

**IMPACTO DO PROGRAMA CIÊNCIA SEM FRONTEIRAS:
EVIDÊNCIAS A PARTIR DO ENADE 2019**

Autor: Guilherme Gomes Teixeira¹
Coautor: Walcir Soares Júnior²

Área 8: Econometria

JEL: C01

¹ Egresso da Universidade Positivo

² Prof. Dr. na Universidade Positivo

RESUMO

O presente estudo avalia quantitativamente o efeito da participação no programa Ciência Sem Fronteiras (CSF) sobre o desempenho dos alunos, expresso em suas notas no Exame Nacional de Desempenho dos Estudantes (ENADE) de 2019. As estimativas são feitas a partir das metodologias de *Propensity Score Matching* (PSM) e Mínimos Quadrados Ponderados (MQP). Os resultados sugerem, *ceteris paribus*, um efeito positivo do programa sobre a nota no componente de formação geral do ENADE da ordem de 4 a 5 pontos ou 13 a 14%, porém, sem significância estatística para o componente específico do exame. A abordagem do estudo se restringe a alunos de graduação e tecnólogo em cursos de engenharia, ciências da saúde, da natureza e da terra. Os resultados contribuem para elucidar aspectos de retorno do programa e reforçar a relevância de políticas públicas de educação. O estudo é uma abordagem alternativa à de Conceição et al. (2016) com maior abrangência de cursos, dados mais recentes e arranjo metodológico mais complexo.

Palavras-chave: Ciência Sem Fronteiras. Avaliação de impacto. *Propensity Score Matching*.

ABSTRACT

This study evaluates quantitatively the effect of the participation in the *Ciência Sem Fronteiras* (CSF) program on the performance of the students present in their grades on the *Exame Nacional de Desempenho dos Estudantes* (ENADE) of 2019. The estimates are made through the *Propensity Score Matching* (PSM) and *Weighted Least Squares* (WLS) methodologies. The results suggest, *ceteris paribus*, a positive effect of the program on the grades on the ENADE's general formation component between 4 and 5 points or 13 and 14%, with no statistically significant estimates for the exam's specific component. The study's approach is limited to undergraduate and technologist students on engineering, health, nature and earth sciences. The results contribute to clarify return aspects of the program and to stress the relevance of education policies. This study is an alternative approach to Conceição et al. (2016) with a wider range of courses, latest data and a more complex methodological arrangement.

Key words: *Ciência Sem Fronteiras*. Impact evaluation. *Propensity Score Matching*.

1. INTRODUÇÃO

A teoria do capital humano, desenvolvida ao longo de várias obras como as de Schultz (1961), Mincer (1962), Denison (1962), Becker (1964), Blaug (1976) e Becker et al. (1990), argumenta que a formação de capital humano (educação e treinamento) ocorre por meio de investimentos assim como em outras formas de capital, e que como tal, é fator determinante para o crescimento econômico de um país.

A abordagem clássica dos retornos educacionais sugere que a educação tem papel fundamental para a economia dos países por meio do melhoramento da qualidade do capital humano da força de trabalho, ou seja, sua produtividade (BECKER 1964). De acordo com Hausmann et al. (2011) ainda, o grau de complexidade de uma economia depende da formação de capital humano de seus indivíduos e a forma como esses interagem na produção de bens e serviços (conhecimento colaborativo).

Juntamente à teoria do capital humano, corroboram exemplos de países, dentre outros, como Coréia do Sul e Cingapura, cujo rápido crescimento associa-se a melhorias em seus estoques de capital humano orientadas à produção (MERK, 2015; MAITRA; BISWAJIT, 2016). Esses fatores têm contribuído para uma ávida busca por abordagens semelhantes, através de investimentos em educação, ciência, pesquisa e desenvolvimento em países desenvolvidos (como o programa *Erasmus Mundus* na União Europeia) e em países emergentes (como o programa *Becas Chile* no Chile, o *Bec.Ar* na Argentina e o *Top China* na China, etc.) (CONCEIÇÃO et al., 2016).

Nesse âmbito, o Governo Federal do Brasil lançou o programa de intercâmbio e mobilidade internacional Ciência Sem Fronteiras (CSF) em 2011, cujo objetivo era consolidar a capacidade científica, inovadora e tecnológica do Brasil por meio do oferecimento de bolsas de estudo e pesquisa para estudantes brasileiros em universidades e institutos do exterior, bem como atrair pesquisadores estrangeiros ao território brasileiro para finalidades de produção científica.

O programa previa o oferecimento de até 101 mil bolsas de graduação e pós-graduação a serem implementadas ao longo de quatro anos (2011-2015). Apesar do alto custo do programa (cerca de 10,5 bilhões no total ou 103 mil reais ao ano por aluno³), os retornos efetivos do programa permanecem pouco explorados. A literatura de avaliações do programa dispõe de estudos qualitativos, que fazem uso de questionários ou que abordam localidades específicas, mas que carecem de evidências quantitativas de causalidade, esse é o caso dos trabalhos de Azevedo et al (2013), Mari e Thiengo (2014), Cunha (2016) e Fagundes et al. (2019).

Por meio da metodologia de Propensity Score Matching (PSM), Conceição et al. (2016) estimaram de forma quantitativa o efeito do CSF sobre o desempenho dos alunos no Exame Nacional de Desempenho dos Estudantes (ENADE), porém, restrito ao ENADE 2013 e aos cursos de ciências da saúde apenas. De modo semelhante, Conceição (2017) analisou também a taxa de conclusão e evasão dos egressos do programa a partir dos microdados do Censo da Educação Superior (CENSUP) de 2014 e 2015.

Dada a abrangência do programa CSF e o alto dispêndio de recursos públicos em sua execução, justifica-se o desenvolvimento do presente estudo sob uma perspectiva teórica, pois enriquece a literatura brasileira de avaliações de políticas públicas educacionais e o debate sobre o papel da educação na economia e vice-versa; e sob uma perspectiva prática, pois contribui com a produção de futuros trabalhos, elucida aspectos de retorno do programa e colabora com o planejamento estratégico do Estado brasileiro em suas políticas de educação.

³ Relatório nº 21 da CCT do Senado Federal (2015)

Diante disso, este trabalho tem como objetivo geral estimar a partir de microdados do ENADE 2019, o efeito do programa CSF sobre o desempenho de seus beneficiários dos cursos de ciências da saúde, ciências da natureza e da terra, e engenharias, dos quais a última categoria integra importante área priorizada pelo programa e cujos cursos não foram considerados pelo estudo de Conceição et al. (2016). Justifica-se o uso de dados do ENADE 2019 por se tratar do exame mais recente no momento, por refletir informações de cursos ainda não abordados quantitativamente por outros estudos, além de referir-se a possivelmente os últimos egressos do CSF a prestarem o ENADE, dada a suspensão do programa em 2015 e o não oferecimento de novas bolsas de graduação desde então.

Apesar do uso da mesma metodologia que o estudo de Conceição et al. (2016), conforme mencionado, a contribuição deste trabalho se dá pela abordagem a partir de uma base de dados mais atual e referente também a estudantes de outros cursos, além do uso da metodologia de Mínimos Quadrados Ponderados pelos escores de propensão do PSM (MQP), possibilitando assim estimar o efeito do programa CSF para grupos distintos de beneficiários em uma abordagem duplamente robusta.

