

NEDUR

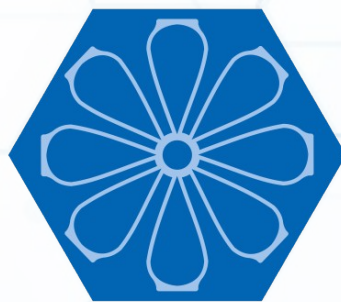
TD NEDUR-UFPR N° 01-2024

**Implicações do Mismatch Educacional por
Gênero no Mercado de Trabalho Formal:
uma análise para egressos da UFPR**

Cibele De Biasi, Kênia Barreiro de Souza

Dezembro/2024





NEDUR

O NEDUR – Núcleo de Estudos em Desenvolvimento Urbano e Regional é um núcleo de pesquisa da Universidade Federal do Paraná, vinculado ao Programa de Pós-Graduação em Desenvolvimento Econômico e ao Departamento de Economia, que tem por finalidade realizar pesquisas aplicadas de excelência no campo da Economia e sua interface com a Ciência Regional e Urbana, visando contribuir para o desenvolvimento socioeconômico do Brasil, como também de suas regiões e cidades. Dentre os princípios que regem a prática de pesquisa no âmbito do NEDUR, destaca-se a rigorosidade técnico-científica, privilegiando a sinergia entre fundamentos teóricos e métodos quantitativos de análise empírica, e o compromisso com a ética e transparência no processo de produção e disseminação do conhecimento científico. O NEDUR-UFPR foi criado em setembro de 2013, congregando um conjunto de docentes, discentes e pesquisadores que desenvolvem pesquisas direta e indiretamente relacionadas com a Ciência Regional e Urbana.

Na utilização ou citação de partes do documento é obrigatório referenciar os autores do trabalho:

De Biasi, C.; Souza, K. B. Implicações do Mismatch Educacional por Gênero no Mercado de Trabalho Formal: uma análise para egressos da UFPR. Texto para Discussão NEDUR-UFPR N° 01-2024, Núcleo de Estudos em Desenvolvimento Urbano e Regional (NEDUR) da Universidade Federal do Paraná, Curitiba, Dezembro/2024.

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade do(s) autor(es), não exprimindo, necessariamente, o ponto de vista do Núcleo de Estudos em Desenvolvimento Urbano e Regional e da Universidade Federal do Paraná.



Implicações do *Mismatch* Educacional por Gênero no Mercado de Trabalho Formal: uma análise para egressos da UFPR

Cibele De Biasi^Δ, Kênia Barreiro de Souza^Φ

Resumo

O artigo investiga o *mismatch* educacional entre homens e mulheres no mercado de trabalho formal brasileiro. A partir de um painel de dados que integrou informações de egressos da UFPR, como área de formação, às suas informações no mercado de trabalho formal da base de dados da Relação Anual de Informações Sociais (RAIS) identificada entre 2003 e 2021, como ocupação e remuneração, os resultados encontrados apontam que o *mismatch* educacional influencia negativamente os resultados no mercado de trabalho para todos os indivíduos, não obstante, as mulheres são ligeiramente mais penalizadas do que os homens. Além disso, quando as análises são feitas separadamente para cada área, os resultados indicam que áreas caracterizadas por habilidades específicas, como a Saúde e Engenharias, apresentam efeitos negativos significativamente mais acentuados no que diz respeito ao *mismatch* horizontal em comparação às demais áreas de habilidades menos específicas, como Humanas, Linguística; Letras e Artes e Sociais Aplicadas.

Palavras-chave: Desigualdade de gênero; *Mismatch* educacional; Gap salarial.

Código JEL: J31, J71.

^Δ Doutora em Desenvolvimento Econômico pelo Programa de Pós-Graduação em Desenvolvimento Econômico (PPGDE) da UFPR. E-mail: cibelebiasi@hotmail.com.

^Φ Professora do Departamento de Economia e do Programa de Pós-Graduação em Desenvolvimento Econômico (PPGDE) da UFPR, Pesquisadora PQ-2 do CNPq e Pesquisadora do NEDUR. E-mail: keniadesouza@ufpr.br.



