDONÇA, Mário Jorge. Políticas de financiamento estudantil: análise de impacto do Fies no tempo de conclusão do ensino superior. **Texto para Discussão do IPEA nº 2507**, Setembro, 2019.

BLACKWELL, M.; IACUS, S.; KING, G.; PORRO, G. Cem: *Coarsened Exact Matching in Stata*. **The Stata Journal**, v. 9, n. 4, p. 524–546, 2009. Disponível em: <<https://dash.harvard.edu/bitstream/handle/%20handle/1/cem-stata.pdf;jsessionid=A62A4FE9C13799BC51F6E93524337447?sequence=4314511>>. Acesso em: 22 dez. 2021.

BRAND, Joice Backes. **Análise do impacto dos benefícios PNAES no desempenho acadêmico dos estudantes nos cursos noturnos das áreas sociais e aplicadas da Universidade Federal do Rio Grande.** Monografia (Graduação) – Curso de Ciências Econômicas. 61f. ICEAC, FURG, Rio Grande, 2017.

BRASIL. Ministério da Ciência e Tecnologia. Conselho Nacional de Desenvolvimento Científico e Tecnológico. **Anexo II da RN-017/2006 – Bolsas por Quota no País.** Programa Institucional de Bolsas de Iniciação Científica - PIBIC - Norma Específica. Brasília: BRASIL, 2012. Disponível em: <http://www.cnpq.br/view/-/journal_content/56_INSTANCE_0oED/10157/100352>. Acesso em: 22 dez. 2021.

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. Disponível em: <https://www.cgee.org.br/documents/10195/734063/PIBIC_relatorio_completo.pdf/259e9590-08fc-40a6-9c1c-86646a6132e3?version=1.0>. Acesso em: 22 dez. 2021.

CONCEIÇÃO, Otavio Canozzi; FRANÇA, Marco Túlio Aniceto; JACINTO, Paulo de Andrade. O Programa Ciência sem Fronteiras contribui na formação dos estudantes? Uma análise a partir do ENADE 2013. *In:* Anais do XLIV Encontro Nacional de Economia, Rio de Janeiro, 2018.

DE CASTRO GARCIA, Renato; ARAÚJO, Veneziano de Castro; SANTOS, Emerson Gomes dos; MASCARINI, Suelene; COSTA, Ariana. Uma análise dos efeitos da interação da universidade com empresas sobre a produtividade acadêmica. **Economia Aplicada**, v. 21, n. 1, p. 5-28, 2017.

DE FELÍCIO, Fabiana; VASCONCELLOS, Lígia. O efeito da educação infantil sobre o desempenho escolar medido em exames padronizados. *In:* Anais do XXXV Encontro Nacional de Economia, Recife, 2007.

FAVA-DE-MORAES, Flávio; FAVA, Marcelo. A iniciação científica: muitas vantagens e poucos riscos. **São Paulo em Perspectiva**, v. 14, n. 1, p. 73-77, jan./mar. 2000.

FIRESTONE, Rebecca. *Evaluating Program Effectiveness: Key Concepts and How to Use Coarsened Exact Matching.* Washington (D.C.): PSI, 2015. *Population Services International.* Disponível em: <<https://www.psi.org/publication/evaluating-program-effectiveness-key-concepts-and-how-to-use-coarsened-exact-matching/>>. Acesso em: 22 dez. 2021.

FRANÇA, Marco Túlio Aniceto; GONÇALVES, Flávio de Oliveira. Provisão pública e privada de educação fundamental: diferenças de qualidade medidas por meio de propensity score. **Economia Aplicada**, v. 14, n. 4, p. 373-390, 2010.

FREITAS, Tiarajú Alves de; SCHIRMER, Sirlei; GONCALVES, Helen Sibelle Nogueira; SILVA, Juliano Pereira. Análise do desempenho dos discentes beneficiados pelo programa PNAES na Universidade Federal do Rio Grande em 2015. *In:* 3º Simpósio de Avaliação da Educação Superior, Florianópolis, 2017.

IACUS, Stefano M.; KING, Gary; PORRO, Giuseppe. *Causal Inference without Balance Checking: Coarsened Exact Matching.* **Political Analysis**, v. 20, Issue 1, p. 1-24, Winter 2012.

IACUS, Stefano M.; KING, Gary; PORRO, Giuseppe. *Matching for causal inference without balance checking.* June 2008. Disponível em: <<http://ssrn.com/abstract=1152391>>. Acesso em: 22 dez. 2021.

KANNEBLEY JÚNIOR, Sérgio; CAROLO, Murilo Damião; DE NEGRI, Fernanda. Impacto dos Fundos Setoriais sobre a produtividade acadêmica de cientistas universitários. **Estudos Econômicos (São Paulo)**, v. 43, n. 4, p. 647-685, 2013.

LOMBARDI FILHO, Stélio Coêlho. **Ensaio sobre educação na rede pública de ensino fundamental**: análises a partir dos Microdados da Prova Brasil. Dissertação (Mestrado em Economia). 77f. Curso de pós-graduação em Economia. UFPB, João Pessoa, 2015.

MARCUSCHI, Luiz. **Avaliação do Programa Institucional de Bolsas de Iniciação Científica (PIBIC) do CNPq e Propostas de Ação (Relatório Final)**. Recife: UFPE, 1996.

MASSI, Luciana; QUEIROZ, Saete Linhares. Estudo Sobre Iniciação Científica no Brasil: uma revisão. **Cadernos de Pesquisa**, v. 40, n. 139, p. 173-197, jan./abr. 2010.

NIKAJ, Silda; LUND, P. Kay. *The Impact of Individual Mentored Career Development (K) Awards on the Research Trajectories of Early-Career Scientists*. **Academic Medicine**, v. 94, n. 5, p. 708-714, 2019.

PINHO, Maria José de. Ciência e ensino: contribuições da iniciação científica na educação superior. **Avaliação**, 2017, v. 22, n. 3, p. 658-675, 2017.

ROCHA, Wilsimara M.; MONASTERIO, Leonardo M.; EHRL, Philipp. Qual foi o impacto do FIES nos salários? **Radar**, n. 46, p. 33-38, 2016.

SOHN, Eunhee. *The Impact of Local Industry R&D on Academic Science: Evidence From the Agricultural Biotechnology Industry*. **Organization Science**, v. 32, n. 3, p. 675-707, 2014.

SOUSA, José Eduardo Gonçalves de. **Avaliação de impacto do FIES sobre a taxa de ocupação dos beneficiários**. Monografia de Conclusão de Curso. 36f. Departamento de Economia. Universidade de Brasília, Brasília, 2017.

TEIXEIRA, Gibran; NETO, Giácomo Balbinotto. Seguro-desemprego brasileiro e salário de reinserção: análise empírica com regressão com descontinuidade e propensity score matching. **Nova Economia**, v. 26, n. 3, p. 943-980, 2016.

UNIVERSIDADE DE BRASÍLIA (UnB). Você Pesquisa? Então... Mostre!. **In**: Anais do 2º Seminário de Pesquisa da UnB – graduação e pós-graduação. Brasília: UnB, 1993.

UNIVERSIDADE DE BRASÍLIA (UnB). Decanato de Pesquisa e Pós-Graduação. **Resolução da CPP N.001/2011**. Brasília: UnB, 2011.

VILLAS BÔAS, Gláucia. Currículo, iniciação científica e evasão de estudantes de Ciências Sociais. **Tempo Social**, v. 15, n. 1, p. 45-62, abr. 2003.