

**PENALIZAÇÃO SALARIAL DO *OVEREDUCATION*: EVIDÊNCIAS PARA OS
EGRESSOS DA FURG PELO *PROPENSITY SCORE MATCHING* (PSM)¹**

Larissa Castagna Gonçalves²
Vivian dos Santos Queiroz Orellana³
Gibran da Silva Teixeira⁴
Rodrigo Rocha Gonçalves⁵
Jorge Alberto Orellana Aragón⁶

RESUMO: Este estudo tem como objetivo investigar a penalização salarial associada à condição de *overeducation* entre egressos da Universidade Federal do Rio Grande (FURG). Para tanto foi construída uma base de dados com egressos da FURG entre 2017 e 2022, cruzada com as informações da Relação Anual de Informações Sociais (RAIS). O método aplicado foi o Propensity Score Matching (PSM). Os resultados indicam que os egressos nessa situação sofrem, em média, uma penalização salarial de aproximadamente 42%.

Palavras-chaves: *Overeducation*. PSM. Salário. FURG.

ABSTRACT: This study aims to investigate the wage penalty associated with the condition of *overeducation* among graduates from the Federal University of Rio Grande (FURG). To achieve this, a dataset was constructed using information on FURG graduates between 2017 and 2022, cross-referenced with data from the Annual Social Information Report (RAIS). The methodology applied was Propensity Score Matching (PSM). The results indicate that graduates in a situation of *overeducation* experience, on average, a wage penalty of approximately 42%.

Keywords: *Overeducation*. PSM. Wage. FURG.

Classificação JEL: J23, J31

¹ Projeto financiado pelo Edital Fapergs 09/2023 Programa Pesquisador Gaúcho - PqG

² Mestranda do Programa de Pós-Graduação em Economia Aplicada (PPGE) da Universidade Federal do Rio Grande (FURG). Email: castagna.larissa@gmail.com

³ Professora do Programa de Pós-Graduação em Economia Aplicada (PPGE) da Universidade Federal do Rio Grande (FURG). Email: vivianqueiroz@furg.br

⁴ Professor do Programa de Pós-Graduação em Economia Aplicada (PPGE) da Universidade Federal do Rio Grande (FURG). Email: tgibran@furg.br

⁵ Professor do Programa de Pós-Graduação em Economia Aplicada (PPGE) da Universidade Federal do Rio Grande (FURG). Email: rrochagoncalves@furg.br

⁶ Professor do Programa de Pós-Graduação em Economia Aplicada (PPGE) da Universidade Federal do Rio Grande (FURG). Email: jorgeorellana@furg.br

1. Introdução

O número de vagas no ensino superior aumentou expressivamente no país nas últimas décadas, pois, segundo Pereira et al. (2016), o número de matrículas em cursos de nível universitário passou de aproximadamente 480 mil em 2002 para mais de um milhão em 2012. Essa expansão foi impulsionada por políticas e programas educacionais bem-sucedidos que proporcionaram maior inclusão e formação de trabalhadores qualificados, tais como o Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI), o Fundo de Financiamento Estudantil (FIES), o Programa Universidade para Todos (PROUNI) e a oferta de cursos de graduação pelos Institutos Tecnológicos (IFs) (Ferreira et al., 2021).

Apesar dos avanços na educação superior, é importante destacar que algumas evidências apontam que há um descompasso entre a expansão do número de graduados e a oferta de vagas no mercado de trabalho. Segundo Simões (1985), mais de 80% dos formandos de cursos de ensino superior ingressavam no mercado de trabalho até um ano após a formatura no início da década de 1970. No entanto, entre 2014 e 2018, apenas 25% dos egressos se inseriram no mercado de trabalho após a formatura (Lima, 2021).

Diante da dificuldade de inserção no mercado de trabalho, muitos graduados acabam ingressando em ocupações diferentes das suas respectivas áreas de formação (Almeida et al., 2020). Tal fenômeno é conhecido na literatura como *overeducation* (ou sobre-educação) é o termo usado para explicar a inserção de trabalhadores que possuem um grau de formação maior do que o exigido pela ocupação em que estão inseridos (Annegues; Souza, 2019). Diaz e Machado (2008) fizeram uma análise para o Brasil e regiões e observaram que aproximadamente 17,3% da população brasileira se encontra em situação de *overeducation*.

A inadequação entre escolaridade e ocupação pode trazer consequências negativas tanto para os indivíduos quanto para o mercado de trabalho. Entre os principais impactos estão o desperdício de capital humano, a redução da satisfação e da motivação profissional, a oferta de salários inferiores ao esperado para o nível de qualificação, menor produtividade e maior rotatividade no emprego (Almeida, 2023). Algumas áreas de estudo, porém, podem ser mais restritas, o que dificulta a inserção do egresso no seu campo de formação (Reimer et al., 2008). No entanto, é possível que o *overeducation* seja temporário, ou seja, alguns trabalhadores podem ocupar funções que requerem menor qualificação por necessidade ou expectativa de se inserirem em ocupações equivalentes à sua formação no futuro, aproveitando para ganhar experiência e ampliar a rede de contatos (Annengues et al., 2018).

Santos (2002) ressalta que a sobre-educação é particularmente preocupante diante da escassez de profissionais com formação de qualidade no contexto brasileiro. Entre os fatores apontados para esse cenário estão, de um lado, as estratégias empresariais de contratação de trabalhadores com maior escolaridade que são considerados mais adaptáveis e facilmente treináveis e, de outro, a baixa qualidade da educação oferecida no país e a falta de alinhamento entre o sistema educacional e as demandas do mercado de trabalho. Diante dessas condições, torna-se mais viável para as empresas contratarem indivíduos superqualificados mesmo que seja para ocupações que não exijam tal nível de instrução (Santos, 2002).

Alguns estudos se destacam pela investigação da *overeducation* na literatura nacional, como o feito por Annegues et al (2018) que investigaram a sobre-educação entre os recém-formados na Universidade Federal da Paraíba (UFPB). O estudo desenvolvido por Cavalcanti et al (2020) contribuíram sobre o tema ao analisar os

efeitos do *overeducation* sobre os rendimentos dos trabalhadores graduados do Estado de São Paulo. Almeida (2023), por sua vez, analisou a sobre-educação entre os jovens graduados do estado de Minas Gerais.

Diante do exposto, observa-se uma lacuna na literatura no que se refere à análise do fenômeno do *overeducation* entre egressos recém-formados, especialmente no contexto brasileiro. Nesse sentido, o presente estudo tem como objetivo principal investigar a penalização salarial associada à condição de *overeducation* entre os egressos recém-formados da Universidade Federal do Rio Grande (FURG), bem como estimar o tempo médio de permanência nessa condição após a inserção no mercado de trabalho. Especificamente, será feita uma análise por áreas do conhecimento. Para atingir os propósitos desse estudo, adota-se o método de *Propensity Score Matching* (PSM), com o intuito de controlar possíveis vieses de seleção entre os grupos comparados, e análise de sobrevivência, visando estimar a duração do *mismatch* educacional ao longo do tempo. A base de dados utilizada integra informações da Relação Anual de Informações Sociais (RAIS) identificada vinculadas aos registros acadêmicos de ex-alunos da FURG formados entre 2017 e 2022 que foram extraídos do Sistema Acadêmico Institucional.