Implications of Educational Mismatch by Gender in the Formal Labor Market: an analysis for UFPR graduates

Cibele De Biasi, Kênia Barreiro de Souza

Abstract

The essay investigates educational mismatch between men and women in the formal Brazilian labor market. Using a panel of data that integrated information from UFPR graduates, such as field of study, with their information in the formal labor market from the database of the identified Annual Social Information Report (RAIS) from 2003 to 2021, such as occupation and remuneration, the findings indicate that educational mismatch negatively influences outcomes in the labor market for all individuals, nevertheless women are slightly more penalized than men. Additionally, when analyses are conducted separately for each field, the results indicate that areas characterized by specific skills, such as Health and Engineering, present significantly more pronounced negative effects regarding horizontal mismatch compared to other areas with less specific skills, such as Humanities, Linguistics, Letters and Arts, and Applied Social Sciences.

Keywords: Gender inequality; Educational mismatch; Wage gap.

JEL Code: J31, J71.



1. Introdução

A média de anos de estudo da força de trabalho no Brasil experimentou um aumento significativo nas últimas décadas e esse padrão se reflete, também, no aumento constante da participação de trabalhadores com educação superior. Conforme dados do Censo Demográfico, na década de 1980, apenas 5% dos indivíduos ocupados no Brasil possuíam ensino superior. Esse percentual aumentou para 8,5% na década de 1990, passando para 9,5% no início do novo milênio e chegando em 15% em 2010 (REIS, 2021).

O crescimento observado a partir dos anos 2000 foi promovido, entre outras coisas, a partir de políticas sociais voltadas para atender a parcela marginalizada e de baixa renda da população que não tinha condições de frequentar níveis mais elevados de ensino. Por exemplo, o Programa Universidade Para Todos (PROUNI) concedeu aos alunos de baixa renda financiamento total ou parcial para o ingresso em universidades particulares; já o Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais (REUNI), por sua vez, fomentou o investimento público no ensino superior e objetivou torná-lo mais inclusivo ao expandir o número de vagas e reduzir as desigualdades regionais. Além disso, contribuíram, também, as políticas afirmativas de cotas para alunos das escolas públicas e para a população negra. Conforme o INEP, essas políticas foram responsáveis por aumentar em aproximadamente 45% o número de matrículas no ensino superior entre 2003 e 2013 (MARIONI, 2020).

Contudo, quando a demanda por trabalho qualificado não acompanha o aumento da oferta de mão de obra educada, surge o problema do *mismatch* educacional. Este pode ser tanto de natureza vertical, quando um indivíduo está trabalhando em uma ocupação que requer mais (*undereducation* – subeducação) ou menos (*overeducation* – sobre-educação) anos de estudo do que efetivamente possui, quanto de natureza horizontal, que ocorre quando o indivíduo está empregado em uma ocupação não relacionada com a sua área de formação acadêmica. O presente artigo focará nos sobre-educados (*mismatch* vertical) e no *mismatch* horizontal.

De acordo com Carvalho e Reis (2023), cerca de 25% dos trabalhadores ocupados no Brasil em 2012 eram sobre-educados e essa porcentagem aumentou em 11 pontos percentuais em 2022. Os autores apontam também que, considerando o nível educacional



dos indivíduos, os sobre-educados com ensino médio completo eram aproximadamente 42% em 2012, aumentando em 9 pontos percentuais em 2022 e os sobre-educados com ensino superior completo eram cerca de 28%, aumentando em 7 pontos percentuais dez anos depois.

A sobre-educação é um problema na medida em que indica que parte do investimento em educação não está sendo aproveitado pelo mercado de trabalho. A literatura existente aponta que os indivíduos sobre-educados não são remunerados da mesma forma que aqueles que possuem a escolaridade adequada à função que desempenham (BAUER, 2002; KORPI, TÅHLIN, 2009; TSAI, 2010; REIS, 2017; MARIONI, 2020). Além disso, eles são também mais propensos a ser insatisfeitos em relação aos seus trabalhos (VERHOFSTADT; DE WITTE; OMEY, 2007; ALLEN; VAN DER VELDEN, 2001) e a apresentar maior rotatividade de empregos (SICHERMAN, 1991).