Portanto, têm-se como objetivos específicos: examinar a literatura de trabalhos anteriores que abordaram aspectos do programa CSF, compreender os retornos econômicos da educação e coletar os dados necessários para a estimativa dos resultados. Como hipótese, espera-se que o efeito do programa CSF sobre o desempenho de seus beneficiários seja positivo, de forma que esses tenham resultados, expressos por suas notas no ENADE, superiores aos de seus contrafactuais (grupo de controle) em função do programa (HECKMAN et al., 1998).

Este trabalho está dividido em cinco seções a contar com essa introdução. A primeira seção introduz o tema a ser abordado e suas implicações; a segunda seção descreve os retornos socioeconômicos da educação, o programa CSF no contexto histórico da educação no Brasil e os trabalhos que o avaliaram; a terceira seção descreve as metodologias PSM e MQP e apresenta os dados utilizados; a quarta seção expõe as estimativas e os resultados obtidos; finalmente, a quinta e última seção apresenta as conclusões do trabalho.

2. EVIDÊNCIAS DA FORMAÇÃO DE CAPITAL HUMANO: O PROGRAMA CIÊNCIA SEM FRONTEIRAS (CSF) E A EDUCAÇÃO NO BRASIL

Esta seção constrói a parte teórica do estudo, alude às principais obras da literatura quanto aos retornos educacionais e à teoria do capital humano na seção 2.1, apresenta o programa Ciência Sem Fronteiras no contexto histórico da educação no Brasil na seção 2.2 e faz uma breve revisão de literatura na seção 2.3.

2.1 A TEORIA DO CAPITAL HUMANO E OS RETORNOS DA EDUCAÇÃO

A abordagem literária com enfoque no capital humano teve seus primórdios a partir da segunda metade do século XX. De acordo com Schultz (1961) a colaboração dos indivíduos em um sistema econômico se dá por meio de sua acumulação de capital humano, a saber, a aquisição de conhecimento e habilidades que possuam valor econômico, através de investimentos que logram retornos.

Becker (1964) argumenta que o grau de investimento em capital humano depende da racionalidade dos indivíduos ante à expectativa de retornos e do tamanho das famílias, nas quais o investimento familiar *per capita* é maior quanto menor o número de membros.

Segundo Becker, na corrente tradicional da teoria do capital humano a educação aumenta a produtividade e conseqüentemente a renda dos indivíduos, enquanto que na “corrente credencialista”, os ganhos de renda se devem a credenciais que sinalizam domínio sobre conhecimentos e habilidades aos indivíduos, que justamente por serem mais produtivos originalmente, tendem a adquirir mais educação que os demais.

No modelo de Becker et al. (1990), as taxas de retorno aos investimentos em capital humano são diretamente relacionadas ao estoque de capital humano de uma economia. Desse modo, é comum aos países subdesenvolvidos ter pouco estoque e baixos retornos de capital humano, e aos países desenvolvidos ter alto e crescente estoque de capital humano com altas taxas de retorno (BARRO e LEE, 2010). Em sua concepção, esse fator explicaria a concentração de pesquisa e desenvolvimento nos países mais ricos e a fuga de cérebros a partir dos países de menor renda.

Robert Topel (1988) argumenta que economias com uma razão capital por trabalhador muito abaixo da média global alcançam rápido e intenso crescimento por meio de pesados investimentos em capital físico, aproximando-se à média global em uma “convergência condicional” (caso da Alemanha e do Japão no pós-guerra). Em seu modelo, a produção industrial orientada à exportação aumenta a demanda por mão-de-obra qualificada, cujo processo de formação de capital humano se retroalimenta.

No que se refere ao Brasil, de acordo com Menezes-Filho (2001), os retornos econômicos a cada ano adicional de escolaridade formal têm seguido um ritmo de queda ao longo do tempo, possivelmente pela expansão do número de pessoas com maior qualificação no país. O autor mostra que o diferencial salarial médio para um ano adicional de estudo formal no Brasil passou de 17% em 1977 para cerca de 14% em 1997, permanecendo, porém, como o maior entre os países da América Latina.

2.2 O PROGRAMA CSF NO CONTEXTO HISTÓRICO DA EDUCAÇÃO BRASILEIRA

Até a década de 1930, a educação no Brasil era rudimentar, restritiva e dependente do exterior para níveis mais avançados. A partir dessa época até o regime militar (1964-1985) houveram mudanças significativas como: a criação do Ministério da Educação e Cultura (MEC), do Instituto Nacional de Estudos Pedagógicos (INEP), expansão das universidades públicas e privadas, a formação universitária dos professores, elevação dos anos de escolaridade compulsória, etc (OLINDA; SÍLVIA, 2003; SCHWARTZMAN; SIMON, 2004).

A partir de 1988, com o processo de redemocratização e a promulgação da Constituição Cidadã (1988), a educação passou a ser um direito fundamental dos indivíduos através do artigo 6º, solidificando a democratização da educação no país de forma gratuita, flexível e universal (PLANALTO, 1988).

Na década de 1990 houve a promulgação da Lei de Diretrizes e Bases (LDB)⁴, a criação do programa Bolsa Escola, o Exame Nacional do Ensino Médio (ENEM) e o Fundo de Financiamento Estudantil. Nos anos 2000 e início dos anos 2010 foram criados: o Plano de Desenvolvimento da Educação (PDE), o Fundo de Manutenção e Desenvolvimento da Educação Básica e Valorização dos Profissionais da Educação (FUNDEB), o Índice de Desenvolvimento da Educação Básica (IDEB), além do Programa Universidade Para Todos (PROUNI), o Sistema de Seleção Unificada (SISU) e o Programa Nacional de Acesso ao Ensino Técnico e Emprego (PRONATEC).

⁴ PLANALTO (1996). Lei nº 9.364 de 1996

Dados do Banco Mundial apontam que: os investimentos públicos em educação no Brasil passaram de 3 a 5% do Produto Interno Bruto (PIB) na segunda metade da década de 1990 para 6,32% até 2017; a média de anos de escolaridade formal da população brasileira com mais de 25 anos de idade passou de 3,4 em 1970 para cerca de 8 anos em 2018; a taxa de alfabetização da população com mais de 15 anos passou de 74,59% em 1980 para 93,23% até 2018. De acordo com o Censo da Educação Superior (CENSUP) do INEP, o número de matrículas em cursos de graduação passou de 1,37 milhões em 1980 para 8,6 milhões em 2019 (INEP, 2000; 2017; 2019).

Nesse contexto, foi instituído o programa Ciência Sem Fronteiras (CSF) em 2011, com o objetivo de promover alta qualificação de brasileiros em universidades, centros educacionais e institutos de pesquisa internacionais de alto renome e atrair pesquisadores estrangeiros para a produção de conhecimento científico e tecnológico no Brasil (PLANALTO, 2011).

O programa previa para o período entre 2011 e 2015 o oferecimento de 101 mil bolsas de graduação, tecnólogo e pós-graduação para estudantes brasileiros de áreas classificadas como prioritárias: engenharias, ciências exatas, ciências da natureza e da terra, ciências energéticas e ciências da saúde, a serem financiadas 74% com recursos da União e 26% com recursos privados (BRASIL, 2011 apud CONCEIÇÃO et al., 2016; PLANALTO, 2011).

A proporção de modalidades de bolsas oferecidas pelo programa CSF entre 2011 e 2014 foi de: 78% para graduação sanduíche, 9% para doutorado sanduíche, 6% para pós-doutorado, 3% para doutorado pleno e 4% para as demais modalidades; a proporção para o ano de 2015 até o mês de agosto era a mesma para as modalidades de doutorado pleno e pós-doutorado, 1% maior para graduação e doutorado sanduíche e 2% menor para as demais modalidades (BRASIL, 2015).