A relevância deste estudo justifica-se pela escassez de análises que articulem a penalização salarial e a duração da condição de *overeducation* entre egressos do ensino superior brasileiro, especialmente em instituições públicas fora dos grandes centros. A literatura nacional tem avançado na identificação da incidência do fenômeno, mas ainda carece de investigações que combinem abordagens causais e temporais para mensurar seus efeitos econômicos e estruturais. Ao utilizar dados administrativos da FURG integrados com a RAIS e empregar métodos citados acima, esta pesquisa contribui para preencher essa lacuna ao oferecer evidências inéditas sobre os impactos do *overeducation* no início da trajetória profissional dos egressos. Os resultados fornecerão subsídios empíricos para o aprimoramento das políticas educacionais e para a avaliação da efetividade da formação superior frente às exigências do mercado de trabalho, especialmente em um contexto de expansão recente do ensino superior no Brasil.

O presente estudo propõe uma abordagem distinta do que foi feito por Annegues et al. (2018) por investigar o efeito médio da situação de *overeducation* sobre os salários, controlando o viés de seleção entre os grupos de egressos sobre-educados e adequadamente alocados. Ao estimar também a duração da permanência nessa condição, este trabalho oferece uma perspectiva dinâmica que ainda é pouco explorada na literatura nacional. Além de fazer uma atualização temporal e metodológica em relação a pesquisas anteriores, a presente pesquisa contribui de forma original ao aprofundar a compreensão sobre os custos individuais do *mismatch* educacional ao fornecer subsídios empíricos para o debate sobre a efetividade da formação superior e sua adequação ao mercado de trabalho.

Este estudo se encontra dividido em cinco seções, incluindo esta introdução. A segunda seção apresenta o referencial teórico sobre *overeducation* e faz uma breve revisão dos estudos empíricos. A seção três expõe a metodologia e a base de dados utilizada e a seção quarto exhibe e discute os resultados. Por fim, na última seção tem-se as considerações finais.

2. Referencial Teórico

O fenômeno da sobre-educação ou *overeducation* suscita importantes discussões sobre a articulação entre o sistema educacional e o mercado de trabalho.

De acordo com Diaz e Machado (2008), a presença de trabalhadores cuja escolaridade excede os requisitos das ocupações que exercem desafia as teorias tradicionais de alocação de trabalho. A literatura aponta três abordagens principais para explicar esse desajuste: a Teoria do Capital Humano, o Modelo de Competição por Emprego (*Job Competition*) e os Modelos de Atribuição (*Assignment Models*).

A Teoria do Capital Humano, conforme apresentada por Silva (2009), estabelece que os salários refletem os investimentos realizados em educação e treinamento, de modo que anos adicionais de escolaridade resultam, em média, em maiores rendimentos. Nessa perspectiva, a sobre-educação é interpretada como uma condição temporária: os trabalhadores superqualificados estariam apenas aguardando melhores oportunidades ou um processo de ajuste produtivo por parte das empresas, que permitiria o aproveitamento pleno do capital humano disponível (Diaz e Machado, 2008). Assim, o descompasso entre qualificação e ocupação seria resultado de desequilíbrios conjunturais entre oferta e demanda de trabalho, e não uma falha estrutural da teoria.

Por outro lado, os Modelos de Competição por Emprego consideram que, em contextos de alta escolarização da força de trabalho, os indivíduos acumulam mais anos de estudo para manter sua posição competitiva no mercado. Nesse modelo, o nível educacional funciona como um critério de seleção, mesmo quando a ocupação não exige tal qualificação formal (Anegues e Souza, 2019). Logo, o fenômeno da sobre-educação se intensifica quando há um número considerável de pessoas escolarizadas na população, forçando os indivíduos a investirem mais em instrução para manter sua posição no mercado de trabalho (Diaz; Machado, 2008).

Já os *Assignment Models* propõem que os salários resultam da alocação dos trabalhadores a diferentes tipos de ocupações com base em atributos individuais como escolaridade, habilidades e produtividade, e nas características dos postos disponíveis. De acordo com Sattinger (1993) e McGuinness (2006), essa abordagem reconhece que a sobre-educação não decorre apenas do nível de qualificação do trabalhador, mas também da escassez de vagas adequadas à sua formação. Assim, trabalhadores altamente qualificados podem ser alocados a funções que demandam níveis educacionais inferiores, gerando perdas salariais que refletem uma má alocação de recursos humanos, podendo ser temporário ou persistente.

Alguns estudos empíricos se destacam na literatura nacional sobre o tema, como Santos (2002), por exemplo, que investigou a existência de *overeducation* ou *undereducation* em diferentes áreas de ocupação, além disso também analisou o fenômeno por gênero e região. Para a realização dessa análise, o autor fez uso dos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD/IBGE), dos anos de 1992, 1995, 1997 e 1999. Como resultado, o autor constatou que as mulheres são mais supereducadas do que os homens, pois apresentaram melhor *matching* entre educação e ocupação, ou seja, são menos subeducadas. Já em relação aos resultados da análise por região, o autor obteve que as regiões Sul e Sudeste apresentam maior proporção de trabalhadores supereducados.

Diaz e Machado (2008) analisam a ocorrência e os impactos da *overeducation* e da *undereducation*, que obtêm salários mais altos em uma ocupação do que os trabalhadores com a mesma escolaridade, no Brasil nas grandes regiões e no Estado de São Paulo, para isso utilizaram-se dados do Censo 2000. Os resultados mostraram que o Brasil está, como um todo, numa situação de *overeducation* (17,3% da população). E, em relação às diferentes regiões, constatou-se que a região Sul apresentou a maior taxa de sobre-educação (18,6%), já a região Sudeste apresentou a maior taxa de adequação, enquanto a região Nordeste apresentou a maior taxa de

undereducation, evidenciando os mesmos resultados encontrados por Santos (2002).

Annegues et al (2018) fizeram um estudo sobre a situação de *overeducation* dos egressos da Universidade Federal da Paraíba (UFPB) e analisaram as diferentes áreas de formação e como essas áreas afetam o tempo em que os indivíduos permanecem na situação de sobre-educação. Para fazer essa análise, os autores utilizam os métodos de *Propensity Score Matching* (PSM) para múltiplos tratamentos e modelos de análise de sobrevivência. Os resultados obtidos indicaram que os egressos das áreas de humanas, artes e ciências sociais aplicadas têm maiores chances de os egressos enfrentarem a situação de *overeducation* e menores chances de superar essa situação durante os primeiros anos de formados.