De acordo com Reis (2021), a participação das mulheres no total de trabalhadores com nível superior completo aumentou de 41% em 1980 para 57% em 2010. Além disso, o autor destaca que, entre 1980 e 2000, a taxa de sobre-educação era maior para as mulheres. No entanto, essa dinâmica se inverteu em 2010, quando os homens se tornaram relativamente mais sobre-educados.

Os resultados encontrados pela literatura existente que combina a análise do *mismatch* educacional com questões gênero parecem depender bastante das técnicas utilizadas. Por isso, muitos estudos apontam para a importância do controle da heterogeneidade não observada dos indivíduos nas análises. Existem pesquisas que apontam evidências de que o *mismatch* impacta tanto homens quanto mulheres, mas que estas são penalizadas de maneira mais significativa (SALINAS-JIMÉNEZ, RAHONA-LÓPEZ; MURILLO-HUERTAS, 2013). Outras pesquisas indicam que o *mismatch* educacional horizontal afeta exclusivamente as mulheres (MAHUTEAU *et al.*, 2015), e também existem resultados que mostram que o *mismatch* não afeta nem mulheres nem homens (BOTO-GARCÍA; ESCALONILLA, 2022).

Diante desse breve contexto, apesar do *mismatch* educacional ser um problema atual no Brasil, caracterizado por ser um país em desenvolvimento e apresentar uma tendência crescente do número de indivíduos, especialmente as mulheres, com ensino superior completo, poucos ainda são os estudos que objetivam analisá-lo com foco em



questões de gênero. Nesse sentido, o presente artigo busca auxiliar no preenchimento dessa lacuna e se debruça na investigação da hipótese que as mulheres são mais penalizadas em termos salariais no mercado de trabalho formal por estarem em situação de *mismatch* educacional quando comparadas aos homens.

Para tal, os objetivos dividem-se em duas partes. Primeiramente, pretende-se investigar a proporção de indivíduos egressos da Universidade Federal do Paraná (UFPR) que se encontram em situações de *mismatch* educacional vertical e horizontal. Em um segundo momento, a partir de um painel de dados que permite controlar as características individuais não observáveis e invariantes no tempo, busca-se analisar se os efeitos decorrentes do *mismatch* educacional vertical e horizontal exercem impactos diferentes sobre os salários de mulheres e homens no mercado de trabalho formal.

As investigações do presente artigo se diferenciam das demais existentes na literatura uma vez que realiza uma análise dinâmica a partir de uma base de dados que integra informações da Relação Anual de Informações Sociais (RAIS) identificada e da Classificação Brasileira de Ocupações (CBO) com informações de egressos da UFPR, possibilitando a identificação das áreas de formação desses indivíduos e suas posteriores ocupações e remunerações no mercado de trabalho formal ao longo do tempo. O período de análise vai do ano de 2003 ao ano de 2021.

O artigo está estruturado em outras quatro seções, além desta introdução. A seção dois traz a revisão de literatura. A seção três apresenta a base de dados e a metodologia utilizada na condução das análises. Os resultados e discussões estão expostos na seção quatro. Por fim, a seção cinco apresenta as considerações finais do artigo.

2. Revisão de literatura

A revisão de literatura a seguir sintetiza os principais trabalhos encontrados sobre o tema. A partir dela, pode-se ressaltar dois aspectos principais. Em primeiro lugar, os resultados sugerem, na grande maioria, efeitos negativos do *mismatch* educacional no salário dos indivíduos. Em segundo lugar, contudo, os estudos apontam para a importância do controle da heterogeneidade não observada dos indivíduos e para a necessidade de



cautela na interpretação dos resultados de trabalhos que não tiveram a possibilidade de fazê-lo.

O trabalho de Bauer (2002) aplica o modelo OLS *pooled* a um painel de dados de indivíduos da Alemanha entre os anos de 1984 e 1998. Os resultados encontrados apontam que os trabalhadores sobre-educados ganham menos, e trabalhadores subeducados ganham mais do que trabalhadores com o mesmo nível de formação educacional, mas que trabalham em ocupações que requerem plenamente seus níveis de educação. Contudo, quando a heterogeneidade não observada dos indivíduos é controlada, os resultados se alteram e, na maioria dos casos, os efeitos do *mismatch* educacional desaparecem completamente. Dessa forma, o autor acende alerta para a importância do controle da heterogeneidade não observada nas análises.

