

# Política Afirmativa e *Overeducation*: Uma Análise Para Egressos da UFPB

Francisco Danilo da Silva Ferreira <sup>1</sup> Aléssio Tony Cavalcante Almeida <sup>2</sup>

Ana Cláudia Annegues <sup>3</sup>

## Resumo

O artigo examinou a probabilidade que egressos cotas da educação superior possam estar em situação de *overeducation*, isto é, trabalhando em ocupações que exigem menor grau de escolaridade. O alcance do objetivo se deu pela associação do balanceamento por entropia com o *Propensity Score Matching* (PSM). Para tanto, consideram-se os dados dos alunos egressos de cursos da graduação da UFPB, cuja entrada no curso ocorreu no período de 2010 e 2011 e a conclusão até 2017, conjuntamente com informações do mercado formal de trabalho para o ano de 2018. Os resultados indicam para o conjunto de cursos da UFPB que a reserva de vagas não apresenta efeitos sobre a probabilidade dos egressos cotistas possam ser *overeducation*. Por outro lado, os ex-alunos cotistas, que se encontram no topo da distribuição de salários, bem como aqueles egressos cotistas de Ciências Humanas e Ciências Sociais Aplicadas apresentam maior probabilidade de ocupar postos de trabalho que não necessitam de nível superior, isto é, são potencialmente subutilizados no mercado de trabalho.

**Palavras-chave:** Política Afirmativa. Balanceamento por Entropia. PSM. *Overeducation*.

## Abstract

The paper analyzed the probability that higher education quota former students can be in overeducation situation, that is, working in occupations that require lower level of education. The reach of the objective gave up for the entropy balancing with Propensity Score Matching(PSM) association. Therefore, it considered the UFPB graduates courses former students data, whose entrance in the curso occurred in the period 2010 and 2011 and the completion until 2017, jointly with formal labor market data to the year 2018. The results indicate to the UFPB courses set that the vacancies reservation does not present effects on the probability of the quota former students can be overeducation. On the other hand, the quota former students, that which meet in the top in the wage distribution, as well as those ones quota former students of the humanities sciences and applied social sciences present most probability to occupy labor post that are not need higher education, that is, the are potentially underused in the labor market.

**Keywords:** Affirmative Policy. Entropy Balancing. PSM. *Overeducation*.

Área 12 - Economia Social e Demografia Econômica

**JEL classification:** C54, J20, I23

<sup>1</sup> Doutorando em Economia Aplicada, Programa de Pós-Graduação em Economia, Universidade Federal da Paraíba, Professor do Departamento de Economia da UERN/Campus Assú, ffdanilloferreira@gmail.com

<sup>2</sup> Doutor em Economia Aplicada, Professor do Programa de Pós-Graduação em Economia, Universidade Federal da Paraíba, E-mail: alessio@ccsa.ufpb.br

<sup>3</sup> Pós-doutora em Economia Aplicada - UFRN, Economista da UFPB, annegues.ana@gmail.com

# 1 Introdução

Nas últimas décadas, o Brasil vivenciou a expansão do ensino superior público e privado, mediante políticas e programas educacionais, a exemplo do Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI), Fundo de Financiamento Estudantil (FIES), Programa Universidade para Todos (PROUNI), bem como a oferta de graduação pelos Institutos Tecnológicos (IFs), ampliando, assim, a oferta de trabalhadores qualificados. Entretanto, há evidências que a oferta de vagas no mercado de trabalho não acompanhou a ampliação do número de graduados (SCHWARTZMAN, 2004), produzindo chamado *overeducation*.

O referido fenômeno é caracterizado pela existência de trabalhadores em ocupações que exigem um nível de escolaridade inferior a que eles efetivamente dispõem. Embora as discussões teóricas sobre *overeducation* tenham se iniciado em meados dos anos 70, McGuinness (2006) aponta que as literaturas teórica e empírica ainda não apresentam consenso quanto as causas do mesmo. Alguns autores sugerem que o excesso de educação reflete o desequilíbrio ou a ineficiência no mercado de trabalho e no sistema educacional, isto é, há uma oferta excessiva de trabalhadores qualificados frente a uma pequena demanda do mercado de trabalho (MCGUINNESS, 2006), (LEUVEN; OOSTERBEEK, 2011). Já outros argumentam que o fenômeno representa uma fase temporária e necessária para que alguns trabalhadores mostrem suas habilidades e ascendam futuramente na carreira (SICHERMAN, 1991).

Uma das consequências do excesso de educação é que este impõe custos, não só ao trabalhador, como também para a economia de forma geral. Os indivíduos sobre-educados são penalizados com salários mais baixos em comparação aos trabalhadores que possuem o nível de instrução adequando a ocupação que desempenham. Da mesma forma, há redução da satisfação do trabalhador, produzindo desmotivação do mesmo em aprimorar sua qualificação posto que o investimento em educação não gera o retorno esperado. Assim, segundo Groot (1996) há evidências para sugerir que os trabalhadores *overeducated* são relativamente improdutivos e possuem uma maior taxa de rotatividade (SLOANE; BATTU; SEAMAN, 1999)

Além disso, o excesso de educação conduz a economia a operar abaixo do seu potencial, dado que as habilidades dos trabalhadores sobre-educados são subutilizadas (MCGUINNESS, 2006). Acrescenta-se ainda que o *overeducation* submete a sociedade a elevados custos, tendo em vista o uso do dinheiro público para qualificar tais trabalhadores, levando a subutilização dos gastos públicos em capital humano.

Em paralelo à expansão do ensino superior houve a adoção de políticas afirmativas para ingresso nas universidades públicas, com reserva de vagas a minorias raciais (negros, pardos e indígenas) ou a alunos de escolas públicas (cotas sociais). As universidades estaduais do Rio de Janeiro<sup>4</sup> foram as primeiras a adotar a política de cotas, no início dos anos 2000. Inicialmente foram implementadas cotas para estudantes da rede pública e em seguida estas foram estendidas a candidatos autodeclarados negros e pardos. Posteriormente, a política ganhou espaço e se consolidou, com diversas instituições públicas estaduais e federais adotando por leis ou resoluções de política de cotas.

O objetivo central da política de cotas consiste em corrigir injustiças históricas e romper os ciclos de perpetuação da pobreza e das desigualdades por meio do acesso ao ensino superior, o qual costuma estar associado a ocupações com maiores ganhos salariais no mercado de trabalho, fornecendo um caminho para empregos melhores ou de *status* ocupacionais mais elevados com a ascensão dos níveis de educação. Contudo, o debate sobre os reais benefícios e custos da política de cotas é controverso e a literatura econômica do tema não é unânime quanto aos impactos na qualidade do ensino nas universidades e na inserção dos egressos beneficiários no mercado de trabalho. Para alguns, a política permitiria o acesso de indivíduos que não teriam chances de

<sup>4</sup> Universidade Estadual do Rio de Janeiro (UERJ) e Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF)

ingressar por ampla concorrência (ESTEVAN; GALL; MORIN, 2018), por outro lado, a política de cotas não seria suficiente para promover igualdade de oportunidades, tendo em vista que o problema estaria no ensino básico deficitário e assim deveria haver políticas complementares (LOPES, 2017).

Desse modo, o ensino universitário não seria suficiente para compensar as diferenças de habilidades dos cotistas com os demais pares, estando estes em posição inferior no mercado de trabalho, não apenas em termos salariais, mas também em relação às ocupações, Battu e Sloane (2002) aponta que as minorias tendem a ser mais afetados pelo *overeducation*. Assim, uma porcentagem significativa de cotistas com excesso de educação implicaria que as políticas de ação afirmativa podem não ser bem sucedidas quanto ao seu objetivo de promover mobilidade social às minorias e igualdade no mercado de trabalho.

Diante o exposto, o presente artigo tem por objetivo investigar a influência da política de cotas adotada pela UFPB sobre a probabilidade do egresso beneficiário da política ser *overeducated*. Assim, para alcançar o objetivo proposto, associou-se o balanceamento por entropia com o *Propensity Score Matching* (PSM). O primeiro permite implementar uma reponderação para ajustar a distribuição amostral antes de realizar o pareamento, no intuito de alcançar grupos mais equilibrados. O segundo se traduz em um método não-experimental de *matching*, o qual mediante critérios estatísticos e características observáveis viabiliza a construção artificial do melhor grupo de comparação para os tratados. Por último, utilizaram-se os limites de Rosenbaum (*Rosenbaum bounds*) para verificar a sensibilidade do efeito médio do tratamento sobre os tratados em relação a fatores omitidos.