Annegues e Souza (2019) estudaram os efeitos salariais do *overeducation*, ou seja, investigaram qual foi o retorno salarial do trabalhador que se encontrava na situação de *overeducated*. Para alcançar tal objetivo, foram utilizados dados dos egressos da UFPB de 2003 a 2013, em conjunto com a base de dados da RAIS para os mesmos anos. Como resultado principal, mostrou-se que os alunos graduados que estão em situação de *overeducation* recebem em média um retorno salarial menor se comparado aos egressos que possuem um bom *match* entre qualificação e emprego.

Figueiredo Neto (2020) investigou os efeitos que a situação de *overeducation* causam sobre os rendimentos, uma vez que há diversos custos quando se está buscando um emprego e esse custo seria um dos motivos do trabalhador aceitar uma oferta de emprego que exija um nível abaixo ao possuído. Para esse estudo, o autor utilizou dois modelos: probit bivariado recursivo e uma equação minceriana. O autor utilizou dados da RAIS nos anos de 2006 e 2015 e da Classificação Brasileira de Ocupações (CBO). Os resultados encontrados foram que o setor de construção civil teve menor associação com *overeducation*, ao contrário do setor de agricultura e pesca que apresentaram maior associação.

Ferreira et al (2021) analisaram a probabilidade de os egressos cotistas da educação superior estarem em situação de *overeducation*. Para atingir tal objetivo foi utilizado o método *Propensity Score Matching* (PSM) e foram considerados como base de dados os alunos egressos dos cursos de graduação da UFPB em conjunto com dados do mercado de trabalho formal referentes ao ano de 2018. Os resultados apresentados mostraram que, ao analisar o conjunto de cursos da UFPB, a reserva de vagas destinadas às cotas não afetou a probabilidade de os egressos cotistas estarem em situação de *overeducation*. Outro resultado obtido através das análises foi que os egressos cotistas que concluíram cursos nas áreas de Ciências Humanas e Ciências Sociais Aplicadas apresentaram maiores chances de ocupar empregos que não exigem formação superior, indicando uma possível subutilização no mercado de trabalho.

Cavalcanti et al (2020), contribuíram sobre o tema fazendo um estudo sobre os efeitos do *overeducation* sobre os rendimentos dos trabalhadores graduados do Estado de São Paulo, a duração da condição de *overeducated* e a incidência da incompatibilidade educação-ocupação entre os anos de 2003 a 2013. A metodologia utilizada para tal objetivo foi uma análise de sobrevivência, através do modelo não-paramétrico de Kaplan-Meier e do modelo semi-paramétrico de Cox. Como resultado, os autores constataram que aproximadamente 27% da força de trabalho encontrava-se em estado de *overeducation*, porém com uma tendência de redução ao longo do período analisado. Em relação a análise de duração da situação de sobre-educação, os resultados apontaram que a probabilidade de permanecer no estado de *overeducated* diminui ao longo do tempo.

Almeida (2023) analisou a situação de sobre-educação entre jovens graduados

em Minas Gerais no período de 2012 a 2019. Para tanto, o autor utilizou dados da Relação Anual de Informações Sociais (RAIS) e o modelo de análise de sobrevivência, aplicando o conceito de *Job Analysis* para definir a situação de *overeducation*. O autor obteve como resultado principal que a sobre-educação está presente entre os jovens graduados do estado, mas apenas 32,1% dos indivíduos que estavam no contexto de *overeducation* conseguiram deixar essa condição ao longo dos oito anos.

Soares (2024) aplica o estudo de *overeducation* em uma análise conjunta com a política de ações afirmativas na UFRN, o objetivo principal foi investigar a compatibilidade, ou não, entre os egressos inseridos no mercado de trabalho e o grau de escolaridade adquirido com a formação universitária. Para tal, foram cruzados dados das admissões na UFRN com as informações contidas na RAIS e utilizado o método de Regressões por Descontinuidade (RDD) como estratégia de identificação. Os principais resultados que a autora obteve foram que ao fazer uma análise por gênero tem-se que os homens ingressantes na universidade e beneficiados pelas ações afirmativas ao concluírem o curso, possuem maior probabilidade de *overeducation* se comparados à mulheres que enfrentam a mesma situação.

Apesar do crescente número de estudos sobre a sobre-educação no Brasil, ainda são escassas as investigações que acompanham a trajetória dos indivíduos ao longo do tempo. A maioria das pesquisas concentra-se em analisar a incidência do fenômeno ou sua relação com características individuais e institucionais, sem explorar com profundidade os impactos salariais e a duração da condição de *overeducation* após a entrada no mercado de trabalho. Diante dessa lacuna, o presente estudo contribui ao examinar, de forma inédita no contexto da Universidade Federal do Rio Grande (FURG), a penalização salarial associada ao *overeducation* entre egressos recentes, bem como a duração dessa condição por meio de análise de sobrevivência. Este trabalho oferece evidências relevantes para o debate sobre a eficácia da formação superior e sua compatibilidade com as exigências do mercado de trabalho.

3. Metodologia

3.1 Pareamento por Score de Propensão (*Propensity Score Matching* - PSM)

Para analisar o impacto da condição de *overeducation* sobre a remuneração dos egressos da Universidade Federal do Rio Grande (FURG), este estudo emprega o método de *Propensity Score Matching* (PSM), desenvolvido por Rosenbaum e Rubin (1983). Essa técnica permite estimar efeitos causais a partir de dados observacionais, controlando o viés de seleção com base em características observáveis.

O PSM consiste em comparar dois grupos: o grupo de tratamento, composto por indivíduos que se encontram sobre-educados (isto é, com escolaridade acima da exigida para a ocupação exercida), e o grupo de controle, formado por egressos empregados em ocupações compatíveis com sua formação. O objetivo é verificar se há diferenças salariais sistemáticas entre esses grupos atribuíveis à condição de *overeducation*.

Conforme Cavalcanti (2015), o PSM permite estimar as consequências de um “tratamento” (neste caso, estar sobre-educado) ao comparar os indivíduos tratados com não tratados que tenham características semelhantes. Para isso, o método calcula, por meio de um modelo de regressão logística (*logit* ou *probit*), a probabilidade condicional de um indivíduo ser tratado, dado um conjunto de variáveis observáveis X. A variável dependente da regressão é uma *dummy* indicando se o indivíduo está ou não em condição de *overeducation*:

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

Em que D_i é uma variável indicadora de tratamento, que assume valor 1 se o indivíduo está sobre-educado e 0 caso contrário; X_i é um vetor de covariáveis observáveis que afetam a probabilidade de estar sobre-educado (como sexo, idade, desempenho acadêmico, instituto de formação, entre outros); β é o vetor de parâmetros associados a cada covariável; μ_i é o termo de erro aleatório; $G(\cdot)$ é uma função de distribuição acumulada (CDF) logística.