O custo estimado do programa era de R\$ 3,2 bilhões⁵, porém, de acordo com o relatório nº 21/2015 do Senado Federal os gastos com o programa CSF haviam custado cerca de R\$ 10,5 bilhões entre 2012 e novembro de 2015, dos quais o MEC e o MCTI haviam financiado 66% e 34% respectivamente (BRASIL, 2015). A contribuição do setor privado na composição de recursos do CAPES e CNPq para financiamento do programa foi de R\$ 601.012.136,16 desde o início do programa até abril de 2015, e mais R\$ 1.054.654.851,64 a serem repassados após o encerramento do programa em 2015 (BRASIL, 2015).

O custo médio de cada bolsista do programa era de cerca de R\$ 103 mil por ano, considerado elevado se comparado com o custo público anual médio por estudante de ensino superior no Brasil no mesmo período (2012 a 2015), cerca de R\$ 21 mil (INEP, 2017). Por outro lado, o programa CSF proporcionou um envio massivo de acadêmicos brasileiros ao exterior, um estreitamento de laços colaborativos entre universidades brasileiras e internacionais em prol da produção científica e um maior destaque do ensino superior brasileiro no exterior (BRASIL, 2015).

2.3 AVALIAÇÕES DO PROGRAMA CSF

O programa CSF promoveu uma quebra de paradigma ao oferecer bolsas de graduação em abundância, especialmente para a modalidade de graduação sanduíche, em contraste à histórica ênfase brasileira em bolsas de pesquisa e pós-graduação (CASTRO ET AL., 2012; BRASIL, 2015). O programa também teve um papel importante em otimizar a competitividade internacional das universidades brasileiras, estreitando laços com universidades estrangeiras (CASTRO ET AL., 2012).

⁵ CASTRO et al. (2012). Cem Mil Bolsistas no Exterior.

Assim como em programas similares na América Latina, o programa CSF também teve problemas como: dificuldade por parte dos estudantes no domínio de idiomas estrangeiros, tendência de ingresso em instituições de ensino que não as melhores, baixo aproveitamento das bolsas devido a curta duração, incompatibilidade entre o conhecimento estrangeiro e a comunidade acadêmica brasileira e baixa absorção dos estudantes egressos no mercado de trabalho brasileiro, corroborando evidências de Schwartzman (1972).

As pesquisas de Pereira (2015) e Fagundes et al. (2019) também apontam problemas com o processo seletivo do programa, dificuldade com o requerimento não-padronizado de certificações de proficiência linguística por parte das IES de destino, dificuldades associadas a disponibilização de auxílio financeiro no exterior, pouco suporte das universidades estrangeiras e falta de acompanhamento do desempenho dos beneficiários.

Além de reforçar tais evidências, a pesquisa de Azevedo et al. (2013) apontou também melhorias na estrutura organizacional das IES e mais intensa troca de conhecimento entre os alunos em virtude do CSF e destacou também a importância das ciências sociais no progresso tecnológico, que por sua vez não eram classificadas como prioritárias pelo programa.

O trabalho de Cunha (2016) demonstra um excesso de bolsas de graduação sanduíche oferecidas pelo programa e enfatiza a importância das bolsas de pós-graduação, especialmente pós-doutorado, de desburocratizações legislativas que permitam maior atração de pesquisadores ao Brasil e o aprimoramento dos laboratórios das universidades brasileiras a fim de incentivar o retorno dos alunos.

Por meio da metodologia PSM, Conceição et al. (2016) estimaram o efeito do programa CSF sobre o desempenho dos alunos no Exame Nacional de Desempenho dos Estudantes (ENADE) de 2013. Os resultados obtidos apontaram um desempenho 3 a 6% maior em função do programa, tanto nas questões de formação geral do exame, como nas questões de formação específica (CONCEIÇÃO et al., 2016).

Conforme já mencionado, embora o trabalho de Conceição et al. (2016) tenha realizado uma avaliação do efeito do programa CSF também utilizando a metodologia PSM, a proposta do presente estudo complementa-a e difere-se em termos de periodicidade, aplicação e abrangência. Isso se deve ao fato de que o presente trabalho aborda um período mais recente (2019), utiliza também a metodologia MQP e faz uso de dados referentes também a outros cursos atendidos pelo programa CSF, a partir dos quais os resultados estimados não de ser diferentes em grau, interpretação e aplicação.

3. ESTRATÉGIA EMPÍRICA

Esta seção discorre sobre a estratégia empírica empregada no estudo, descrevendo a metodologia *Propensity Score Matching* (PSM) na seção 3.1, a metodologia de Mínimos Quadrados Ponderados (MQP) na seção 3.2 e apresentando estatísticas descritivas dos dados utilizados na estimativa de resultados na seção 3.3.

3.1 O PROPENSITY SCORE MATCHING (PSM)

Diferentemente do campo da medicina, no qual experimentos quanto ao efeito de um tratamento ou droga são realizados de forma plenamente experimental, no campo da economia, tal abordagem é geralmente rara. Existem, porém, metodologias que semelhantemente, possibilitam mensurar o efeito de determinadas intervenções ou tratamentos de interesse de maneira quase-experimental, como a metodologia utilizada neste estudo, o *Propensity Score Matching* (PSM) (ROSENBAUM e RUBIN, 1983; CAMERON e TRIVEDI, 2005).

Cameron e Trivedi (2005) explicam que o problema da não possibilidade de observação de dados sob diferentes estados simultaneamente pode ser contornado ao inferir-se causalidade na relação entre variáveis por meio de contrafactuais. Nesse caso, controla-se o efeito de outras variáveis que não o tratamento de interesse, buscando estimar qual resultado um indivíduo teria caso houvesse sido submetido ao tratamento em questão (HECKMAN et al., 1998).

Com relação ao efeito do tratamento, é estimado o Efeito Médio do Tratamento sobre os Tratados (*Average Treatment Effect on the Treated* - ATET ou ATT), que mensura a contribuição média de um determinado tratamento sobre o resultado dos indivíduos do grupo dos tratados. Tal estimativa está condicionada a pressupostos como:

- I. Hipótese da ignorabilidade: uma vez que o vetor das características observáveis (x) está incluso na regressão, não existem variáveis omitidas que distorçam o seu erro, tornando os seus resultados independentes (CAMERON e TRIVEDI, 2005).
- II. Hipótese da correspondência ou *matching*: a proporção de indivíduos tratados na composição da população ou amostra deve variar entre 0 e 100%, pois, para cada indivíduo tratado com um dado vetor de x , deve haver outro indivíduo não-tratado com um vetor de x semelhante (contrafactual) (CAMERON e TRIVEDI, 2005).
- III. Hipótese da independência média condicional: o resultado médio dos indivíduos com características semelhantes (mesmo vetor de x) deve ser igual. Controlando-se o efeito de características observáveis, qualquer diferença entre o resultado dos indivíduos tratados e não-tratados é imputado como efeito do tratamento (CAMERON e TRIVEDI, 2005).
- IV. Hipótese dos escores de propensão: caso a implementação do tratamento não seja aleatória, a probabilidade condicional de participação dos indivíduos no tratamento é calculada por meio de escores de propensão com base em suas características observáveis de x que possivelmente determinam sua participação no tratamento. O vetor de x terá independência condicional se os escores de propensão de cada uma de suas variáveis x apresentarem independência condicional (CAMERON e TRIVEDI, 2005).