A estratégia acima é viabilizada através da utilização de duas fontes de dados. A primeira se refere aos dados da UFPB, o qual fornece informações sobre a trajetória dos egressos durante a graduação. Apesar de a instituição ter implantado sua política de cotas em 2011, o presente trabalho também utiliza informações de 2010 com a finalidade de minimizar o viés de autosseleção, para tanto os egressos de 2010 farão parte do grupo de controle. Já as informações do mercado de trabalho são provenientes da Relação Anual de Informações Sociais (RAIS) identificada de 2018, que através do CPF permitiu mapear a situação do egresso no mercado de trabalho formal.

A literatura nacional sobre as políticas afirmativas é ampla, mas se restringe aos seus efeitos no desempenho e o esforço dos beneficiários na graduação (PEREIRA; BITTENCOURT; JUNIOR, 2013), (PEIXOTO et al., 2016), assim como sobre as chances de ingresso e evasão no ensino superior, (JUNIOR; SOUZA; WALTENBERG, 2016), Estevam et al. (2018), (SILVA et al., 2019). Quanto ao *overeducation* tem se focado na área de formação (ANNEGUES et al., 2018), e nos efeitos sobre os rendimentos (CAVALCANTI et al., 2020), (ANNEGUES; SOUZA, 2020). Todavia, não foi encontrado evidências empíricas a respeito dos efeitos das políticas afirmativas sobre a probabilidade do indivíduo ser sobre-educado. Assim, o artigo tenta preencher a lacuna da literatura nacional em relação ao desequilíbrio entre a realização educacional e ocupacional de egressos que usaram políticas afirmativas, bem como fornecer subsídios aos formuladores de políticas educacionais.

O presente trabalho, além dos aspectos introdutórios, é composto por quatro seções. Na sequência é apresentado a literatura sobre *overeducation*. A terceira seção trata das particularidades metodológicas. Posteriormente é discutido os resultados. Na última seção são apresentados as principais considerações em relação aos resultados obtidos.

## 2 Revisão de Literatura

### 2.1 *Overeducation*

A literatura interpreta o *overeducation* como o excesso de escolaridade do trabalhador em relação à exigida pela sua ocupação atual. Contudo, a mesma não apresenta unanimidade quanto

a teoria que explica o referido fenômeno. À vista disto, os achados empíricos têm sido analisados segundo algumas teorias do mercado de trabalho, a exemplo da teoria do capital humano, os modelos de competição por emprego, os *assignment models* e as teorias sobre mobilidade ocupacional.

Para a teoria do capital humano o problema de *overeducation* seria um fenômeno temporário derivado do desequilíbrio entre oferta e demanda no mercado de trabalho (TSANG; LEVIN, 1985). Conforme a referida teoria haverá *overeducation* ao longo do processo de ajustamento do mercado, ou seja, o intervalo temporal necessário para que o trabalhador encontre uma ocupação pertinente ao seu nível educacional ou o tempo que as empresas precisam para promoverem o ajuste no processo produtivo o adequando ao capital humano disponível no mercado de trabalho.

Ressalta-se que o excesso de educação não é visto por outras teorias como um fenômeno transitório, ou seja, de curto prazo, como sugerido pela teoria do capital humano. Por seu turno, os modelos de competição por emprego pressupõem que o *overeducation* é produto da competição dos indivíduos por postos de trabalho. Segundo o modelo de competição de Thurow e Lucas (1972), há uma fila de indivíduos no mercado de trabalho competindo por ocupações, sendo as posições determinadas pelo nível de escolaridade dos mesmos. Neste sentido, para manter ou progredir sua posição na fila o indivíduo se qualifica cada vez mais, o que resultara em excesso de educação em relação a exigida pelo cargo.

Nos modelos de designação (*assignment models*) a economia enfrenta um problema de alocação entre trabalhadores heterogêneos e ocupações de diferentes complexidades (SATTINGER, 1993). O processo de alocação não é uma loteria, dado que as escolhas dos trabalhadores quanto a ocupação são orientadas pela maximização da utilidade e os salários são determinados pelas características dos trabalhadores e dos empregos. Contudo, o equilíbrio no mercado de trabalho não necessariamente deve ser de compatibilidade entre trabalhadores alocados a uma determinada ocupação e a qualificação requerida para tal. Neste sentido, alguns indivíduos estarão em situação de *overeducation*, ou seja, sua qualificação está acima do exigido para sua atual ocupação.

As teorias de mobilidade de carreira são outras alternativas de explicação do *overeducation* (SICHERMAN, 1991), as mesmas sugerem que os trabalhadores aceitam ocupações inicialmente incompatíveis com seu nível educacional, mas que lhes permita adquirir habilidades necessárias, através de treinamentos e experiência profissional, permitindo ao indivíduo alcançar promoções e ascensão da carreira na empresa atual ou em outra do mercado.

Apesar de a literatura apresentar muitos trabalhos sobre *overeducation*, as evidências em relação às minorias são escassas. Entre os primeiros estudos destaca-se o realizado por Duncan e Hoffman (1981), que utilizaram uma amostra dos trabalhadores americanos e investigaram os ganhos econômicos do *overeducation*. Os autores evidenciam que 50% dos homens negros dos EUA têm mais educação que o necessário para a sua ocupação, em comparação aos 40% do total de trabalhadores americanos. Contudo, os mesmos encontram efeitos positivos do excesso de educação para todos os grupos, mas o retorno é a metade do ganho adicional de um ano de educação necessária para a ocupação.

Alpin, Shackleton e Walsh (1998) apresenta evidências para a Grã-Bretanha, examinando a subutilização dos trabalhadores graduados. Para tanto, os autores utilizam dados da *Labour Force Survey data* de 1995. Alpin, Shackleton e Walsh (1998) apontam diversos fatores que aumentam a probabilidade do indivíduo ser *overeducated*, um deles é ser de grupos minoritários, evidenciando que trabalhadores não brancos são mais propensos a ter excesso de educação, em torno de 29,9%, em comparação aos trabalhadores brancos (26,8%). Alpin, Shackleton e Walsh (1998) ainda destaca que o problema de *overeducation* seria temporário para muitos indivíduos, pois à medida que os mesmos adquirem habilidades e experiência relacionadas a ocupação, eles são capazes de melhorar sua correspondência profissional.

Já Blackaby et al. (1999) investigam a perspectiva de emprego de grupos minoritários no Reino Unido com base em um painel (1987-1991) de microdados do *Labour Force Survey*. Os



autores mostram que o problema do excesso de educação entre as minorias é fruto da distância em relação ao local de trabalho, reduzindo as chances de melhores correspondências entre nível educacional possuído e o exigido pela ocupação, elevando o excesso de educação entre as minorias.

Assim como Blackaby et al. (1999), Battu e Sloane (2002), e Battu e Sloane (2004) realizaram estudos para o Reino Unido. Desse modo, a partir da Pesquisa Nacional Britânica de Minorias Étnicas, Battu e Sloane (2002) investigam se as minorias étnicas sofrem de maior excesso de educação, frente aos trabalhadores brancos em função de uma possível discriminação. Os resultados encontrados mostram que em média os trabalhadores não brancos apresentam mais de 30% de excesso de educação, já para os brancos esse valor fica em torno dos 20%. Os autores ainda destacam a tendência das minorias estrangeiras serem menos recompensadas em caso de *overeducation*.

Battu e Sloane (2004) ampliam o estudo realizado em 2002 e examinam a incidência de incompatibilidade entre o nível educacional e a realização ocupacional sobre os ganhos e outros resultados do mercado de trabalho. Battu e Sloane (2004) encontram que os grupos étnicos têm níveis variados de excesso de educação com maior incidência entre os grupos indianos e afro-asiáticos. Quando os autores introduzem variáveis de controles nas equações, os mesmos encontram que os afro-asiáticos são mais propensos a ter mais educação em relação aos indianos. Além disso, os autores destacam que ter um diploma estrangeiro, ser nascido no Reino Unido e ter fluência no idioma aumentam a probabilidade do indivíduo não branco ser supereducado ou subeducado, indo de encontro aos argumentos da teoria da assimilação.