Para assegurar que o PSM produza estimativas válidas e não viesadas, é recomendável que duas hipóteses fundamentais sejam satisfeitas: a hipótese da independência condicional e a hipótese de suporte comum. Segundo Ferreira et al (2021), a hipótese da independência condicional estabelece que os resultados potenciais não dependem do status de tratamento, $(Y_{i0}, Y_{i1}) \perp D_i \vee X_i$, sendo $D_i \in \{0,1\}$ o indicador de tratamento; 1 = sobre-educado; 0 = adequadamente educado e X_i é o vetor de características observáveis. A segunda hipótese, denominada de suporte comum, implica que para cada indivíduo no grupo tratado tenha pelo menos um indivíduo correspondente no grupo de controle com score de propensão semelhante, isso garantirá uma comparação válida entre os grupos, $0 < Pr(D_i=1 \vee X_i) < 1$.

Uma vez satisfeitas essas condições, é possível estimar o efeito médio do tratamento sobre os tratados (ATT), definido como:

$$\hat{t}_{ATT} = E[Y_{i1} - \hat{Y}_{i0} \vee D_i=1] = E[Y_{i1} \vee D_i=1] - E[\hat{Y}_{i0} \vee D_i=1] \quad (2)$$

O efeito médio é obtido através da diferença de resultado do grupo de tratamento e do grupo de controle, ou seja, o ATT corresponde à diferença média nos salários entre os indivíduos sobre-educados e seus contrafactuais adequadamente educados, com características observáveis semelhantes.

Este estudo utiliza três técnicas de pareamento para estimar o ATT: Pareamento por Vizinho Mais Próximo (Nearest Neighbor Matching), em que cada indivíduo tratado é pareado com aquele do grupo de controle que possui o escore de propensão mais próximo. Serão consideradas versões com e sem reposição, bem como pareamento com os cinco vizinhos mais próximos; Radius Matching que impõe um limite (raio) na distância máxima entre os escores de propensão, permitindo que apenas controles dentro desse intervalo sejam usados como comparadores; Kernel Matching que utiliza uma média ponderada de todos os não tratados, com pesos atribuídos inversamente à distância do escore de propensão.

3.2 Análise de Sensibilidade de Rosenbaum

Embora o método de *Propensity Score Matching* (PSM) seja eficaz para controlar o viés de seleção em relação a características observáveis, ele não elimina a possibilidade de viés de omissão de variáveis não observadas. Para lidar com esse risco, este estudo emprega a análise de sensibilidade de Rosenbaum (2002), que avalia o quanto os resultados do pareamento são robustos à presença de fatores não

observados que possam afetar simultaneamente a probabilidade de tratamento e o desfecho.

A lógica da análise baseia-se na ideia de que, mesmo após o pareamento, dois indivíduos i e j com as mesmas covariáveis observáveis X podem ter probabilidades distintas de tratamento devido a variáveis não observadas u . A razão entre as probabilidades p de tratamento dos indivíduos i e j pode ser expressa como:

$$\frac{\frac{p_i}{1-p_j}}{\frac{p_j}{1-p_j}} = \exp(\gamma(u_i - u_j)) \quad (3)$$

Em que p_i e p_j são as probabilidades de tratamento dos indivíduos i e j , respectivamente; γ representa o impacto da variável não observada u sobre a probabilidade de tratamento. Como u_i e u_j são não observáveis, Rosenbaum propõe uma abordagem baseada em limites. Assume-se que, para indivíduos pareados, a diferença máxima entre suas probabilidades de tratamento é controlada por um parâmetro $\Gamma \geq 1$, de forma que:

$$\frac{1}{\Gamma} \leq \frac{\frac{p_i}{1-p_j}}{\frac{p_j}{1-p_j}} \leq \Gamma \quad (4)$$

Quando $\Gamma = 1$, não há viés de omissão e os pares são perfeitamente comparáveis; à medida que Γ aumenta, cresce a possibilidade de viés oculto afetar o resultado. O objetivo é encontrar o valor crítico de Γ a partir do qual o resultado estimado (por exemplo, o efeito médio do tratamento sobre os tratados, $\hat{\tau}_{ATT}$) deixa de ser estatisticamente significativo. Portanto, se os resultados permanecerem significativos até, por exemplo, $\Gamma = 1,5$, isso indica que seria necessário um viés não observado com força moderada para invalidar a inferência causal do estudo. Por outro lado, se a significância for perdida com Γ próximo de 1, os resultados são considerados sensíveis à omissão de variáveis.

A robustez dos resultados pode ser sumarizada assim: i) Alta robustez: resultado permanece significativo até valores elevados de Γ (ex. $\Gamma \geq 2$); ii) Robustez moderada: resultado se mantém até $\Gamma \in [1,3,1,5]$; iii) Baixa robustez: resultado perde significância com $\Gamma \approx 1$, indicando alta sensibilidade.

3.4 Base de Dados e Tratamentos

Para este estudo são usadas duas fontes de dados: i) informações acadêmicas dos egressos da Universidade Federal do Rio Grande (FURG) que se formaram entre 2017 e 2022 e que são fornecidas pelo Centro de Gestão de Tecnologia da Informação

(CGTI) da FURG e ii) informações sobre a participação do egresso no mercado de trabalho formal presentes na Relação Anual de Informações Sociais (RAIS) Identificada de 2022. As informações que constam no sistema acadêmico da FURG sobre os egressos são turno, curso, coeficiente de rendimento acadêmico, ano e forma de ingresso, cor, município de nascimento etc. Já a RAIS apresenta as características dos trabalhadores, além de o tipo de ocupação, salário, horas de trabalho, classificação da atividade e características sobre a empresa, como a natureza jurídica, seu porte etc.

Para identificar o egresso na RAIS foi usado o Cadastro de Pessoas Físicas (CPF), que consta nas informações acadêmicas do egresso, permitindo construir uma base de dados geral em que conste informações dos egressos e do mercado de trabalho formal. Para calcular a penalização salarial, identificou-se os egressos formados entre 2017 e 2022 apenas na RAIS de 2022 e para a análise de sobrevivência foi necessário acompanhar os formados em cada RAIS dos anos de 2018 até 2022.

Para definir a variável de *overeducation* foi usada a abordagem do *Job Analysis* que define o nível de escolaridade requerido para cada ocupação, assim como em Ferreira et al (2021). Na RAIS é possível identificar a ocupação do egresso usado a Classificação Brasileira de Ocupações (CBO) que, em sua última edição de 2002, utiliza uma nova metodologia de classificação e faz a revisão e atualização completas de seu conteúdo, dividindo as ocupações registradas em 10 grandes grupos por semelhança e competência nas atividades desenvolvidas, 49 subgrupos principais, 195 subgrupos, 625 famílias e mais de 2600 ocupações (Almeida, 2023), como pode ser conferido no Quadro1. Por meio da CBO, será verificado o nível de escolaridade exigido para cada ocupação, dessa forma a variável *overeducation* assume o valor 1 se a ocupação do egresso na CBO exigir nível de escolaridade menor que o ensino superior e zero caso contrário.