As inferências proporcionadas por uma simples comparação entre um grupo de indivíduos que receberam um tratamento e um grupo de indivíduos que não receberam podem estar comprometidas caso haja um viés de seleção. Isso ocorre caso os indivíduos já possuam diferentes características entre si antes mesmo do tratamento. Características essas, que determinam a sua participação ou abstenção na exposição ao tratamento ou que por si só façam com que os resultados finais dos indivíduos sejam diferentes (HECKMAN et al., 1988; CAMERON e TRIVEDI, 2005).

Cameron e Trivedi (2005) explicam que esse viés de seleção pode ser eliminado ao se controlar o efeito de variáveis que não sejam o tratamento. No caso de variáveis observáveis omitidas, a eliminação pode ser feita ao se utilizar a função controle do estimador baseada em Mínimos Quadrados Ordinários (MQO) ou Mínimos Quadrados Generalizados (MQG), que leva em consideração todas as variáveis observáveis que determinam a participação dos indivíduos no tratamento:

$$y_i = x_i' \beta + \alpha D_i + u_i, \quad (1)$$

na qual y_i é o resultado de interesse, x são os regressores (variáveis determinantes na participação ou não no tratamento, como idade, idade ao quadrado e anos de escolaridade por exemplo), σ é o efeito estimado do tratamento, D é uma variável *dummy* que assume o valor 1 caso o indivíduo faça parte do grupo dos tratados e 0 caso for do grupo de controle, e por fim, u é o erro ou termo estocástico da regressão (CAMERON e TRIVEDI, 2005).

O PSM proporciona uma eliminação do viés de seleção ao tratamento, permitindo que a análise econométrica seja feita a partir de uma amostra de indivíduos composta por tratados e não-tratados (grupo de controle), pareados com base em seus respectivos escores de propensão a serem elegíveis ou não ao recebimento do tratamento. Dessa maneira as hipóteses condicionais à estimativa do impacto do tratamento de interesse são atendidas, uma vez que o efeito das características dos indivíduos sobre seus resultados é controlado, isto é, os indivíduos são semelhantes com base em suas características observáveis antes da exposição ao tratamento (CAMERON e TRIVEDI, 2005).

Na metodologia PSM, o pareamento dos indivíduos não é feito diretamente com base nas variáveis de características x , mas com base em suas probabilidades condicionais a serem elegíveis ao tratamento (escores de propensão) $p(x)$, que são funções específicas de x (CAMERON e TRIVEDI, 2005). O pareamento dos indivíduos com base em seus escores de propensão pode ser feito utilizando critérios como:

- I. *Nearest neighbor* ou vizinho mais próximo: abordagem que pareia indivíduos tratados a não-tratados com o escore de propensão mais próximo, *matching* do vizinho mais próximo (CAMERON e TRIVEDI, 2005).
- II. Radial ou *Cliper Matching*: os indivíduos tratados são pareados exclusivamente com indivíduos não-tratados cujos escores de propensão estão dentro de um intervalo de tolerância pré-definido por um radial (CAMERON e TRIVEDI, 2005).
- III. *Stratification*: a abordagem que estratifica os dados dos indivíduos de acordo com seus escores de propensão, pareando indivíduos tratados a indivíduos não-tratados com escores de propensão dentro de um intervalo decimal de 10 unidades que contenha o escore de propensão do indivíduo tratado (CAMERON e TRIVEDI, 2005).
- IV. *Kernel*: pareamento através de uma regressão *Kernel*, que pode ser aplicada para valores específicos de x . Porém, caso esse x seja muito dimensional o número de pareamentos pode ser muito baixo (CAMERON e TRIVEDI, 2005).

De acordo com Dehejia e Wahba (2002) apud Cameron e Trivedi (2005) cada método de *matching* será mais ou menos adequado à análise dependendo do modelo de escores de propensão utilizado. Por essa razão, os resultados obtidos neste estudo são exibidos conforme pelo menos dois dos métodos de pareamento aqui mencionados, a fim de verificar a precisão das estimativas realizadas.

3.2 MÍNIMOS QUADRADOS PONDERADOS (MQP)

Além do modelo PSM, este trabalho também faz uso da metodologia de Mínimos Quadrados Ponderados pelos escores de propensão do modelo PSM (MQP), que consiste em uma regressão de Mínimos Quadrados Ordinários (MQO) cujos componentes são ponderados de acordo com a propensão à seleção ao tratamento para cada bloco amostral estimado no modelo PSM (HILL e REITER, 2006).

De acordo com Imbens e Wooldridge (2008), dessa maneira, um problema de má especificação passível de acometer a regressão de MQO ou o modelo PSM isoladamente, é contornado, reduzindo a correlação entre as variáveis incluídas e omitidas da regressão ao ponderá-las pelos escores de propensão do modelo PSM de maneira “duplamente robusta”.

Portanto, após a realização do pareamento dos indivíduos tratados e não-tratados e a estimativa da propensão à seleção ao tratamento de seus respectivos blocos no modelo PSM, o efeito médio do programa CSF sobre o desempenho dos estudantes que dele participaram será estimado por meio da metodologia MQP controlando-se também o efeito de variáveis de desempenho não utilizadas no modelo PSM.

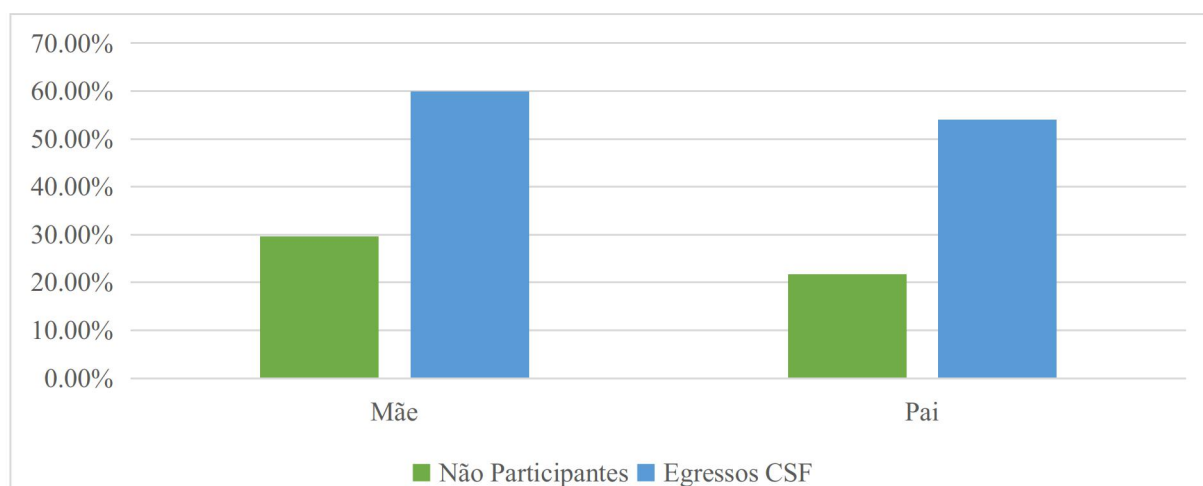
3.3 ESTATÍSTICAS DESCRITIVAS

Os dados são do Exame Nacional de Desempenho dos Estudantes (ENADE) referente ao ano 2019, que contém informações detalhadas quanto a nota dos estudantes, suas instituições de ensino e características socioeconômicas nas quais os egressos do CSF e os não participantes do programa diferem-se, como: cor, gênero, renda familiar, escolaridade dos pais, entre outras.