Por sua vez, Lindley (2009) explora a incidência de *Overeducated* e *Undereducated*, bem como o efeito sobre os rendimentos de imigrantes, nativos e grupos étnicos minoritários por sexo. Para tal, o autor emprega informações da Pesquisa Trimestral da Força de Trabalho de 1993 a 2003 do Reino Unido. Os achados mostram que, em comparação com brancos, negros africanos, outros não-brancos e indianos, os homens são mais propensos a ter excesso de educação, enquanto para as mulheres é mais provável que indianas e paquistaneses/bangladeshianas sejam *overeducated*. No entanto, Lindley (2009) aponta que há grandes retornos às habilidades ocupacionais para alguns grupos étnicos e imigrantes minoritários.

Rafferty (2012) também realiza uma análise para o Reino Unido examinando as diferenças étnicas no excesso de educação, desemprego e salários como marcadores potenciais de discriminação ou "penalidades étnicas", definidas como as diferenças nos resultados do mercado de trabalho que persistem após contabilizar o capital humano observável e as características demográficas. Para estimar as referidas penalidades étnicas, Rafferty (2012) usa a abordagem de *matching* de covariável. Os resultados revelam que os níveis de penalidades de realização educacional são persistentes entre vários grupos étnicos minoritários. Rafferty (2012) ressalta que as desvantagens sociais anteriores ao mercado de trabalho explicam os padrões de *overeducation*.

Já Clark, Joubert e Maurel (2017) são os primeiros a examinarem a dinâmica de carreira de trabalhadores em situação *overeducated* em relação aos seus pares na mesma ocupação e seus efeitos sobre os salários. Para tanto, os mesmos usam dados longitudinais do *National Longitudinal Surveys* (NLSY79) combinados com o *Current Population Survey* (CPS) para os Estados Unidos. Os autores evidenciaram que o *overeducation* não é apenas mais comum, mas também mais persistente entre os trabalhadores negros e indivíduos com AFQT baixo, indo de encontro as ideias de Alpin, Shackleton e Walsh (1998). Além disso, a taxa de risco de excesso de educação cai em cerca de 60% durante os primeiros 5 anos com educação excessiva. Por último, Clark, Joubert e Maurel (2017) destacam que o *overeducation* está associado a salários atuais e futuros significativamente mais baixos, e que o desemprego passado está associado a uma maior duração do *overeducation* no futuro.

## 3 Metodologia

### 3.1 Estratégia Empírica

Com o objetivo de avaliar os efeitos da política de cotas sobre a probabilidade dos egressos beneficiários da UFPB ser *overeducation*, o presente estudo lançou mão do *Propensity Score Matching* (PSM) associado ao balanceamento por entropia. Logo após, realizou-se a análise de sensibilidade, proposta por Rosenbaum (2002), para verificar se o viés de variáveis omitidas afetou os resultados do Efeito Médio de Tratamento sobre os Tratados (ATT).

#### 3.1.1 Balanceamento por Entropia

Proposto por Hainmueller (2012), o balanceamento por entropia pode ser entendido como uma generalização da abordagem convencional do *propensity score*, porém, com procedimentos metodológicos distintos. Enquanto que no pareamento por escore de propensão calcula-se o *score* por intermédio de uma regressão (*logit/probit*), para logo após realizar o pareamento das unidades, e verificar se os pesos estimados equilibram as distribuições das covariadas, na entropia é permitido estimar os pesos diretamente para equalizar as distribuições das amostras, isto é, o método de entropia tem a vantagem de implementar diretamente o equilíbrio exato, bem como é duplamente robusto em relação à regressão linear segundo Zhao e Percival (2016).

Segundo Hainmueller (2012), a entropia permite ao pesquisador fixar um equilíbrio mais preciso estabelecendo um conjunto de restrições de equilíbrio, o que implica que os momentos (média, variância e assimetria) da amostra no grupo de controle ponderado equivalem exatamente aos momentos correspondentes no grupo de tratamento. Desse modo, de acordo com Hainmueller (2012), torna-se dispensável a verificação do equilíbrio no sentido convencional, pelo menos nos momentos incluídos nas restrições do equilíbrio.

Em contraste com outros métodos de pré-processamento, o balanceamento por entropia assegura alto equilíbrio das covariáveis entre os grupos de tratamento e controle, mesmo em pequenas amostras, dado que não exige que uma unidade não correspondida seja descartada, como no *Coarsened Exact Matching* (CEM). Assim, um grupo de controle sintético é projetado para representar uma imagem virtualmente perfeita do grupo de tratamento sem perdas de observações, garantindo o máximo de informações.

Considere  $w_i$  o peso do balanceamento por entropia escolhido para cada unidade do grupo de controle, encontrado pelo seguinte esquema de reponderação que minimiza a distância métrica de entropia.

$$\min_{w_i} H(w) = \sum_{\{i|D=0\}} w_i \log(w_i/q_i) \quad (1)$$

Sujeito as restrições de equilíbrio e normalização.

$$\sum_{\{i|D=0\}} w_i c_{ri}(X_i) = m_r \quad \text{Com } r \in 1, \dots, R \quad (2)$$

$$\sum_{\{i|D=0\}} w_i = 1 \quad (3)$$

$$w_i \geq 0 \text{ Para todo } i \text{ tal que } D = 0 \quad (4)$$

Em que  $q_i = 1/n$  é um peso base amostral e  $c_{ri}(x_i) = m_r$  descreve um conjunto de  $R$  restrições de equilíbrio impostas aos momentos das covariáveis no grupo de controle reponderado.

Operacionalmente, após a escolha da variável que será incluída na reponderação, é especificado um conjunto de restrições de balanceamento para equalizar os momentos das

distribuições das covariáveis entre grupos de tratamento e controle ponderados. Desse modo, o método de entropia procura para o conjunto de observações do controle, pesos  $w = [w_1, \dots, w_{n_0}]$  que minimizam a equação 1, a distância de entropia entre  $w$  e o vetor base de pesos  $Q = [q_1, \dots, q_{n_0}]$ , sujeita as restrições de balanceamento (equação 2), restrição de normalização (Equação 3) e restrição de não negatividade (Equação 4).

O balanceamento por entropia foi desenvolvido a princípio para avaliar os efeitos do tratamento em dados observáveis, porém, conforme [Watson e Elliot \(2016\)](#) o referido método também pode ser aplicado ao objetivo de ajustar a amostra para estimativas subsequentes dos efeitos do tratamento. Segundo [Hainmueller \(2012\)](#), a entropia pode ser facilmente combinado com quase qualquer estimativa (padrão) que o pesquisador queira utilizar para modelar o resultado nos dados pré-processados.

Assim, para atingir o objetivo proposto, o presente estudo segue a estratégia proposta por [Hainmueller \(2012\)](#) e [Watson e Elliot \(2016\)](#) usando os pesos alcançados pelo balanceamento por entropia com o *Propensity Score Matching* (PSM), apresentado na subseção a seguir.

### 3.2 *Propensity Score Matching*

Seguindo a estratégia adotada, o presente artigo recorreu ao *Propensity Score Matching* (PSM) após o balanceamento por entropia. O referido método busca criar um grupo de controle estatisticamente semelhante com base em um vetor de características observáveis, representando o contrafactual do grupo de tratamento, neste caso os egressos que não utilizaram a política de cotas.

A princípio, o PSM foi introduzido por [Rosenbaum e Rubin \(1983\)](#), no intuito de solucionar o problema da dimensionalidade, em que a inclusão de mais características torna difícil a ocorrência do *matching*. Assim, as unidades de comparação são aqueles cujos escores de propensão são próximos da unidade tratada, e não exatamente igual ([CAMERON; TRIVEDI, 2005](#)).