Quadro 1: Grandes Grupos de Ocupações

Código	Título
0	Membros das forças armadas, policiais e bombeiros militares
1	Membros superiores do poder público, dirigentes de organizações de interesse público e de empresas, gerentes
2	Profissionais das ciências e das artes
3	Técnicos de nível médio
4	Trabalhadores de serviços administrativos
5	Trabalhadores dos serviços, vendedores do comércio em lojas e mercados
6	Trabalhadores agropecuários, florestais e da pesca
7	Trabalhadores da produção de bens e serviços industriais
8	Trabalhadores da produção de bens e serviços industriais
9	Trabalhadores em serviços de reparação e manutenção

Fonte: CBO - Ministério do Trabalho e do Emprego

Após realizar os procedimentos de seleção já mencionados a amostra

construída para esse trabalho resultou em 3.152 observações.

Em relação as variáveis descritivas do modelo, foram empregadas características dos egressos como sexo, idade, forma de ingresso, e características dos cursos como turno e duração do curso, incluindo também o coeficiente de rendimento dos alunos, entre outras variáveis que podem ser visualizadas abaixo no quadro 2:

Quadro 2: Estatística descritiva das variáveis usadas no pareamento

Variáveis demográficas	Mean (Std.dev)
homem	.4109249 (.492078)
idade	30.8473 (6.98459)
Branca	.7811918 (.4135024)
Campus	
SVP	.0450031 (.2073431)
SLP	.0180012 (.1329763)
SAP	.0270019 (.1621139)
RGr/Saúde..	.0841092 (.2775947)
Forma de ingresso	
PSVO	.0704531 (.2559489)
VagaRema	.0170701 (.1295529)
Outros	.0595903 (.236763)
Turno	
integral	.4972067 (.5000698)
manhã	.0996276 (.2995493)
duracaocurso	4.613284 (1.684356)
Variáveis Setoriais	
Indústria	.1731844 (.3784653)
Construção	.0325885 (.1775844)
Comércio	.0822471 (.2747834)
Serviços	.7073246 (.4550613)
Variáveis de Grupo Ocupacional	
Dirigentes	.0490379 (.2159806)
Profis.	.3432651 (.4748728)
Técnicos	.2420857 (.4284124)
Administr.	.2091868 (.4067912)
Serviços	.0853507 (.2794462)
Operadores	.0158287 (.1248318)
Outras variáveis	
naturalrio	.446617 (.4972193)
lcoefrend	2.002103 (.2800113)
tempoemprego	28.36872 (48.25359)

trab_riogr	.3258845 (.4687771)
lsalhora	3.059712 (.8558627)
Grandes áreas de estudo	
ccsa	.1945996 (.3959536)
ch	.273743 (.4459479)
cexatas	.3798883 (.4854341)

Fonte: Elaboração própria através de dados da RAIS e dos dados da FURG, software Stata.

Ao fazer uma análise descritiva das covariáveis podemos inferir que em relação as variáveis demográficas, tem-se que cerca de 41,1% da amostra é composta por homens, a idade média dos indivíduos é de aproximadamente 30,85 anos, cerca de 78,1% dos indivíduos se declaram brancos e 44,7% são naturais da cidade de Rio Grande. Uma observação importante é que a Universidade Federal do Rio Grande possui os campus Carreiros e campus Saúde localizados na cidade de Rio Grande, porém por ser uma universidade multi campis, existem outros campus que estão localizados nas cidades de Santa Vitória do Palmar, São Lourenço do Sul e Santo Antônio da Patrulha, porém a maioria dos egressos (82%) concluíram suas graduações no campus Carreiros. Analisando pela ótica do mercado de trabalho, observou-se que a maioria dos egressos (70,7%) se encontra empregado no setor de serviços e uma minoria (3,2%) no setor de construção, onde 32,6% dos indivíduos da amostra estão inseridos no mercado de trabalho da cidade. Além disso, destaca-se que foram agrupados os cursos ofertados pela FURG em quatro grandes áreas, são elas: ccsa (cursos relacionados a ciências sociais aplicadas) que conta com um percentual de 19,5% dos egressos formados, ch (cursos voltado a área de humanas), cexatas (cursos relacionados as ciências exatas) e cs_bio_oceano (cursos relacionados à saúde, biológica e oceano), que representa cerca de 15% da amostra.

4. Resultados

Para testar a relevância dos resultados aqui apresentados, foi feita uma análise do balanceamento das covariáveis após o pareamento realizado pelo PSM, que pode ser conferido através da análise gráfica de viés padronizado (antes e depois do pareamento). A Figura A1 do apêndice permite verificar graficamente as diferenças antes e depois do pareamento, no qual foram analisados os diferentes métodos para cada análise desenvolvida neste trabalho e anexadas apenas aquela que mostrou-se mais adequada.

Ao analisar essas estatísticas, é possível inferir que, com exceção do método *Radius Matching* que não foi eficaz em melhorar o balanceamento entre os grupos tratados e controle, todos os indicadores permaneceram praticamente inalterados em relação à amostra original (Unmatched), enquanto todos os outros métodos se mostraram eficazes e forneceram estatísticas significativas de pareamento das covariáveis. O método que se mostrou mais eficaz foi *Nearest Neighbor Matching* (SR), pois apresentou melhor desempenho, com menor viés médio e maior homogeneidade entre os grupos. No caso dos demais métodos *Nearest Neighbor Matching* (CR), *Nearest Neighbor Matching* (CR); K=5 e *Kernel Matching* mostraram-

se eficazes uma vez que também produziram estimativas consistentes e com qualidade de pareamento satisfatória.

O resultado do efeito médio do tratamento pode ser conferido no quadro 3. O ATT indica quanto o salário-hora médio aumenta ou diminui para o egresso em situação de *overeducation*, se comparado com o que receberia caso estivesse adequadamente educado. É possível inferir que há uma penalização salarial significativa para os indivíduos classificados como sobre-educados se comparados àqueles cuja escolaridade é compatível com o nível exigido para o posto de trabalho. Após a conversão dos coeficientes logarítmicos (ATT) para variações percentuais reais, verifica-se que os indivíduos que estão na situação de *overeducation* recebem, em média, salários-hora entre 39,9% e 47,5% menores do que seus pares adequadamente alocados, dependendo do método utilizado. Após essa primeira análise, pode-se concluir que esses resultados reforçam a evidência de que a sobre-educação está associada a uma expressiva perda salarial.