Observa-se que a proporção de pessoas do sexo masculino era superior dentre os egressos do CSF (61,04%) em relação aos não participantes (45,07%). Dentre os egressos do CSF e os não participantes do programa respectivamente, eram brancos (63,77% e 53,33%), pardos (23,51% e 33,89%), negros (5,27% e 8,21%), amarelos (2,69% e 2,45%), indígenas (0,21% e 0,32%) e não quiseram declarar ou não responderam (4,53% e 1,77%), observa-se dentre os egressos do CSF uma menor proporção de negros, pardos e indígenas e uma maior proporção de brancos em relação aos não participantes do programa.

Quanto à escolaridade dos pais, observa-se que 60,01% dos egressos do CSF eram alunos cuja mãe possuía ensino superior (graduação ou pós-graduação) e 54,05% cujo pai possuía ensino superior, enquanto que para os não participantes essa proporção era de apenas 29,59% e 21,7% respectivamente (gráfico 1).

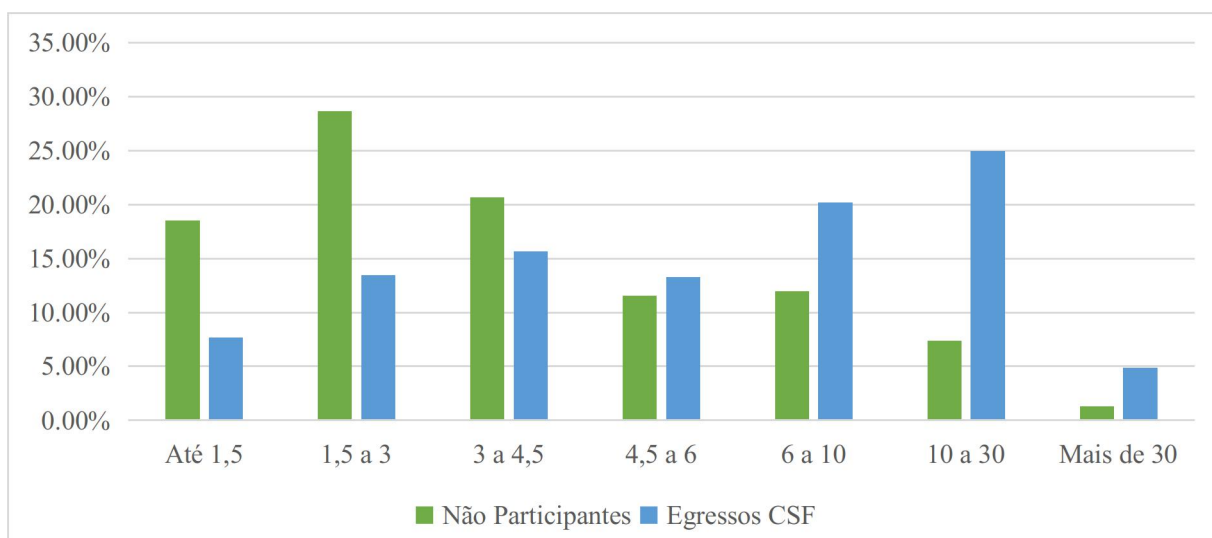
Gráfico 1 – Escolaridade dos pais dos não participantes e egressos do CSF (%)



Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

Conforme mostra o gráfico 2, aproximadamente 50% dos egressos do CSF estavam distribuídos entre os três níveis mais altos de renda familiar, enquanto 67,84% dos alunos não participantes estavam distribuídos entre os três níveis mais baixos de renda, ou seja, os egressos do CSF tinham em média, renda familiar superior aos demais.

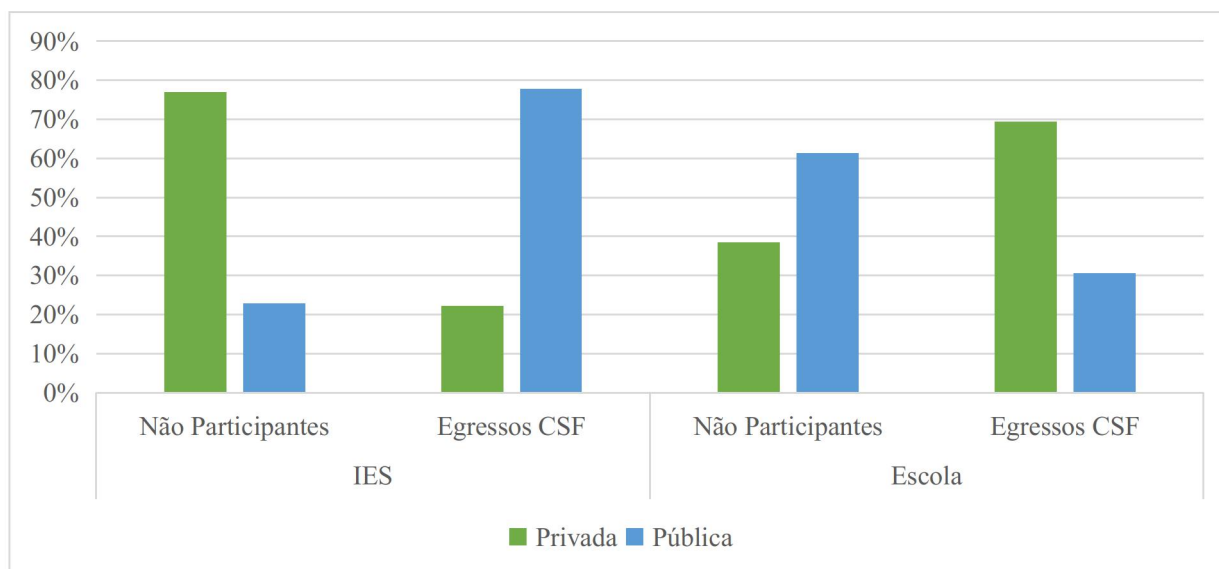
Gráfico 2 - Renda familiar dos não participantes e egressos do CSF (2019) (salários mínimos)



Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

Observa-se também que a maior parte dos egressos do CSF eram de IES públicas (77,8%), enquanto a maior parte dos não participantes eram de IES privadas (77,8%), uma proporção exatamente oposta (gráfico 3). A maioria dos egressos do CSF também cursou todo ou a maior parte do ensino médio em escola privada (69,4%), enquanto a maior parte dos não participantes cursou em escola pública (61,4%) (gráfico 3).