O escore de propensão, por sua vez, é definido como a probabilidade condicional de receber o tratamento dado um vetor de características observáveis  $X$ . Esse escore pode ser estimado por uma regressão *logit* ou *probit*. Nessa regressão, a variável explicativa é uma *dummy*, que assume valor um caso a unidade receba o tratamento e zero caso contrário. O condicionamento no escore de propensão é geralmente mais simples do que o condicionamento em uma grande dimensão de características. Formalmente temos:

$$Pr[D_i = 1 | X_i] = G(X_i\beta + \mu_i) \quad (5)$$

Cabe destacar, que a estimação por PSM se fundamenta em duas hipóteses-chave. A primeira, diz respeito a ignorabilidade, também conhecida como hipótese de independência condicional, nessa hipótese assumimos que os resultados potenciais são independentes do *status* de tratamento, condicional a  $X$ :

$$(Y_{i0}, Y_{i1}) \perp D_i | X_i \quad (6)$$

No entanto, é possível que uma unidade tratada não encontre seu correspondente no grupo de não tratados, baseado no *propensity score* estimado. Neste caso, temos a segunda hipótese, denominada de suporte comum entre os dois grupos. A referida hipótese implica que para cada unidade tratada tenha correspondente no grupo de controle para cada característica  $X$  para a qual deseja-se comparar.

$$0 < Pr(D_i = 1 | X_i) < 1 \quad (7)$$

A partir das hipóteses acima, os indivíduos são pareados conforme suas características observáveis disponíveis, viabilizando as estimações do efeito médio do tratamento sobre os

tratados (ATT) sobre a variável de resultado, o qual é obtido pela diferença nos resultados médios entre os grupos de tratamento e controle, como a seguir:

$$\begin{aligned}\hat{\tau}_{ATT} &= E[Y_{i1} - \hat{Y}_{i0} \mid D_i = 1] \\ &= E[Y_{i1} \mid D_i = 1] - E[\hat{Y}_{i0} \mid D_i = 1]\end{aligned}\quad (8)$$

Em que o primeiro termo representa a chance de ser *overeducation* dos que participaram da política de cotas, já o segundo equivale ao contrafactual dos indivíduos tratados. Ressalta-se ainda, que o artigo utiliza três diferentes técnicas de pareamento. O primeiro corresponde ao método do vizinho mais próximo (*Nearest Neighbor Matching*), com e sem reposição e *caliper*, bem como considerando um e dez vizinhos, o mesmo seleciona as melhores correspondências do controle para cada indivíduo no grupo de tratamento. O segundo é o método de *Matching Kernel*, onde cada unidade tratada  $i$  é pareada com várias observações  $j$ , com pesos inversamente proporcionais à distância entre a unidade tratada e de controle. O terceiro é o *Matching Radius*, onde para cada unidade tratada  $i$  é pareada com a unidade  $j$  que está contida em um determinado raio.

### 3.3 Análise de Sensibilidade

Como definido na subseção 3.2, o *Propensity Score Matching* fundamenta-se em duas hipóteses, entre estas a de independência condicional (CIA), a mesma preconiza que, dado o vetor de covariáveis observáveis os resultados potenciais são independentes da designação do tratamento. Contudo, os estimadores do efeito médio do tratamento podem ser inconsistente, caso existam fatores/variáveis não observáveis que afetam o grupo de tratamento, como também a variável de resultado.

Diante disto, Rosenbaum (2002) propôs uma análise de sensibilidade, chamada de limites de Rosenbaum (*Rosenbaum bounds*), o mesmo avalia o quão danoso é a omissão de uma variável sobre os resultados do efeito casual do tratamento sobre o tratado na análise de correspondência. Neste sentido, se propõe a verificação da robustez do modelo através da hipótese de viés de seleção.

A probabilidade do egresso ter participado da política de cotas é dado por:

$$Pr(X_i) = Pr(D_i = 1 \mid X_i) = G(\beta X_i + \gamma \mu_i) \quad (9)$$

Assumimos que dois indivíduos  $i$  e  $j$  sejam pareados segundo suas características observáveis, bem como  $G$  tenha uma distribuição logística, a razão de chances *odds ratio* de dois indivíduos pareados receberem o tratamento pode ser escrita como:

$$\frac{Pr(X_i)[1 - Pr(X_j)]}{Pr(X_j)[1 - Pr(X_i)]} = \frac{G(\beta X_i + \gamma \mu_i)}{G(\beta X_j + \gamma \mu_j)} \quad (10)$$

Caso existam fatores omissos os dois indivíduos apesar de idênticos nas características observáveis, terão diferentes probabilidades de serem expostos ao tratamento, logo a razão de probabilidade diferirá de um, implicando na existência de viés de seleção.

De acordo com Rosenbaum (2002), a equação (10) implica nos limites da *odds ratio*, expressa pela desigualdade abaixo:

$$\Gamma^{-1} \leq \frac{Pr(X_i)[1 - Pr(X_j)]}{Pr(X_j)[1 - Pr(X_i)]} \leq \Gamma \quad (11)$$

Sendo  $\Gamma = e^\gamma$ . Para os indivíduos pareados que possuem a mesma probabilidade de participar do tratamento, o *odds ratio* será igual a um, caso contrário, os indivíduos que supostamente se assemelham em características observáveis podem divergir em suas chances de receber o



tratamento. Assim, essencialmente,  $\Gamma$  estima o grau de afastamento que uma estimativa realizada por PSM está livre de fatores omissos. Em outras palavras, os limites de Rosenbaum indicam o quanto alterações dos fatores não-observados podem afetar a robustez dos resultados.

### 3.4 Dados

Para alcançar o objetivo proposto, lançou-se mão de dois conjuntos de microdados. O primeiro diz respeito aos dados administrativos da UFPB, a qual agrega informações acadêmicas e socio-econômicas dos egressos ao longo da graduação. Neste caso, a amostra das informações da UFPB corresponde aos ex-alunos admitidos em 2010 e 2011, tendo em conta que a política de cotas tem início em 2011. Neste sentido, com intuito de minimizar problemas de viés de autoseleção, adotou-se um grupo de egressos que ingressaram na UFPB em 2010 como parte do grupo de controle, em outras palavras, caso houvesse a reserva de vaga já em 2010, estes teriam probabilidade de se inscrever na política.

Já a Relação Anual de Informações Sociais (RAIS) identificada corresponde a segunda fonte de microdados, a mesma compreende as características dos trabalhadores formais, bem como informações sobre sua ocupação e rendimentos. Desse modo, o conjunto de dados finais é resultado da junção dos dados da UFPB e da RAIS identificada de 2018, para tal se utilizou o CPF do egresso, viabilizando a identificação dos ex-alunos na base da RAIS. Ainda se ressalta que o estudo se limitou a 2018 por dois motivos. O primeiro trata do período de conclusão, tendo em vista que para alguns cursos os egressos concluíram apenas em 2017, segundo, a RAIS identificada de 2018 é a mais atual disponível.

Quanto a determinação da incidência de *overeducation*, a literatura estabelece basicamente três métodos fundamentais. O primeiro corresponde à *job analysis*, o mesmo se fundamenta em uma classificação ocupacional elaborada por analistas de trabalho profissionais, que especificam o grau de instrução requerida para cada ocupação. O segundo método é o *self-assessment*, neste caso o próprio trabalhador reporta o nível de escolaridade exigida por sua ocupação; já o terceiro corresponde ao *realized matches*, nessa técnica o nível de instrução requerida é obtida pela média da distribuição da escolaridade dos trabalhadores inseridos em cada ocupação. Desta forma, quem possuir escolaridade superior à média mais o desvio-padrão de sua ocupação é classificado como *overeducated*.

Para o objetivo do estudo, bem como para os dados disponíveis, o presente trabalho empregou o *job analysis*, o mesmo é conceitualmente superior aos demais métodos (HARTOG, 2000). No Brasil, a codificação das ocupações é realizada pela Classificação Brasileira das Ocupações (CBO) de 2002 presente nos microdados da RAIS, o mesmo é descrito por famílias ocupacionais que constitui um conjunto de ocupações semelhantes quanto as atividades realizadas.

A CBO 2002 realiza algumas adaptações da metodologia adotada pela *International Standard Classification of Occupations* de 1988 (ISCO 88), a mesma utiliza a escolaridade como critério para determinar o nível de competência. Desse modo, por meio da CBO, será verificado o grau de instrução requerido para cada ocupação, assim caso o grau de instrução exigido seja menor que o nível superior, o egresso será considerado sobre educado, sendo uma *dummy* igual a um.

A tabela 1 a seguir apresenta a estatística descritiva da amostra de egressos utilizadas no balanceamento por entropia e no PSM. As informações estão subdivididas em uma amostra completa, *overeducation* geral, cotista e não cotista. A princípio em relação às características individuais e socioeconômicas, observa-se que em todos os casos as amostras são compostas em sua maioria por mulheres e egressos não brancos. Quanto ao ensino médio, nota-se que em todas as situações a maioria dos egressos concluiu na rede pública, com destaque para os *overeducation* cotistas, com 93%. Já em relação à escolaridade dos pais, constata-se que parte significativa não possui ensino superior, com ênfase para os sobre-educados cotistas, que descola consideravelmente dos demais grupos. Quanto a renda familiar, observa-se que não há muita

distinção entre as amostras, contudo entre os cotistas o percentual é significativamente maior. No que diz respeito as variáveis acadêmicas verificam-se que não há uma diferença significativa na duração média da graduação, já em relação ao cra, nota-se que os cotistas em *overeducation* tem baixo rendimento acadêmico em relação aos seus pares.