Ao desagregar os efeitos do *overeducation* ao longo da distribuição de salários-hora, observou-se um padrão desigual de penalização. Entre os indivíduos pertencentes ao primeiro quartil (Q25), os efeitos médios estimados foram positivos ou nulos, indicando que, em contextos de baixa remuneração, a escolaridade acima da exigida para o cargo não acarreta penalização imediata e pode, inclusive, representar um diferencial competitivo. No entanto, no terceiro quartil (Q75), correspondente aos salários mais elevados, os efeitos da *overeducation* foram sistematicamente negativos e significativos. O método de *Radius Matching*, por exemplo, estimou uma penalidade de $-0,1825$ no logaritmo do salário-hora ($p < 0,01$), o que corresponde a uma perda real superior a 16%. Tais evidências sugerem que os efeitos adversos da *overeducation* são mais intensos para trabalhadores qualificados que não conseguem alocação condizente com sua formação, o que reforça a hipótese de subutilização do capital humano nos segmentos mais competitivos do mercado de trabalho.

Quadro 3: Efeito Médio do Tratamento (ATT)

Método de pareamento	Efeito Médio	Porcentagem real (%)	Quantil salarial	
			0.25	0.75
Nearest Neighbor Matching (CR), K=1	-0.5106***	≈ -39.98%	0.0275	-0.1000
Nearest Neighbor Matching (CR); K=5	-0.5388***	≈ -41.65%	0.0313	-0.0882
Nearest Neighbor Matching (SR)	-0.5235***	≈ -40.75%	-0.0245	-0.0651
Radius Matching	-0.6457***	≈ -47.57%	-0.0135	-0.1825***
Kernel Matching	-0.5472***	≈ -42.14%	-0.0128	-0.1313**

Fonte: Elaboração própria com base nos dados da RAIS, FURG através do software Stata. Nota: CR: com reposição; SR: sem reposição; caliper 0.001; K: número de vizinhos próximos. *: $p < 0.1$; **: $p < 0.05$; ***: $p < 0.01$.

Como mencionado anteriormente, é recomendado que se faça o teste de sensibilidade de Rosenbaum para verificar se há variáveis omitidas que poderiam impactar negativamente no modelo, ou seja, que poderiam gerar resultados viesados. O quadro 4 reúne os resultados do teste nos diferentes pareamentos especificando as estatísticas dos limites (Γ) que variaram de 1,0 a 2,0 e para cada valor de Γ são calculados a significância estatística. Observou-se que os resultados em relação ao

efeito médio, evidenciaram que, mesmo sob o limite extremo testado ($\Gamma=2$), os valores de p- permaneceram iguais a zero ($p<0.01$) isso significa que a inferência sobre o impacto da *overeducation* não está comprometido por omissão de variáveis relevantes, conclui-se então que as covariáveis observadas explicam adequadamente a atribuição ao tratamento, indicando alta robustez dos modelos.

Ao observar os resultados de robustez dos efeitos do tratamento para o quantil de renda de 0.75, que nesse caso representam os indivíduos que recebem maiores salários, é possível constatar que os métodos *Radius* ($p<0.01$) e *Kernel* ($p<0.05$), apresentam resultados robustos, comprovando a existência de um efeito negativo estatisticamente significativo para o grupo de maior renda.

Quadro 4: Análise de sensibilidade para o efeito médio e quantil 0,75

Modelo		Efeito médio	Quantil 0,75
Nearest Neighbor Matching (CR); K=1	Limites (Γ)	2	
	p-	3.7e-09	
Nearest Neighbor Matching (CR); K=5	Limites (Γ)	2	
	p-	0	
Nearest Neighbor Matching (SR)	Limites (Γ)	2	
	p-	3.0e-07	
Radius Matching	Limites (Γ)	2	2
	p-	0	0.0095
Kernel Matching	Limites (Γ)	2	1.9
	p-	0	0.0693

Fonte: teste rbounds, elaboração própria com base nos dados da RAIS, FURG através do software Stata. Nota: CR: com reposição; SR: sem reposição; caliper 0.001; K: número de vizinhos considerados na estimação

O quadro 5 mostra o resultado desagregados dos efeitos do *overeducation* por área do conhecimento e revela penalidades salariais relevantes e estatisticamente significativas em todas as grandes áreas, com destaque para os egressos das Ciências Exatas e Tecnológicas (CEXATAS) e das Ciências Sociais Aplicadas (CCSA). Para estes grupos, os efeitos médios variaram entre -0.5099 a -0.5332 e 0.4529 a 0.5782 , respectivamente no logaritmo do salário-hora, o que equivale a reduções reais entre 36% e 44% nos rendimentos dos sobre-educados. Já nas Ciências Humanas (CH), a penalidade também foi sistematicamente negativa, com valores entre -0.4406 e -0.5665 . A maior penalidade individual foi observada no método *Radius Matching* para a quarta área analisada, no qual englobam cursos voltados para área da saúde, biológicas e oceanográficas (-1.179), indicando uma perda de mais de 60%, embora esse valor deva ser interpretado com cautela, dada a possível heterogeneidade ou tamanho reduzido da amostra. Esses achados reforçam que os efeitos da alocação inadequada variam conforme a área de formação, sendo mais severos em cursos com maior exigência técnica e formação especializada.

Quadro 5: Efeito Médio do Tratamento por área do conhecimento

Método de pareamento	CCSA	CH	CEXATAS	CS_BIO_O CE
Nearest Neighbor Matching (CR), K=1	-0.5061***	-0.4607***	-0.5257***	-0.5448**
Nearest Neighbor Matching (CR); K=5	-0.5078***	-0.4911***	-0.5099***	-0.4309*
Nearest Neighbor Matching (SR)	-0.5782***	-0.4795***	-0.5332***	-0.7530***
Radius Matching	-0.4871***	-0.5665***	-0.5177***	-1.1798***
Kernel Matching	-0.4529***	-0.4406***	-0.5319***	-0.4879***

Fonte:Elaboração própria com base nos dados da RAIS, FURG através do software Stata. Nota: CR: com reposição; SR: sem reposição; caliper 0.001; K: número de vizinhos considerados na estimação; *: p<0.1; **: p<0.05; ***: p<0.01. Áreas de conhecimento: CCSA: Ciências Sociais Aplicadas; CH: Ciências Humanas; CEXATAS: Ciências Exatas e Tecnológicas; CS_BIO_OCE: Ciências da Saúde, Biológicas e Oceanográficas.

Com base na análise de sensibilidade de Rosenbaum (rbounds) para as diferentes grandes áreas de conhecimento, buscou-se avaliar a robustez dos resultados do efeito médio, rodados anteriormente, frente a potenciais vieses de seleção não observados. Analisando o quadro 6, verifica-se que os resultados demonstraram que todos os métodos foram capazes de manter a significância estatística de seus respectivos efeitos (com p-valores extremamente baixos, frequentemente próximos de zero) mesmo sob cenários de elevado viés de seleção não observado, onde o parâmetro Γ atingiu seu limite máximo de 2. Dessa forma, ao fazer a análise de sensibilidade podemos inferir que os efeitos da *overeducation* nos salários quando observado nas grandes áreas de conhecimento produziram resultados válidos e robustos.