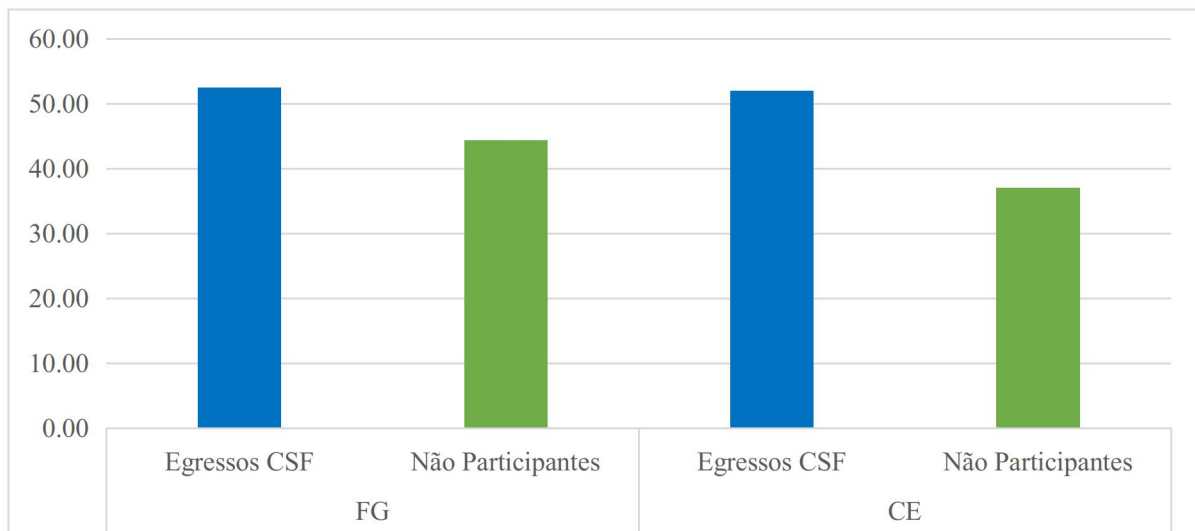
Gráfico 3 – Alunos de IES e escola pública e privada (não participantes e egressos do CSF) (%)



Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

A nota média dos egressos do CSF, em média, foi superior à dos não participantes no componente de Formação Geral (FG) do ENADE em 11,81 pontos com um Intervalo de Confiança (IC = 11,26 a 12,35) e no Componente Específico (CE) em 5,77 pontos (IC = 5,21 a 6,33) (gráfico 4).

Gráfico 4 – Nota média estimada dos não participantes e egressos do CSF no ENADE 2019 (pontos)



Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

Destaca-se que, os egressos do CSF e os não participantes do programa tinham características socioeconômicas muito distintas, sobretudo, que o grupo dos egressos era superior em proporção de brancos, escolaridade dos pais, renda familiar, tipo de IES e escola de origem. Sendo assim, uma comparação entre a nota dos alunos tal qual apresentada no gráfico 4 não reflete exclusivamente o efeito do programa, logo não é passível de inferência causal (HECKMAN et al., 1988; CAMERON e TRIVEDI, 2005).

Justifica-se, portanto, o uso da metodologia PSM, apresentada na seção 3.1, a fim de segregar os dados dos alunos de ambos os grupos de forma a equipará-los em termos de propensão a fazer parte do tratamento (programa CSF) com base em suas características observáveis. Desse modo, será então possível estimar de forma quantitativa o efeito do programa CSF sobre o desempenho dos alunos uma vez que esses sejam suficientemente semelhantes para que nenhum outro fator além da participação no programa adultere o resultado final estimado (CAMERON e TRIVEDI, 2005).

Por fim, destaca-se que tais dados e conseqüentemente tais resultados fazem referência a cursos das áreas da saúde, engenharias e ciências da natureza e da terra, são eles: medicina veterinária, odontologia, medicina, agronomia, farmácia, arquitetura e urbanismo, enfermagem, fonoaudiologia, nutrição, fisioterapia, zootecnia, biomedicina, tecnologia em radiologia, agronegócios, gestão hospitalar, gestão ambiental, estética e cosmética, segurança do trabalho, educação física (bacharelado), engenharia da computação, civil, elétrica, de controle e automação, mecânica, de alimentos, química, de produção, ambiental, e florestal.

4. DISCUSSÃO E RESULTADOS

A partir desses dados, foram estimados através da metodologia PSM e MQP os resultados dispostos nesta seção. Para fazer o pareamento dos indivíduos com base em suas características observáveis em comum para a seleção ao tratamento (participação no CSF), foram utilizados: uma variável *dummy* para alunos do gênero feminino (*feminino*), alunos de IES pública (*ies_publica*), alunos que cursaram todo ou a maior parte do ensino médio em escola pública (*escpub*), alunos da região sul ou sudeste (*sul_sudeste*), alunos cuja mãe possui ensino superior (graduação ou pós-graduação) (*mae_ensup*), e alunos com renda familiar

igual ou maior que 3 salários mínimos (*renda_mais_3*).

O quadro 1 mostra o número de blocos nos quais os não participantes e egressos do CSF foram segregados de acordo com o limite inferior do escore de propensão a serem elegíveis ao tratamento, a saber, 15 blocos com propensão entre 0,05 e 5%.

Quadro 1 – Blocos propensão à seleção ao tratamento (controle e egressos do CSF)

Propensão	Controle	CSF	Total
.000507	62246	74	62320
.0007813	84512	143	84655
.0015625	66640	125	66765
.003125	29402	83	29485
.0046875	26662	134	26796
.00625	50368	373	50741
.0125	1200	4	1204
.0140625	6759	103	6862
.015625	4991	82	5073
.01875	6195	102	6297
.021875	8547	135	8682
.025	9019	270	9289
.03125	4849	193	5042
.0375	17850	910	18760
.05	11329	810	12139
Total	390569	3541	394110

Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

Na tabela 1, tem-se o efeito médio estimado do programa CSF sobre o desempenho dos alunos na nota do componente de formação geral do ENADE 2019 (*nt_fg*) o qual foi de +4,532 pontos em média (ATT pelo critério *stratification*). O efeito médio do CSF se mostrou estatisticamente significativo em todos os critérios de pareamento utilizados.

Tabela 1 – Estimativa do efeito médio do CSF sobre a nota no componente FG do ENADE 2019

Critério de Pareamento	CSF	Controle	ATT	Erro-Padrão	t
<i>Nearest Neighbor</i>	3641	373907	4,511	0,274	16,437
<i>Stratification</i>	3641	404232	4,532	0,262	17,291

Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

A tabela 2 apresenta o efeito médio estimado do programa CSF sobre o desempenho dos alunos na nota do componente de formação específica do ENADE 2019 (*nt_ce*). Entretanto, a estimativa do efeito do programa CSF sobre a nota em tal componente não apresentou significância estatística em nenhum critério de pareamento utilizado.

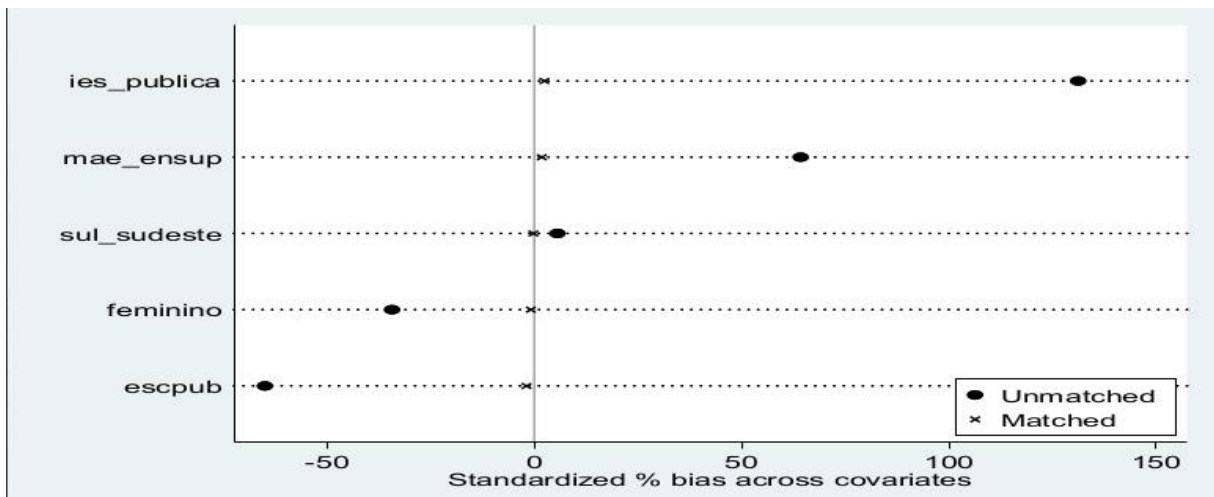
Tabela 2 – Estimativa do efeito médio do CSF sobre a nota CE do ENADE 2019