Com respeito as variáveis de mercado de trabalho, pode-se identificar que em média os egressos em situação de *overeducation* auferem as menores remunerações em relação aos demais grupos, por outro lado, parte expressiva dos egressos trabalham no estado da Paraíba, e mais de 50% possui vínculo no setor público e conservam em torno de um vínculo. Quanto a classificação Brasileira de ocupações (CBO) 2002, pode-se verificar que a amostra completa se concentra nos CBO's 2 e 4, já entre os *overeducation* o CBO 4 reuni a maioria.

Tabela 1 – Estatística descritiva das variáveis usadas no pareamento

<i>Varáveis</i>	<i>Completo</i>	<i>Overeducation</i>	<i>Overeducation Cotista</i>	<i>Overeducation Não Cotista</i>
<b><i>Características individuais e Socioeconômicas</i></b>				
Mulher	0.59	0.58	0.57	0.58
Homem	0.41	0.42	0.43	0.42
Branco	0.44	0.42	0.41	0.42
Não Branco	0.56	0.58	0.59	0.58
Idade	29.55	30.11	30.64	30.05
Ensino Médio Público	0.53	0.60	0.93	0.57
Pai Superior	0.19	0.17	0.03	0.18
Mãe Superior	0.24	0.22	0.07	0.23
Natural de Paraíba	0.87	0.89	0.88	0.88
Renda até Dois Salários Mínimos	0.47	0.50	0.74	0.48
cra	8.10	8.04	7.96	8.05
Duração da Graduação	4.50	4.55	4.39	4.56
<b><i>Mercado de trabalho</i></b>				
Salário	2430.01	2193.57	1953.22	2220.06
Local de Trabalho	0.83	0.85	0.89	0.85
Setor Público	0.51	0.55	0.50	0.56
Quantidade de Vínculos	1.12	1.07	1.06	1.07
<b><i>Classificação Brasileira de Ocupações CBO 2002</i></b>				
CBO1	0.15	0.23	0.24	0.24
CBO2	0.37	0.03	0.01	0.04
CBO3	0.09	0.13	0.12	0.14
CBO4	0.27	0.41	0.43	0.42
CBO5	0.05	0.09	0.11	0.09
CBO6	0.00	0.00	0.00	0.00
CBO7	0.01	0.02	0.02	0.02
CBO8	0.00	0.00	0.00	0.00
CBO9	0.00	0.00	0.00	0.00
<b><i>Classificação Nacional de Atividades Econômicas - CNAE 2.0</i></b>				

Agricultura	0.00	0.00	0.00	0.00
Indústria	0.05	0.07	0.08	0.07
Serviços	0.83	0.81	0.77	0.82
Comercio	0.10	0.10	0.13	0.11
Construção Civil	0.01	0.00	0.01	0.00
<b>Grandes áreas de conhecimento</b>				
Ciências Agrária	0.03	0.04	0.06	0.04
Ciências da Saúde	0.19	0.09	0.12	0.08
Ciências Exatas e da Terra	0.04	0.04	0.01	0.04
Ciências Humanas	0.21	0.22	0.18	0.22
Ciências Sociais Aplicadas	0.29	0.39	0.42	0.38
Engenharias	0.06	0.05	0.04	0.05
Linguística, Letras e Artes	0.15	0.15	0.15	0.15

Fonte: Elaboração própria com base nos microdados da STI/UFPB e RAIS 2018. Erro-padrões em parenteses.

Nota: \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

Com relação aos grandes setores da economia, observa-se que parte considerável dos egressos trabalha no setor de serviços, seguido pelo comércio. Quanto às grandes áreas do conhecimento, nota-se que os egressos se concentram nas áreas de ciências sociais aplicadas e Humanas.

## 4 Resultados

### 4.1 Análise do Ajuste do Modelo

A atual seção apresenta e discute os resultados das estimativas de efeitos da política de reserva de vagas sobre a situação dos egressos cotistas da UFPB no mercado de trabalho formal quanto a escolaridade exigida por sua ocupação. Para tanto, o presente artigo utilizou a estratégia proposta por Hainmueller (2012) e Watson e Elliot (2016). Desse modo, em um primeiro momento empreendeu-se o balanceamento por entropia das variáveis observáveis<sup>5</sup>, com o objetivo de tornar os dois grupos mais homogêneos, isso se dá pela obtenção de pesos que minimizam as desigualdades entre os grupos de tratados e controle. Em seguida, implementou-se o pareamento dos grupos de tratados e controle via *Propensity Score Matching*.

Uma vez alcançado os pesos que equilibram os grupos, implementou-se o teste de diferença de médias (teste t) para examinar a condição de balanceamento das variáveis. Os resultados do teste são reportados na tabela 2 e pode-se inferir que após a aplicação da entropia as diferenças estatísticas foram extintas para todas as variáveis consideradas, em outras palavras, as diferenças de médias são estatisticamente iguais a zero entre os dois grupos em análise, indicando o bom ajuste das covariadas pela entropia.

Tabela 2 – Teste de diferença das covariadas antes e após o Balanceamento por Entropia

Varáveis	Antes do balanceamento			Depois do balanceamento		
	Tratamento	Controle	Diferença	Tratamento	Controle	Diferença

<sup>5</sup> No balanceamento foi retirado a variável Agropecuária devido a falta de observações dessas variáveis.

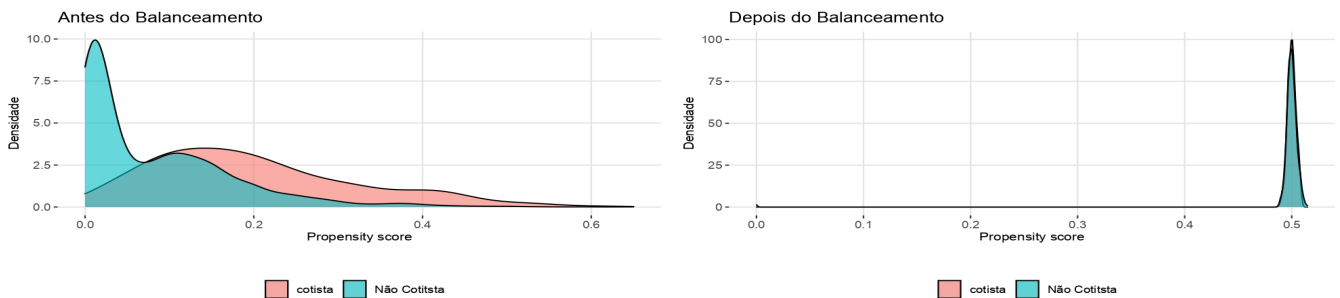
Raça	0.421	0.445	-0.024	0.421	0.422	0.001
Idade	30.222	29.485	0.737*	30.222	30.198	0.024
Escola Pública	0.926	0.499	0.427***	0.926	0.925	0.001
Sexo	0.552	0.595	-0.043	0.552	0.551	0.001
Mãe Ensino Superior	0.069	0.261	-0.192***	0.069	0.070	-0.001
Pai Ensino Superior	0.034	0.211	-0.177***	0.034	0.035	-0.001
Natural da Paraíba	0.860	0.875	-0.015	0.860	0.860	0.000
ra	8.022	8.116	-0.094*	8.022	8.016	0.006
Duração do curso	4.417	4.513	-0.096	4.417	4.413	0.004
Renda até dois salários	0.704	0.450	0.254***	0.704	0.703	0.001
Ocupado	0.321	0.230	0.091***	0.321	0.322	-0.001
Setor público	0.478	0.517	-0.039	0.478	0.478	0.000
Quantidade de vínculos	1.117	1.120	-0.003	1.117	1.120	-0.003
Local Paraíba	0.873	0.833	0.04	0.873	0.873	0.000
Salário	2219.6	2451.2	-231.6*	2219.6	2218.1	1.5
<b><i>Classificação Nacional de Atividades Econômicas - CNAE 2.0</i></b>						
Indústria	0.060	0.055	0.005	0.060	0.060	0.000
Comércio	0.139	0.096	0.043*	0.139	0.139	0.000
Serviços	0.786	0.835	-0.049*	0.786	0.786	0.000
Construção	0.013	0.008	-0.005	0.013	0.013	0.000
<b><i>Grandes Áreas de Conhecimento</i></b>						
Ciências Agrária	0.043	0.034	0.009	0.043	0.043	0.000
Ciências da Saúde	0.239	0.192	0.047*	0.239	0.240	-0.001
Ciências Exatas e da Terra	0.021	0.046	-0.025*	.021	0.021	0.000
Ciências Humanas	0.165	0.218	-0.053*	0.165	0.165	0.000
Ciências Sociais Aplicadas	0.317	0.292	0.025	0.317	0.318	-0.000
Engenharias	0.056	0.063	-0.007	0.056	0.056	0.000
Linguística, Letras e Artes	0.156	0.152	0.004	0.156	0.153	0.003

Fonte: Elaboração própria com base nos microdados da STI/UFPB e RAIS 2018.