Um outro exemplo de estudo que atenda para essa questão é Tsai (2010). A partir de um painel de dados de 1979 a 2005 para indivíduos dos Estados Unidos, o autor não encontra resultados significativos que apontem para efeitos negativos do *mismatch* educacional sobre os salários quando a heterogeneidade não observada dos indivíduos é levada em consideração.

Por sua vez, Korpi e Tåhlin (2009), analisam o *mismatch* educacional a partir de modelos estáticos e dinâmicos em painel com controle de efeitos fixos aplicados a dados de indivíduos da Suécia entre os anos de 1974 e 2000. Os resultados encontrados pelos autores sugerem que os trabalhadores sobre-educados são penalizados desde cedo por uma taxa inferior de retorno ao ensino do qual não se recuperam.

Já os trabalhos de Allen e van der Velden (2001) e Verhofstadt, De Witte e Omeij (2007) analisam a questão do *mismatch* educacional com a satisfação profissional, verificada a partir de um questionário respondido pelos indivíduos analisados. Allen e van der Velden (2001) exploram a relação entre *mismatch* educacional e o *mismatch* de habilidades, caracterizado pela incompatibilidade entre as habilidades adquiridas na formação superior e as requeridas no mercado de trabalho. A partir de um modelo *crosssection* aplicado a uma base de dados de onze países da Europa e do Japão para o ano de 1998, os autores encontram que o *mismatch* educacional afeta significativamente o salário dos indivíduos e apontam para a importância da distinção entre escolaridade e



habilidade, uma vez que essa se relaciona fortemente com a satisfação e a procura por trabalhos.

Por sua vez, Verhofstadt, De Witte e Omeij (2007), a partir de um modelo Logit aplicado a dados de indivíduos da Bélgica em seus primeiros empregos nos anos de 1999 e 2000, encontra que os trabalhadores mais educados estão mais satisfeitos do que seus colegas menos educados, porque possuem empregos de melhor qualidade. Quando se controla por todas as características do trabalho, o estudo encontra uma relação negativa, com os trabalhadores mais educados relatando menor satisfação no trabalho e sugere que tal resultado pode ser devido ao fato de que trabalhadores com mais educação têm expectativas mais altas em relação aos empregos, particularmente quando se trata do primeiro.

Os estudos de Salinas-Jiménez, Rahona-López e Murillo-Huertas (2013), Boto-García e Escalonilla (2022) e Mahuteau *et al.* (2015) são exemplos que combinam a análise do *mismatch* educacional com questões de gênero.

Salinas-Jiménez, Rahona-López e Murillo-Huertas (2013) analisam os efeitos do *mismatch* educacional sobre os salários de mulheres e homens na Espanha e se o *gap* salarial entre os gêneros é explicado por diferenças nas características produtivas ou por retornos diferentes associados a tais características. A partir da abordagem proposta por Mincer (1974), do modelo ORU (*Over, Required and Under Education*) de Duncan e Hoffman (1981) e da decomposição de Blinder-Oaxaca (Blinder, 1973; Oaxaca 1973), as autoras concluem que o *mismatch* educacional contribui para a diminuição dos salários de ambos os gêneros, mas que as mulheres sofrem penalidades mais significativas. Além disso, ao considerar o nível educacional, o estudo revela que a diferença salarial entre os gêneros é menor para indivíduos com baixos níveis de educação, e a porção desse *gap* relacionada às diferenças nos retornos é mais proeminente nesse grupo. Em contrapartida, a disparidade salarial de gênero é mais acentuada entre trabalhadores altamente educados; contudo, nesse cenário, a maior parte do *gap* salarial é atribuído às diferenças nas características produtivas.

Na mesma linha, Boto-García e Escalonilla (2022) investigam a questão do *mismatch* separadamente para mulheres e homens. A partir de um modelo Probit aplicado a uma base de dados de graduados na Espanha nos anos de 2009 e 2010, os autores não



encontramos evidências de que homens e mulheres apresentam riscos diferentes estarem em *mismatch* educacional condicional às suas características observáveis.