Critério de Pareamento	CSF	Controle	ATT	Erro-Padrão	t
<i>Nearest Neighbor</i>	3641	373907	-0,039	0,301	-0,131
<i>Stratification</i>	3641	404232	-0,056	0,294	-0,19

Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

O gráfico 5 apresenta uma comparação entre o grau de viés associado às variáveis utilizadas no modelo PSM antes (*unmatched*) e depois do pareamento dos alunos (*matched*), de modo que quanto mais próximo do eixo central, menor o viés associado a uma variável. Observa-se que o modelo PSM utilizado foi capaz de reduzir o viés associado a cada uma das variáveis utilizadas a zero ou a aproximadamente zero (gráfico 5).

Gráfico 5 – Grau de viés das variáveis do modelo PSM antes e depois do pareamento (%)



Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

Controlando características relacionadas a desempenho, tais como: variável *dummy* do aluno que leu mais de seis livros no ano do exame (*le_mais_6*), *dummy* do aluno que estudou mais de três horas por semana (*estuda_mais_3*) e uma *dummy* do aluno cujo pai possui ensino superior (*pai_ensup*), foi estimado, de modo duplamente robusto, o efeito médio do programa CSF sobre o desempenho dos beneficiários no ENADE por meio de uma regressão de Mínimos Quadrados Ponderados pelos escores de propensão do modelo PSM (MQP).

Conforme a tabela 3, tem-se que o efeito médio do programa CSF, estimado pelo modelo MQP, sobre o desempenho no componente de formação geral do ENADE é de +5,21 pontos em média, com intervalo de confiança cujo limite inferior é próximo às estimativas do modelo PSM. Todas as variáveis utilizadas apresentaram coeficientes positivos e significância estatística, indicando que: alunos que estudam mais de três horas por semana, alunos cujo pai possui ensino superior e alunos que leem mais de seis livros no ano têm desempenho maior em 4,96, 3,91 e 0,7 pontos em média respectivamente, em relação aos demais alunos.

Tabela 3 – Resultados da regressão de MQP sobre a nota de formação geral do ENADE 2019

Formação Geral (<i>nt_fg</i>)	Coefficiente	EP Robusto	t	P> t	[Intervalo de Confiança 95%]	
<i>csf</i>	5,21720	0,32473	16,07	0,000	4,58074	5,85659
<i>le_mais_6</i>	0,70712	0,14078	5,02	0,000	0,43119	0,98305
<i>estuda_mais_3</i>	4,96573	0,10223	48,58	0,000	4,76537	5,16609
<i>pai_ensup</i>	3,91644	0,10165	38,53	0,000	3,71721	4,11566
<i>constante</i>	41,69033	0,08654	481,75	0,000	41,52072	41,85995

Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

A tabela 4 mostra o efeito médio do programa CSF sobre o desempenho dos alunos na nota do componente específico do ENADE 2019, o qual também não apresentou significância estatística, assim como no modelo PSM. A variável *le_mais_6* apresentou significância somente a 5%, as variáveis *estuda_mais_3* e *pai_ensup* também apresentaram coeficientes positivos e significância a 1%, sugerindo desempenho maior em 4,6 e 3,74 pontos em média respectivamente, em relação aos alunos que estudam três horas ou menos por semana e aos alunos cujo pai não possui ensino superior, respectivamente.

Tabela 4 – Resultados da regressão de MQP sobre a nota no componente específico do ENADE 2019

Componente Específico (<i>nt_ce</i>)	Coefficiente	EP Robusto	t	P> t	[Intervalo de Confiança 95%]	
<i>csf</i>	0,10940	0,36831	0,30	0,766	-0,61248	0,83128
<i>le_mais_6</i>	0,31176	0,15135	2,06	0,039	0,01512	0,60840
<i>estuda_mais_3</i>	4,60277	0,10815	42,56	0,000	4,39079	4,81475
<i>pai_ensup</i>	3,74177	0,10918	34,27	0,000	3,52778	3,95577
<i>constante</i>	46,35284	0,09101	509,31	0,000	46,17446	46,53122

Fonte: Microdados do ENADE 2019 (INEP) | Nota: Elaborado pelo autor

5. CONCLUSÃO

Este trabalho teve como objetivo estimar o efeito do programa Ciência Sem Fronteiras sobre o desempenho de seus beneficiários no ENADE 2019. Ao longo do estudo, discorreu-se sobre elementos da teoria do capital humano, os retornos à educação, a inserção do programa CSF no contexto histórico educacional brasileiro, bem como a descrição dos dados e das metodologias utilizadas a fim de alcançar o objetivo principal deste trabalho. Desse modo, comenta-se a seguir os resultados obtidos pelo presente estudo.

As estimativas do modelo PSM, pelo critério *stratification*, sugerem que o efeito do CSF sobre a nota no componente de formação geral do ENADE é de +4,532 pontos em média, o que representa 13,48% da nota média do grupo de controle. Quanto ao componente de formação específica, o efeito estimado não se mostrou estatisticamente significativo, o que sugere que a probabilidade de o efeito do programa para com este componente ser diferente de zero não é significante.

Utilizando o modelo MQP para controlar características específicas de desempenho não utilizadas no modelo PSM, o efeito médio do programa sobre o desempenho de seus beneficiários no componente de formação geral do ENADE foi estimado em +5,21 pontos em média, 0,7 pontos maior do que as estimativas do modelo PSM. O efeito médio estimado para o componente específico também não apresentou significância estatística no modelo MQP assim como no modelo PSM.

Tais resultados sugerem que o programa CSF tem impacto significativo na capacitação dos alunos quanto aos conhecimentos gerais de seus respectivos cursos, porém, sem relação significativa com os conhecimentos específicos de suas áreas. Uma hipótese para esse indício poderia ser que a capacitação específica adquirida no CSF não é significativamente diferente da capacitação oferecida nas IES brasileiras ou que exista uma incompatibilidade entre os conhecimentos específicos lecionados nas diferentes instituições de ensino internacionais e os conhecimentos presentes no ENADE e nas IES brasileiras, conforme sugerido nos estudos de Schwartzman (1972) e Castro et al. (2012).

Vale ressaltar que os resultados obtidos por este trabalho são referentes a cursos de ciências da saúde, engenharias e ciências da natureza e da terra apenas, e que os efeitos estimados dizem respeito apenas ao desempenho dos estudantes, ou seja, suas capacidades cognitivas expressas em seus resultados no ENADE. Portanto, recomenda-se a futuros estudos estimar os efeitos do programa CSF sobre competências não-cognitivas ou socioemocionais que possam ser otimizadas pela experiência do programa, tais como: motivação, autoestima, raciocínio, criatividade, entre outras.

Destaca-se também, que apesar da intenção, devido à limitação dos dados disponíveis, não foi possível testar a sensibilidade dos resultados estimados quanto a efeitos temporais sobre o desempenho dos estudantes, associados a variáveis não-observáveis, permanecendo assim como um ponto em aberto e passível de futuras abordagens.