Nota \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

Por sua vez, a figura 1 apresenta o comportamento da distribuição dos escores de propensão calculados via regressão *logit* antes e após a realização da entropia. A análise visual sugere que a sobreposição dos grupos após a entropia é superior, se comparado aos escores de propensão estimados de forma convencional, evidenciando que os mesmos se tornaram equilibrados após a aplicação dos pesos da entropia ao conjunto de dados do grupo de controle, tornando-os semelhantes nos observáveis.

Figura 1 – Distribuição dos escores de propensão por grupo antes e depois do balanceamento por entropia



Fonte: Elaboração própria com base nos microdados da STI/UFPB e RAIS 2018.

## 4.2 Efeito Médio da Política Afirmativa

Após a entropia, os grupos foram pareados pelos algoritmos de vizinhança mais próxima (*Nearest Neighbor*) com reposição de um e dez observações do grupo de controle e também sem reposição, como também pelas técnicas *radius matching* e *Kernel matching*. Em seguida foi calculado o efeito médio do tratamento sobre os tratados (ATT) da política de reservas de vagas sobre a chance do egresso cotista encontrar-se em condições de sobre-educado. Os resultados alcançados podem ser vistos na tabela 3, e se referem ao conjunto de todos os cursos da UFPB, tal como para os grupos de egressos com os menores e maiores salários.

Inicialmente, na tabela 3, é possível constatar que para o conjunto da UFPB, em todos os algoritmos considerados no trabalho, não foram encontrados efeitos estatisticamente significativos da política de cotas sobre a probabilidade dos egressos cotistas estarem em *overeducation*. Deste modo, segundo os achados do presente trabalho, a princípio, para o agregado de cursos da UFPB, não há indícios de impactos da reserva de vagas sobre a possibilidade do egresso beneficiário está em ocupações que não demandem qualificação de nível superior.

Por sua vez, a literatura empírica revela a associação entre o fenômeno do *overeducation* e a variável salário, assim, de modo a verificar esta relação na presença de política de ações afirmativas, similarmente, estendeu-se a investigação a grupos de egressos que compõem as faixas salariais mais baixas e elevadas da amostra, logo estimou-se, respectivamente, o efeito médio para os quantis 0.25 e 0.75 da distribuição de salários. Os resultados para o quantil 0.25, são similares aos achados para o conjunto da UFPB, conseqüentemente permite-se concluir que não existem indícios estatísticos de efeitos da política em análise sobre o *overeducation* entre os egressos cotistas com remunerações mais baixas.



Tabela 3 – Efeito médio sobre os tratados para todos egresso e para os quantis salariais

<i>Modelo</i>	<i>Efeito</i>	<i>Quantil salarial</i>	
	<i>Médio</i>	<i>0.25</i>	<i>0.75</i>
Nearest (CR, caliper) K=1	0.102 (0.208)	0.252 (0.233)	0.583** (0.263)
Nearest (CR, caliper) K=10	0.222 (0.157)	0.313 (0.176)	0.191 (0.201)
Nearest (SR, caliper)	0.081 (0.202)	0.276 (0.225)	0.510** (0.240)
Radius	0.238 (0.151)	0.275 (0.169)	0.423** (0.186)
Kernel	0.247 (0.151)	0.266 (0.169)	0.423** (0.186)

**Fonte:** Elaboração própria com base nos microdados da STI/UFPB e RAIS 2018. Erro-padrões em parenteses.

**Nota\***  $p < 0.1$ ; **\*\***  $p < 0.05$ ; **\*\*\***  $p < 0.01$ . K é o número de vizinhos considerados na estimação. Foi utilizado um *caliper* de 0.001. No *Kernel* o valor de *bwidth* é de 0.06.

Por outro lado, entre os ex-alunos cotistas que se encontram no topo da distribuição de salários, as estimações são positivas e estaticamente significativas, com exceção da técnica *Nearest* com reposição e dez vizinhos. Assim, diante o exposto, pode-se inferir que os integrantes deste grupo investigado tendem a deter maiores chances de serem subutilizados no mercado de trabalho, porém os mesmos não são potencialmente penalizados em termos salariais. Este último resultado vai de encontro ao que a literatura econômica até então aponta, segunda a mesma a incidência de *overeducation* é menor entre indivíduos com salários mais elevados (ALPIN; SHACKLETON; WALSH, 1998), (BATTU; SLOANE, 2004), (TSAI, 2010), (CLARK; JOUBERT; MAUREL, 2017) e (ANNEGUES; SOUZA, 2020). Todavia, cabe ressaltar que as evidências da literatura correspondem a análises com ausência de políticas afirmativas.

Da mesma forma, examinou-se determinadas áreas do conhecimento, as mesmas foram definidas com base na amostra disponível, sendo consideradas as áreas com mais de trinta observações tanto entre os tratados quanto o grupo de controle, viabilizando as estimações. Desta forma, a tabela 4 traz o efeito médio para as áreas de Ciências da Saúde, Ciências Humanas, Ciências Sociais Aplicadas e Letras, Linguística e Arte.

Tabela 4 – Efeito médio por área do conhecimento

<i>Modelo</i>	<i>CS</i>	<i>CH</i>	<i>CCSA</i>	<i>LLA</i>
Nearest (CR, caliper) K=1	-0.741 (0.467)	1.233** (0.553)	0.769 (0.598)	-0.431 (0.629)
Nearest (CR, caliper) K=10	0.087 (0.362)	0.850* (0.448)	0.869* (0.487)	-0.248 (0.422)
Nearest (SR, caliper)	-0.377 (0.391)	1.310** (0.541)	0.880 (0.569)	-0.431 (0.543)
Radius	0.321 (0.299)	0.844* (0.447)	0.831* (0.479)	0.179 (0.385)
Kernel	0.321 (0.299)	0.798* (0.419)	0.829* (0.476)	0.179 (0.385)

**Fonte:** Elaboração própria com base nos microdados da STI/UFPB e RAIS 2018. Erro-padrões em parenteses.

**Nota\***  $p < 0.1$ ; **\*\***  $p < 0.05$ ; **\*\*\***  $p < 0.01$ . K é o número de vizinhos considerados na estimação. Foi utilizado um *caliper* de 0.001. No *Kernel* o valor de *width* é de 0.06. Áreas do conhecimento: CS: Ciências da Saúde; CH: Ciências Humanas; CCSA: Ciências Sociais Aplicadas; LLA: Linguística, Letras e Arte.

As informações da tabela 4 permitem identificar coeficientes positivos e estatisticamente significativos apenas para Ciências Humanas e Ciências Sociais Aplicadas, sendo que para esta última área os algoritmos *Nearest* com reposição e um vizinhos e *Nearest* sem reposição não apresentaram significância. Desta forma, os resultados alcançados sugerem que os formados cotistas portadores de diplomas das referidas áreas de estudo tendem a serem mais suscetíveis a ocuparem postos de trabalho incompatíveis com seu nível de instrução, seguindo evidências anteriormente encontradas para ambientes sem ações afirmativas como em [Alpin, Shackleton e Walsh \(1998\)](#) e [Barone e Ortiz \(2011\)](#).