Quadro 6: Análise de sensibilidade para grandes áreas de conhecimento

Método de pareamento		CCSA	CH	CEXATA S	CS_BIO_OC E
Nearest Neighbor Matching (CR)	Limites (Γ)	2	2	2	1.55
	p-	0.0242	0.0004	0.0000	0.0938
Nearest Neighbor Matching (CR); K=5	Limites (Γ)	2	2	2	1.35
	p-	0.0058	7.8e-06	1.3e-06	0.0914
Nearest Neighbor Matching (SR)	Limites (Γ)	2	2	2	2
	p-	0.0269	0.0015	0.0004	0.0779
Radius Matching	Limites (Γ)	2	2	2	2
	p-	2.0e-13	0	0	2.3e-15
Kernel Matching	Limites (Γ)	2	2	2	2
	p-	3.1e-11	1.4e-12	0	5.0e-07

Fonte:Elaboração própria com base nos dados da RAIS, FURG através do software Stata. Nota: CR: com reposição;

SR: sem reposição; caliper 0.001; K: número de vizinhos considerados na estimação. Áreas de conhecimento: CCSA: Ciências Sociais Aplicadas; CH: Ciências Humanas; CEXATAS: Ciências Exatas e Tecnológicas; CS_BIO_OCE: Ciências da Saúde, Biológicas e Oceanográficas.

5. Considerações Finais

É observável que o acesso à educação de nível superior vem crescendo ao longo dos anos o que pode ser decorrente de políticas educacionais que facilitaram o acesso para a população. Esse aumento pode ser também visto pelo senso comum brasileiro uma vez que a grande maioria da população não teve avós ou pais que frequentaram uma universidade, e é visível também através de estudos científicos que comprovam que esse acesso vem crescendo, como dados estatísticos disponibilizados pelo Ministério da Educação, assim como grandes estudos na área. Porém, também é perceptível que a oferta de empregos para suprir essa demanda de trabalhadores qualificados não vem crescendo na mesma proporção. É, talvez, por conta desse desequilíbrio existente que ocorra a *overeducation*, ou seja, pessoas que possuem uma qualificação superior, mas que estão empregadas em trabalhos que não exigiriam tal qualificação.

A discussão sobre *overeducation* se faz necessária uma vez que traz consequências negativas para os trabalhadores, como por exemplo o desperdício de capital humano, a redução da satisfação e da motivação profissional e a oferta de salários inferiores ao esperado para o nível de qualificação, são algumas das implicações citadas por Almeida (2023). Enfatiza-se que é através desse tipo de pesquisa que podemos analisar como está a relação entre ocupação e educação e em como isso pode impactar nos investimentos, alocação de recursos e talvez na criação futura de políticas públicas para uma melhor inserção dos trabalhadores no mercado de trabalho.

Posto isso, este trabalho visou contribuir para o tema com uma análise pela ótica dos salários, ou seja, este artigo propôs investigar se ocorreria alguma penalidade sobre os salários caso os egressos da FURG se encontrassem sobre-educados. Para alcançar o objetivo principal se fez o uso do método *Propensity Score Matching*, com os diferentes métodos de pareamento e em conjunto com a Análise de Sensibilidade usada para enfatizar a robustez do modelo. Em síntese, este estudo forneceu uma análise detalhada sobre os desdobramentos salariais associada à *overeducation*, pois verificou-se que há uma perda salarial média de 42% se comparado aos egressos cuja qualificação é adequada à ocupação. Ao fazer um estudo mais detalhado e analisar o impacto por quantil de renda, observa-se que esse efeito é menos acentuado naqueles egressos que possuem menores salários, porém constatou-se indícios de que haveria uma penalização salarial para aqueles que recebem uma maior remuneração, verificado através do método *Kernel Matching*, no qual há uma perda salarial de pelo menos 12%. Ainda evidenciamos qual seria a penalização salarial por grandes áreas de conhecimento, foi constatado que a redução salarial é evidente em todas as áreas, com uma ênfase maior para a área de Ciências Exatas.

Assim, através desses resultados preliminares podemos concluir, inicialmente, que egressos da FURG ao estarem em uma situação de *overeducation* recebem salários menores se comparados àqueles que estão alocados adequadamente na relação trabalho/escolaridade. Dessa forma, os resultados mostram que há uma forte penalização salarial para aqueles indivíduos que se encontram sobre-educados, com variações em sua magnitude ao comparar os

diferentes perfis de renda e as grandes áreas de atuação.

Referências

ALMEIDA, J. C. das M. **A sobreeducação entre os jovens com ensino superior de Minas Gerais entre 2012 e 2019: uma análise da persistência e dos determinantes da saída da condição.** 2023. Dissertação (Mestrado em Economia) – Universidade Federal de Minas Gerais, Belo Horizonte, 2023. Disponível em: <https://repositorio.ufmg.br/handle/1843/64708>

ANNEGUES, A. C.; OLIVEIRA, C.; FIGUEIREDO, E.; PORTO JÚNIOR, S. **Overeducation e área de formação: evidências para os egressos da UFPB.** In: Encontro Nacional de Economia, 46., 2018, Rio de Janeiro. Disponível em: https://www.anpec.org.br/encontro/2018/submissao/files_l/i134854d3bdd4525061d4dca002d4bfa207.pdf

ANNEGUES, A. C.; SOUZA, W. P. S. de F. **Retorno salarial do overeducation: viés de seleção ou penalização ao excesso de escolaridade?** Revista Brasileira de Economia, Rio de Janeiro, v. 72, n. 2, p. 119–138, 2019. Disponível em: <https://periodicos.fgv.br/rbe/article/view/77646/77951>

CAVALCANTI, I. T. do N. **Análise do diferencial de desempenho entre estudantes cotistas e não cotistas da UFBA pelo Propensity Score Matching.** 2015. 126 f. Dissertação (Mestrado em Economia) – Universidade Federal da Bahia, Salvador, 2015. Disponível em: <https://repositorio.ufba.br/handle/ri/18125>

CAVALCANTI, G. da S; ARAÚJO, J. R.; ALVES, J. da S.; SILVEIRA NETO, R. da M. **Overeducation no estado de São Paulo: uma análise com dados em painel.** In: ENCONTRO NACIONAL DE ECONOMIA, 48., 2020. Disponível em: https://www.anpec.org.br/encontro/2020/submissao/files_l/i13-df5ff684620655051f7883c214dbab12.pdf

DIAZ, M. D. M.; MACHADO, L. **Overeducation e undereducation no Brasil: incidência e retornos.** Estudos Econômicos (São Paulo), São Paulo, Brasil, v. 38, n. 3, p. 431–460, 2008. Disponível em: <https://revistas.usp.br/ee/article/view/35946>.