Por sua vez, Mahuteau *et al.* (2015), examinam os impactos dos *mismatches* vertical e horizontal nos salários na Austrália, com atenção especial às disparidades de gênero. Utilizando um painel de dados com efeitos fixos, os autores concluem que o *mismatch* vertical resulta em redução salarial tanto para homens quanto para mulheres. Enquanto isso, o *mismatch* horizontal não afeta os salários por hora dos homens, mas causa uma redução significativa para as mulheres. No entanto, são as pessoas que experimentam *mismatches* vertical e horizontal simultaneamente que enfrentam as penalidades mais severas, especialmente as mulheres.

Em relação ao Brasil, o trabalho de Marioni (2020) analisa o *mismatch* educacional para o mercado de trabalho brasileiro a partir de uma variação do modelo proposto por Verdugo e Verdugo (1989) aplicado a um painel de dados com efeitos fixos e informações da RAIS. A autora encontra proporções semelhantes em relação ao número de indivíduos em situação de *mismatch* educacional vertical – subqualificados e superqualificados. Além disso, o estudo aponta que os trabalhadores superqualificados (subqualificados) ganham significativamente menos (mais) em comparação àqueles que se não estão em *mismatch* e a que a penalidade para a superqualificação é equivalente ao prêmio pela subqualificação.

A partir da utilização dos dados do Censo e da CBO, Reis (2018) fornece uma medida contínua para representar a distância entre as habilidades adquiridas no ensino superior e aquelas exigidas na ocupação de um indivíduo. Essa medida é baseada na semelhança entre as atividades geralmente realizadas na ocupação de cada indivíduo e aquelas necessárias na ocupação considerada a mais intimamente relacionada ao seu campo de estudo. A partir da sua utilização, o autor encontra que indivíduos que têm ocupações com habilidades semelhantes às adquiridas em suas áreas de estudo ganham mais do que aqueles cujas ocupações e área de formação não se relacionam. Por fim, o autor acende alerta para a importância de considerar descrições mais precisas de trabalhadores e ocupações para caracterizar melhor o *mismatch* entre a ocupação e a área de formação e seus impactos nos salários.



O estudo de Reis e Fortes (2019) analisa os efeitos do *mismatch* educacional vertical especificamente para duas modalidades de educação profissional no Brasil: qualificação profissional e curso técnico de nível médio. Os autores encontram que, entre os indivíduos que estão em ocupações que não demandam a conclusão de um curso de educação profissional, aqueles que a concluíram ganham salários mais altos do que aqueles que não a concluíram. Porém, entre os indivíduos que completaram um curso de educação profissional, aqueles que estão em ocupações que exigem essa qualificação ganham mais do que aqueles que estão em ocupações que não exigem. Por fim, entre os indivíduos que estão em ocupações que exigem a conclusão de um curso de educação profissional, aqueles que de fato a possuem ganham mais do que aqueles que não a possuem. Os resultados são observados tanto para cursos de qualificação profissional quanto para cursos técnicos de nível médio.

Por sua vez, Reis (2020) investiga acerca da probabilidade de estar sobre-educado no Brasil. De acordo com os resultados, entre os indivíduos com o ensino médio completo, aqueles egressos de escolas da rede pública têm maior probabilidade de sobre-educação em comparação com os de escolas privadas. Para os que completaram um curso superior, os resultados mostram o contrário, ou seja, os egressos da rede pública possuem uma probabilidade menor de serem sobre-educados. O autor sugere que instituições que apresentam a qualidade média pior na educação estão associadas a uma probabilidade relativamente maior de formarem indivíduos sobre-educados.

A partir de informações de quatro edições do Censo Demográfico, Reis (2021) analisa o comportamento do *mismatch* educacional vertical entre trabalhadores com ensino superior no Brasil entre 1980 e 2010. Conforme os resultados encontrados pelo autor, a proporção de indivíduos com educação superior classificados como sobre-educados apresentou um aumento significativo entre 1980 e 2000, embora já fosse elevada no período inicial da análise. Em 1980, aproximadamente 25% dos trabalhadores estavam em situação de *mismatch* vertical, aumentando em cerca de 10 pontos percentuais vinte anos depois.