O resultado acima pode ser justificado por dois principais fatores, o primeiro se refere a especificação das ocupações das áreas em análise, as quais não são bem definidas, ao contrário de Ciências da Saúde que exige certificação específica para a função. Já o segundo diz respeito a restrição do mercado de trabalho de Ciências Humanas e Ciências Sociais Aplicadas ([ANNEGUES et al., 2018](#)), ([ANNEGUES; SOUZA, 2020](#)). Desse modo, no caso das duas áreas investigadas, garantir o acesso de grupos minoritários ao superior não é suficiente em termos de mobilidade ocupacional, tendo em vista que aspectos do mercado de trabalho e da área de formação podem penalizar os beneficiários da política.

### 4.3 Análise de sensibilidade

As estimações apresentadas na subseção anterior assumiram as hipóteses de sobreposição e independência condicional, esta última admite que os potenciais resultados são independentes da designação do tratamento. Contudo, a presença de variáveis não observáveis na especificação do modelo pode produzir resultados viesados. Isto posto, verificou-se a sensibilidade das estimações do efeito médio do tratamento quanto ao viés de seleção por meio dos limites de *Rosenbaum* para os quantis 0.75, Ciências Humanas e Ciências Sociais Aplicadas, tendo em vista que estes foram os únicos a apresentarem significância estatística.

A tabela 5 reúne as estatísticas dos limites ( $\Gamma$ ) e a probabilidade ( $p^+$ ) de não rejeição da hipótese nula de sobrestimação do efeito para o conjunto de egressos cotistas presentes no quantil 0.75 da distribuição de salários. Os limites apresentados correspondem a pelo menos 10% de

significância estatísticas. DiPrete e Gangl (2004), sugere que os limites abaixo de 1.10 indicam forte influência de fatores não observáveis sobre o efeito médio do tratamento. Desse modo, os resultados presentes na tabela 5 se mostram robustos a uma possível presença de variáveis não observadas, visto que o menor valor é de 1.25 obtido pelos métodos *Nearest* sem reposição e *Kernel*.

Tabela 5 – Análise de sensibilidade do quantil 0.75

<i>Modelo</i>	<i>Limites</i> ( $\Gamma$ )	$p^+$
Nearest (CR, caliper) K=1	1.40	0.1230
Nearest (SR, caliper)	1.25	0.1012
Radius	1.30	0.1427
Kernel	1.25	0.1092

**Fonte:** Elaboração própria com base nos microdados da STI/UFPB e RAIS 2018.

**Nota:**  $\Gamma$  é o primeiro valor com significância estatística do intervalo para a suposição de sobrestimação;  $p^+$  refere-se a probabilidade de não rejeição da hipótese de sobrestimação no efeito da exposição. K é o número de vizinhos considerados na estimação. Foi utilizado um *caliper* de 0.001. No *Kernel* o valor de *bwidth* é de 0.06.

Já os limites de *Rosenbaum* por área de conhecimento são expostos na tabela 6. Para Ciências Humanas os resultados revelam que independentemente do algoritmo aplicado as estimacões são robustas à presença de viés produzida por variáveis omitidas. Entre as técnicas usadas, destaca-se o *Nearest* com reposição e *Nearest* com reposição de um vizinho, onde um viés de 1.75 leva a uma diferença de *odds ratio* do egresso cotista de ciências humanas ser exposto à intervenção e assim questionar se há efeitos positivos da reserva de vagas sobre a chance do egresso ser *overeducation*, por outro lado, o método *Kernel* demonstra ser mais sensível a fatores omitidos, onde o mesmo responde por uma diferença de 10% na razão de chances de exposição a política de cotas.

Tabela 6 – Análise de sensibilidade por grande área

<i>Modelo</i>	<i>Limites</i> ( $\Gamma$ )	<i>CH</i>	<i>CCSA</i>
Nearest (CR, caliper) K=1	$\Gamma$	1.75	
	$p^+$	0.1061	
Nearest (CR, caliper) K=10	$\Gamma$	1.45	4.05
	$p^+$	0.1031	0.1043
Nearest (SR, caliper)	$\Gamma$	1.75	
	$p^+$	0.1061	
Radius	$\Gamma$	1.45	4.0
	$p^+$	0.1032	0.1015
Kernel	$\Gamma$	1.10	4.0
	$p^+$	0.1332	0.1015

**Fonte:** Elaboração própria com base nos microdados da STI/UFPB e RAIS 2018.

**Nota:**  $\Gamma$  é o primeiro valor com significância estatística do intervalo para a suposição de sobrestimação;  $p^+$  refere-se a probabilidade de não rejeição da hipótese de sobrestimação no efeito da exposição. K é o número de vizinhos considerados na estimação. Foi utilizado um *caliper* de 0.001. No *Kernel* o valor de *bwidth* é de 0.06. CH: Ciências Humanas; CCSA: Ciências Sociais Aplicadas

Quanto as estimações para Ciências Sociais Aplicadas, os resultados indicam que os mesmos são robustos à presença de viés de seleção nos três algoritmos que apresentam significância estatística do ATT. Os achados apontam que os três modelos apresentam a mesma sensibilidade em relação às variáveis não observadas, sendo necessário um fator de 4.0 para invalidar os resultados das estimações dos ATT's, ou seja, os fatores omitidos deveriam aumentar as chances dos egressos a se beneficiarem da política de cotas em 4.0 para torna insignificante o efeito da política sobre a possibilidade de ser *overeducation*.

## 5 Considerações Finais

Em décadas recentes, inúmeras políticas educacionais viabilizaram a ampliação de vagas no ensino superior brasileiro. Todavia, o número de ocupações no mercado de trabalho não acompanhou essa expansão, produzindo o fenômeno do *overeducation*. Em paralelo a esta ampliação houve a consolidação das políticas afirmativas, em especial, a reserva de vagas, com objetivo de ampliar a participação de grupos minoritários na universidade. Desse modo, os beneficiários da referida política também estão suscetíveis a sobre-educação.

Posto isto, procurando contribuir com a literatura, o presente trabalho teve por objetivo investigar o efeito da política de reservas de vagas sobre a probabilidade do egresso cotista da UFPB estar em *overeducation*. Para tal, se lançou mão dos dados administrativos da UFPB de 2010 e 2011, dado o início da ação afirmativa em 2010 na instituição em análise, e a RAIS identificada de 2018 que permitiu conhecer a situação ocupacional dos ex-alunos da instituição.

Para alcançar o objetivo proposto, adotou-se como estratégia metodológica a associação do balanceamento por entropia com o *Propensity Score Matching*, proposto por Hainmueller (2012) e Watson e Elliot (2016), a referida associação permite isolar os efeitos da política de cotas sobre a chance do egresso cotista está sendo subutilizado no mercado de trabalho.

A princípio, os resultados para o conjunto dos ex-alunos cotistas da instituição em análise, sugerem que a política de cotas adotada em 2011 não apresentam potenciais efeitos sobre a probabilidade de seus beneficiados serem sobre-educados. Contudo, quando analisado pela distribuição de salários da amostra, os achados revelam efeitos positivos da política de cotas sobre o *overeducation* para os egressos presentes no topo da distribuição (quantil 0.75), ou seja, os formados cotistas com salários mais elevados tendem apresentar maiores chances de ocupar postos de trabalho que não necessitam de formação universitária.

Quanto as grandes áreas de conhecimento, calcularam-se os efeitos apenas para quatro áreas, dado a amostra disponível e a viabilidade das estimações. Dos campos considerados, apenas Ciências Humanas e Ciências Sociais Aplicadas apresentaram indícios de impactos positivos da política em análise. Desse modo, os egressos cotistas das referidas áreas tendem a ter maior probabilidade de estar sendo subutilizados no mercado de trabalho.

Desta forma, os resultados sugerem que apenas o ingresso no ensino superior não é suficiente para viabilizar a mobilidade ocupacional. Em vista disso, mostra-se necessário que a reserva de vagas no ensino superior deva ser acompanhada de políticas complementares ainda na educação básica e graduação na tentativa de minimizar as desigualdades.

Contudo, é relevante destacar que os resultados e considerações apresentados no presente artigo são evidências a curto prazo, tendo em vista o período de início da política em análise e o tempo de conclusão de alguns cursos, em especial nas áreas de Ciências da Saúde e Engenharias, limitando o corte temporal da pesquisa. Neste sentido, se recomenda que os próximos estudos abordem o impacto da política de cotas sobre a probabilidade dos egressos estar em *overeducation* a médio e longo prazo, como também a análise de sobrevivência dos mesmos na situação de subutilização no mercado de trabalho.