FERREIRA, F. D. da S.; ALMEIDA, A. T. C.; ANNEGUES, A. C. **Política afirmativa e overeducation: uma análise para egressos da UFPB.** In: ENCONTRO NACIONAL DE ECONOMIA, 49., 2021. Disponível em: https://www.anpec.org.br/encontro/2021/submissao/files_l/i12-9a3dc4873e3a81fe58b539de515e9b52.pdf

FIGUEIREDO NETO, J. E. **Identificando overeducation no mercado de trabalho brasileiro.** 2020. Monografia (Graduação em Ciências Econômicas) – Universidade Federal do Ceará, Fortaleza, 2020. Disponível em: <https://repositorio.ufc.br/handle/riufc/65687>

LIMA, V. **O drama do mercado de trabalho para recém-formados.** Núcleo Brasileiro

de Estágios, São Paulo, 14 abr. 2021.

PEREIRA, G. M. C. et al. **Panorama de oportunidades para os egressos do ensino superior no Brasil: o papel da inovação na criação de novos mercados de trabalho.** Ensaio: Avaliação em Políticas Públicas Educacionais, v. 24, n. 90, p. 179-198, 2016.

SANTOS, A. M. dos. **Overeducation no mercado de trabalho brasileiro.** Revista Brasileira de Economia de Empresas, São Paulo, v. 2, n. 2, p. 61–80, maio/ago. 2002. Disponível em: <file:///C:/Users/154769-PC/Downloads/4377-Texto%20do%20artigo-16505-1-10-20130913.pdf>

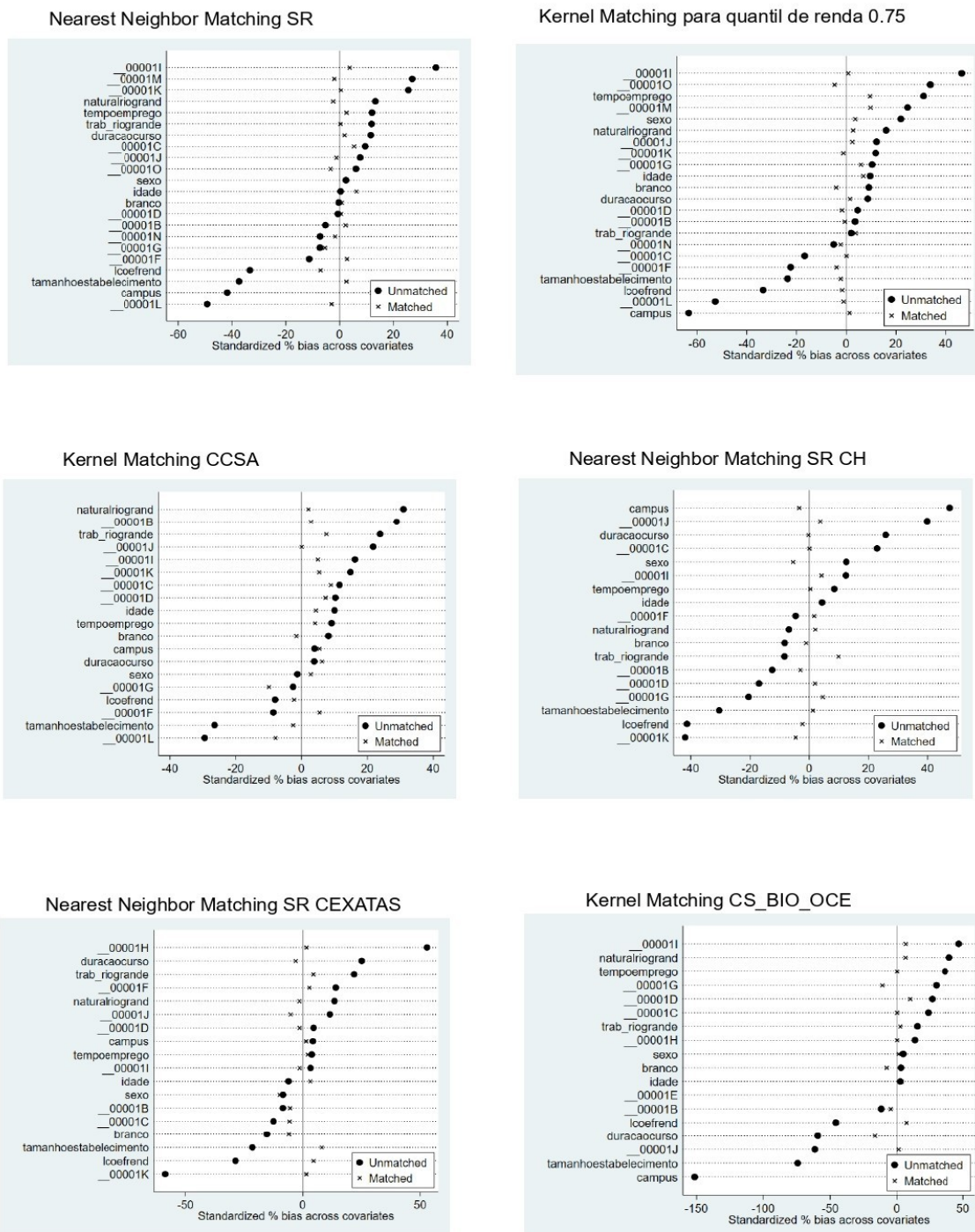
SILVA, E. C. da. **Teoria do capital humano e a relação educação e capitalismo.** In: SEMINÁRIO DE PESQUISA EM EDUCAÇÃO DA REGIÃO SUL – SEPECH, 2009. Londrina: Universidade Estadual de Londrina, 2009. Disponível em: https://www.uel.br/eventos/sepech/sumarios/temas/teoria_do_capital_humano_e_a_relacao_educacao_e_capitalismo.pdf

SIMÕES, R. A. **Ensino superior e mercado de trabalho.** Dissertação (Mestrado em Economia) – Universidade Federal do Ceará, Fortaleza, 1985.

SOARES, A. M. J. **Overeducation entre beneficiários de ações afirmativas: evidências em uma universidade pública brasileira.** Revista de Administração Pública, Rio de Janeiro, v. 58, n. 6, p. e2024–0143, 2024. DOI: 10.1590/0034-761220240143. Disponível em: <https://periodicos.fgv.br/rap/article/view/92646>.

Apêndice

Figura A1: Gráficos de balanceamento das covariáveis



Fonte: Elaboração própria com base nos dados da RAIS, FURG através do software Stata. Nota: CR: com reposição; SR: sem reposição; caliper 0.001; K representa os número de vizinhos próximos. Áreas de conhecimento: CCSA: Ciências Sociais Aplicadas; CH: Ciências Humanas; CEXATAS: Ciências Exatas e Tecnológicas; CS_BIO_OCE: Ciências da Saúde, Biológicas e Oceanográficas.