Carvalho e Reis (2023) procuram investigar acerca da sobre-educação no mercado de trabalho brasileiro em períodos mais recentes. A partir de dados da PNAD Contínua dos anos entre 2012 e 2022, os autores encontram como resultados iniciais o crescimento da



proporção de trabalhadores sobre-educados, sendo este mais intenso entre os trabalhadores com ensino médio completo. Além disso, os autores apontam que o requerimento de escolaridade das ocupações pouco aumentou ao longo dos últimos dez anos, expondo, dessa forma, que o crescimento da demanda por trabalho qualificado cresceu em ritmo bem inferior à oferta.

Em suma, os estudos apontam para os efeitos negativos do *mismatch* educacional nos resultados do mercado de trabalho. Nesse sentido, o presente artigo busca contribuir para a literatura nacional a partir da utilização de uma base de dados única com ênfase nas diferenças entre os gêneros.

3. Metodologia e tratamento de dados

3.1. Base de dados e recorte amostral

A condução das análises do presente artigo parte da integração de dados da Relação Anual de Informações Sociais Identificada (RAIS Identificada) disponibilizados pelo Ministério da Fazenda, da Classificação Brasileira de Ocupações (CBO) e de dados administrativos da Universidade Federal do Paraná (UFPR) concedidos em parceria com a Agência de Tecnologia da Informação e Comunicação (AGTIC). A integração entre as bases de dados foi possível por meio do CPF de cada discente egresso:

A partir da combinação entre essas bases foi possível criar uma amostra que permite identificar a área de formação na UFPR de um indivíduo, suas posteriores ocupações e remunerações no mercado de trabalho formal e a escolaridade máxima requerida para desempenhar as funções nessas ocupações. A amostra conta com 14.853 observações (7.879 mulheres e 6.974 homens) que iniciaram e concluíram um curso de ensino superior na UFPR entre os anos de 1986 e 2020.

Devido à falta de algumas informações da base de dados da UFPR, em alguns casos os anos mínimos e máximos foram calculados a partir da data da primeira e da última matéria cursada, respectivamente, apenas para os indivíduos que apresentaram “Formatura” como forma de evasão. Ainda, foram considerados apenas níveis de graduação (15 semanas, 18 semanas, 20 semanais e cursos anuais). Para as pessoas que



se formaram em mais de um curso nesse período foi considerado apenas o primeiro. Além disso, foram considerados apenas os indivíduos que possuem ensino superior completo como escolaridade máxima, excluindo aqueles com mestrado e doutorado. Por fim, assim como no segundo artigo, os cursos foram classificados em áreas do conhecimento conforme a CAPES.

Os indivíduos presentes na base de dados da UFPR foram posteriormente encontrados na base de dados da RAIS identificada, de modo que fosse possível obter informações acerca de suas remunerações e suas ocupações no mercado de trabalho formal. Para fins da construção do painel de dados, foram considerados apenas o primeiro emprego daqueles indivíduos que trocaram de emprego no mesmo ano ou que estavam empregados em mais de um emprego no mesmo ano. Além disso, foram considerados apenas empregos imediatamente posteriores ao ano de formação na UFPR. O período considerado vai do ano de 2003 ao ano de 2021. Por fim, os dados da RAIS foram combinados com as informações sobre a escolaridade requerida de cada ocupação, conforme a Classificação Brasileira de Ocupações (CBO)¹. O presente artigo considerou como constante a necessidade educacional de cada ocupação ao longo do período analisado, tal qual Carvalho e Reis (2023).