Mediante o exposto, coloca-se em questão que a superioridade de desempenho expressa na diferença de notas entre os não participantes e egressos do CSF parece estar mais associada a características socioeconômicas privilegiadas dos participantes do programa do que ao efeito deste propriamente dito, conforme abordado nas seções 3.3 e 4.

Sobretudo, destaca-se o efeito positivo estimado do programa sobre o desempenho dos alunos no componente de formação geral do ENADE, as melhorias qualitativas e quantitativas apontadas na literatura de avaliações do programa, os possíveis efeitos positivos do programa não abordados por este estudo e a importância de políticas públicas semelhantes no aprimoramento da educação brasileira e no combate às desigualdades econômicas e sociais.

REFERÊNCIAS

AZEVEDO, Márcio; BARRETO, Renata; BEZERRA, Marina Gabrielle; SOCOLOSKI, Mariana; HERMINIO, Pablo. **Análise Política Quanto à Eficiência do Programa Ciência Sem Fronteiras: Relatos IFRN e UFRN.** [S.l.] 2013.

BARRO, Robert J.; LEE, Jong-Wha. **A New Data Set of Educational Attainment in the World, 1950-2010.** Cambridge: Massachusetts, 2010.

BECKER, Gary S. **Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education.** 3rd edition. University of Chicago Press, 1964.

BECKER, Gary S; MURPHY, Kevin M; TAMURA, Robert F. **Human Capital, Fertility and Economic Growth.** Cambridge: Massachusetts, 1990.

BLAUG, Mark. **The Empirical Status of Human Capital Theory: A Slightly Jaundiced Survey.** [S.l.], 1976.

BRASIL. **Ciência Sem Fronteiras: Um Programa Especial de Mobilidade Internacional em Ciência, Tecnologia e Inovação.** Brasília, 2011.

BRASIL. **Relatório nº 21 - CCT de 2015.** Senado Federal: Brasília, 2015.

CAMERON, Colin A.; TRIVEDI, Pravin K. **Microeconometrics: Methods and Applications.** Cambridge: Massachusetts, 2005.

CASTRO, Claudio de Moura; BARROS, Hélio; ITO-ADLER, James; SCHWARTZMAN, Simon. **Cem Mil Bolsistas no Exterior.** [S.l.] 2012.

CONCEIÇÃO, Otávio Canozzi. **Uma Avaliação do Programa Ciência Sem Fronteiras: Efeitos Sobre Desempenho, Trancamento e Conclusão.** Porto Alegre, 2017.

CONCEIÇÃO, Otávio Canozzi; FONTES, Luiz Felipe Campos; SARAIVA, Maurício Vitorino. **Três Anos do Programa Mais Médicos: Uma Análise Econométrica.** Pontifícia Universidade Católica do Rio Grande do Sul, 2016.

CONCEIÇÃO, Otávio 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.** Porto Alegre, 2016.

CUNHA, Dilene Amaral da. **Ciência Sem Fronteiras: Perspectivas da Internacionalização e a Experiência Australiana.** Porto Alegre, 2016.

DEHEJIA, Rajeev H.; WAHBA, Sadek. **Propensity Score Matching Methods for Nonexperimental Causal Studies.** *The Review of Economics and Statistics*, 2002.

DENISON, Edward. **The Sources of Economic Growth in the United States and the Alternatives Before Us.** *Committee for Economic Development*, 1962.

FAGUNDES, Caterine; LUCE, Maria Beatriz; SILVEIRA, Paloma Dias. **A Qualidade da Mobilidade de Estudantes de Graduação no “Ciência Sem Fronteiras”**. Rio de Janeiro, 2019.

HAUSMANN, Ricardo; BUSTO, Sebastian; CHUNG, Sarah; COSCIA, Michele; HIDALGO, Cesar; JIMENEZ, Juan; SIMOES, Alexander; YILDIRIM, Muhammed. **The Atlas of Economic Complexity**. Cambridge: Massachusetts, 2011.

HECKMAN, James; ICHIMURA, Hidehiko; SMITH, Jeffrey; TODD, Petra. **Characterizing Selection Bias Using Experimental Data**. *Econometrica*, 1998.

HILL, Jennifer; REITER, Jerome P. **Interval Estimation for Treatment Effects Using Propensity Score Matching**. *Statistics in Medicine*, 2006.

IMBENS, Guido M.; WOOLDRIDGE, Jeffrey M. **Recent Development in the Econometrics of Program Evaluation**. Cambridge: Massachusetts, 2008.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Evolução do Ensino Superior - Graduação - 1980-1998**. Ministério da Educação: Brasília, 2000.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo da Educação Superior 2017**. Ministério da Educação: Brasília, 2017.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Estimativa do Investimento Público Direto em Educação por Estudante, por Nível de Ensino - Valores Nominais - Brasil 2000-2017**. Ministério da Educação: Brasília, [2017].

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo da Educação Superior: Notas Estatísticas 2019**. Ministério da Educação: Brasília, 2019.

MAITRA, Biswajit. **Investment in Human Capital and Economic Growth in Singapore**. University of Gour Banga: West Bengal, 2016.

MARI, Cezar Luiz De; THIENGO, Lara Carlette. **Ciências e Políticas: Análise do Programa Ciência Sem Fronteiras a Partir da Perspectiva Gramsciana**. Dourados, 2014.

MENEZES-FILHO, Naercio Aquino. **A Evolução da Educação no Brasil e seu Impacto no Mercado de Trabalho**. Universidade de São Paulo, 2001.

MERK, Ministry of Education of the Republic of Korea. **Education, the Driving Force for the Development of Korea**. *Korean Education Centre UK*: London, 2015.

MINCER, Jacob. **On-the Job Training: Costs, Returns and Some Implications**. Columbia University, 1962.

OLINDA, Sílvia Rita Magalhães de. **A Educação no Brasil no Período Colonial: Um Olhar Sobre as Origens Para Compreender o Presente**. Feira de Santana, 2003.

PEREIRA, Vânia Martins. **Relatos de Uma Viagem: Uma Análise Feita Pelos Bolsistas Sobre o Programa Ciência Sem Fronteiras.** Revista Perspectivas do Desenvolvimento: Um Enfoque Multidimensional, 2015.

PLANALTO, Governo Federal. **Constituição da República Federativa do Brasil de 1988.** Casa Civil da Presidência da República: Brasília, 1988.

PLANALTO, Governo Federal. **Decreto nº 7.642, de 13 de dezembro de 2011.** Casa Civil da Presidência da República: Brasília, 2011.

PLANALTO, Governo Federal. **Lei nº 9.394, de 20 de dezembro de 1996.** Casa Civil da Presidência da República: Brasília, 1996.

ROSENBAUM, Paul R.; RUBIN, Donald B. **The Central Role of the Propensity Score in Observational Studies for Causal Effects.** *Biometrika*, 1983.

SCHULTZ, Theodore W. **Investment in Human Capital.** [S.l.] 1961.

SCHWARTZMAN, Simon. **Os Desafios da Educação no Brasil.** Rio de Janeiro, 2004.

SCHWARTZMAN, Simon. **Projeto Retorno - Avaliação do Impacto do Treinamento, no Exterior, de Pessoal Qualificado.** [S.l.], 1972.

TOPEL, Robert. **Labor Markets and Economic Growth.** University of Chicago, 1988.