## Referências

- ALPIN, C.; SHACKLETON, J.; WALSH, S. Over-and undereducation in the uk graduate labour market. *Studies in Higher Education*, Taylor & Francis Group, v. 23, n. 1, p. 17–34, 1998.
- ANNEGUES, A. C. et al. Overeducation e área de formação: Evidências para os egressos da ufpb. *XXIII Encontro Regional de Economia (ANPEC Regional)*, Anais, Fortaleza, CE, 2018. Disponível em: <[https://www.anpec.org.br/nordeste/2018/submissao/arquivos/\\_identificados/057-30545e4aac19d96be6b75873119c721d.pdf](https://www.anpec.org.br/nordeste/2018/submissao/arquivos/_identificados/057-30545e4aac19d96be6b75873119c721d.pdf)>.
- ANNEGUES, A. C.; SOUZA, W. P. S. d. F. Retorno salarial do overeducation: Viés de seleção ou penalização ao excesso de escolaridade? *Revista Brasileira de Economia*, SciELO Brasil, v. 74, n. 2, p. 119–138, 2020.
- BARONE, C.; ORTIZ, L. Overeducation among european university graduates: a comparative analysis of its incidence and the importance of higher education differentiation. *Higher Education*, Springer, v. 61, n. 3, p. 325–337, 2011.
- BATTU, H.; SLOANE, P. J. To what extent are ethnic minorities in britain over-educated? *International Journal of Manpower*, MCB UP Ltd, 2002.
- BATTU, H.; SLOANE, P. J. Over-education and ethnic minorities in britain. *The Manchester School*, Wiley Online Library, v. 72, n. 4, p. 535–559, 2004.
- BLACKABY, D. et al. Unemployment among britain’s ethnic minorities. *The Manchester School*, Wiley Online Library, v. 67, n. 1, p. 1–20, 1999.
- CAMERON, A. C.; TRIVEDI, P. K. *Microeconometrics: methods and applications*. [S.l.]: Cambridge university press, 2005.
- CAVALCANTI, G. d. S. et al. Overeducation no estado de sao paulo: Uma análise com dados em painel. *48º Encontro Nacional de Economia (ANPEC)*, Anais, Brasilia, BR, 2020. Disponível em: <[https://www.anpec.org.br/encontro/2020/submissao/files\\_I/i13-df5ff684620655051f7883c214dbab12.pdf](https://www.anpec.org.br/encontro/2020/submissao/files_I/i13-df5ff684620655051f7883c214dbab12.pdf)>.
- CLARK, B.; JOUBERT, C.; MAUREL, A. The career prospects of overeducated americans. *IZA Journal of Labor Economics*, Springer, v. 6, n. 1, p. 3, 2017.
- DIPRETE, T. A.; GANGL, M. Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological methodology*, SAGE Publications Sage CA: Los Angeles, CA, v. 34, n. 1, p. 271–310, 2004.
- DUNCAN, G. J.; HOFFMAN, S. D. The incidence and wage effects of overeducation. *Economics of education review*, Elsevier, v. 1, n. 1, p. 75–86, 1981.
- ESTEVAM, C. et al. Programa de tutoria por pares no ensino superior: Estudo de caso artigo. *Revista Brasileira de Orientação Profissional*, v. 19, n. 2, p. 185–195, 2018.
- ESTEVAN, F.; GALL, T.; MORIN, L.-P. Redistribution without distortion: Evidence from an affirmative action programme at a large brazilian university. *The Economic Journal*, Oxford University Press, v. 129, n. 619, p. 1182–1220, 2018.



- GROOT, W. The incidence of, and returns to overeducation in the uk. *Applied Economics*, Taylor & Francis, v. 28, n. 10, p. 1345–1350, 1996.
- HAINMUELLER, J. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, JSTOR, p. 25–46, 2012.
- HARTOG, J. Over-education and earnings: where are we, where should we go? *Economics of Education Review*, Elsevier, v. 19, n. 2, p. 131–147, 2000.
- JUNIOR, A. A. F. M.; SOUZA, A. d. M. e.; WALTENBERG, F. D. Affirmative action and access to higher education in brazil: The significance of race and other social factors. *Journal of Latin American Studies*, Cambridge University Press, v. 48, n. 2, p. 301–334, 2016.
- LEUVEN, E.; OOSTERBEEK, H. Overeducation and mismatch in the labor market. *Handbook of the Economics of Education*, Elsevier, v. 4, p. 283–326, 2011.
- LINDLEY, J. The over-education of uk immigrants and minority ethnic groups: Evidence from the labour force survey. *Economics of Education Review*, Elsevier, v. 28, n. 1, p. 80–89, 2009.
- LOPES, A. D. Affirmative action in brazil: how students' field of study choice reproduces social inequalities. *Studies in Higher Education*, Routledge, v. 42, n. 12, p. 2343–2359, 2017.
- MCGUINNESS, S. Overeducation in the labour market. *Journal of Economic Surveys*, Wiley Online Library, v. 20, n. 3, p. 387–418, 2006.
- PEIXOTO, A. d. L. A. et al. Cotas e desempenho acadêmico na ufba: um estudo a partir dos coeficientes de rendimento. *Avaliação: Revista da Avaliação da Educação Superior*, SciELO Brasil, v. 21, n. 2, 2016.
- PEREIRA, J. I. R.; BITTENCOURT, M.; JUNIOR, W. S. Análise do impacto da implantação das cotas na nota enade 2008. *41 ° Encontro Nacional de Economia (ANPEC)*, Anais, Foz do Iguaçu, 2013. Disponível em: <[https://www.anpec.org.br/encontro/2013/files\\_I/i6-1b49f644b3a8e2bbd6923656a5c072fc.pdf](https://www.anpec.org.br/encontro/2013/files_I/i6-1b49f644b3a8e2bbd6923656a5c072fc.pdf)>.
- RAFFERTY, A. Ethnic penalties in graduate level over-education, unemployment and wages: evidence from britain. *Work, employment and society*, SAGE Publications Sage UK: London, England, v. 26, n. 6, p. 987–1006, 2012.
- ROSENBAUM, P. R. Attributing effects to treatment in matched observational studies. *Journal of the American Statistical Association*, Taylor & Francis, v. 97, n. 457, p. 183–192, 2002.
- ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, Oxford University Press, v. 70, n. 1, p. 41–55, 1983.
- SATTINGER, M. Assignment models of the distribution of earnings. *Journal of Economic Literature*, JSTOR, v. 31, n. 2, p. 831–880, 1993.
- SCHWARTZMAN, S. Equity, quality and relevance in higher education in brazil. *Anais da Academia Brasileira de Ciências*, SciELO Brasil, v. 76, n. 1, p. 173–188, 2004.
- SICHERMAN, N. "overeducation" in the labor market. *Journal of Labor Economics*, University of Chicago Press, v. 9, n. 2, p. 101–122, 1991.
- SILVA, A. F. d. et al. Efeitos de políticas afirmativas sobre esforço e abandono: Evidências a partir da universidade federal da paraíba. *XXIV Encontro Regional de Economia*, Anais, 2019. Disponível em: <[https://www.anpec.org.br/nordeste/2019/submissao/arquivos\\_identificados/050-ce1afaf5480238389cb368b73ff0451b.pdf](https://www.anpec.org.br/nordeste/2019/submissao/arquivos_identificados/050-ce1afaf5480238389cb368b73ff0451b.pdf)>.

SLOANE, P. J.; BATTU, H.; SEAMAN, P. T. Overeducation, undereducation and the british labour market. *Applied Economics*, Taylor & Francis, v. 31, n. 11, p. 1437–1453, 1999.

THUROW, L. C.; LUCAS, R. E. *The American distribution of income: a structural problem*. [S.l.]: US Government Printing Office, 1972. v. 7.

TSAI, Y. Returns to overeducation: A longitudinal analysis of the us labor market. *Economics of Education Review*, Elsevier, v. 29, n. 4, p. 606–617, 2010.

TSANG, M. C.; LEVIN, H. M. The economics of overeducation. *Economics of Education Review*, Elsevier, v. 4, n. 2, p. 93–104, 1985.

WATSON, S. K.; ELLIOT, M. Entropy balancing: a maximum-entropy reweighting scheme to adjust for coverage error. *Quality & Quantity*, Springer, v. 50, n. 4, p. 1781–1797, 2016.

ZHAO, Q.; PERCIVAL, D. Entropy balancing is doubly robust. *Journal of Causal Inference*, De Gruyter, v. 5, n. 1, 2016.