Uma vez que a análise prévia dos dados apontou para a presença de *outliers* no que diz respeito às remunerações, foi utilizada a técnica de Z-scores para identificação e exclusão dos indivíduos que recebiam salários muito distantes da média geral. Para estimar os Z-scores, foi subtraído da média o valor de cada observação e o resultado foi dividido pelo desvio-padrão. A distribuição resultante apresentou média zero e a distância entre as observações passou a ser em unidades de desvio-padrão. Quanto maior for o valor do Z-score, em valores absolutos, maior é a distância entre uma determinada observação e a média geral (FIGUEIREDO FILHO *et al.*, 2023). Em conformidade com a literatura existente, o presente artigo definiu com *outliers* os casos que apresentaram valores superiores a 3 e inferiores a -3 unidades de desvio-padrão (ATKINSON; MULIRA, 1993; LEWIS; BARNETT, 1978; WALFISH, 2006).

¹ A escolaridade requerida segue o trabalho de Reis (2020), cujos dados detalhados foram gentilmente fornecidos pelos autores.



3.2. Descrição das variáveis e das hipóteses

O Quadro 1 a seguir traz as informações das variáveis que foram utilizadas para investigar a hipótese de que as mulheres em situação de *mismatch* educacional são mais penalizadas em termos de remuneração no mercado de trabalho formal do que os homens na mesma situação.

Quadro 1 – Descrições das variáveis

Variável	Descrição
Gênero	Variável <i>dummy</i> igual a 1 se o indivíduo for mulher e igual a 0 se for homem.
<i>Mismatch</i> vertical	Variável <i>dummy</i> igual a 1 se o indivíduo está em <i>mismatch</i> vertical e igual a zero caso contrário.
<i>Mismatch</i> horizontal	Variável <i>dummy</i> igual a 1 se o indivíduo está em <i>mismatch</i> horizontal e igual a zero caso contrário.
<i>Mismatch</i> Vertical x <i>Mismatch</i> Horizontal	Variável <i>dummy</i> igual a 1 se o indivíduo está em <i>mismatch</i> vertical e horizontal ao mesmo tempo e igual a zero caso contrário.
Regiões	<i>Dummies</i> de regiões do Brasil (Sul, Sudeste, Centro-Oeste, Nordeste e Norte). Região Sul como referência.
Horas contratadas	Variável contínua de horas contratadas para cada emprego.
Tempo empregado	Variável contínua para o tempo empregado em cada emprego.

Fonte: elaboração própria (2024).

3.3. *Mismatches* vertical e horizontal

Os *mismatches* educacionais podem ser classificados em vertical e horizontal. O *mismatch* vertical acontece quando um indivíduo está trabalhando em uma ocupação que requer mais (undereducation – subeducação) ou menos (overeducation – sobre-educação) anos de estudo do que efetivamente possui. O *mismatch* horizontal, por sua vez, ocorre quando o indivíduo está empregado em uma ocupação não relacionada com a sua respectiva área de formação acadêmica (PARK, 2021).

Para a classificação do *mismatch* vertical foram utilizadas as informações referentes ao nível de escolaridade considerado necessário para desempenhar as atividades de cada



ocupação, conforme a Classificação Brasileira de Ocupações (CBO), ao nível de quatro dígitos. Uma vez que todos os indivíduos da base de dados utilizada nesse artigo possuem nível superior completo e o máximo de escolaridade que a CBO utiliza nas classificações também se refere ao nível superior completo (15 anos de educação formal), apenas os sobre-educados foram analisados, ou seja, todos aqueles que possuem um número de anos de estudo maior do que o máximo considerado necessário para a sua ocupação, conforme definido em Carvalho e Reis (2023).

A estratégia de classificação do *mismatch* horizontal, por sua vez, foi baseada no cálculo da moda das ocupações em cada área conforme Mahuteau et al. (2015). Em outras palavras, foi identificada para cada área a ocupação que concentra a maior parte dos indivíduos. Assim, foram considerados como horizontalmente em *mismatch* todos aqueles empregados em uma ocupação diferente daquela que concentra a maior parte dos indivíduos de sua respectiva área (Agrárias, Biológicas; Engenharias; Exatas e Terra; Humanas; Linguística, Letras e Artes; Saúde; Sociais Aplicadas).

3.4. Estratégia de identificação e modelo empírico

A especificação utilizada no presente artigo é uma variação do modelo proposto por Verdugo e Verdugo (1989), com a diferença que, ao invés de definir os trabalhadores como sobre-educados a partir do valor do desvio-padrão acima da média do nível de escolaridade dos indivíduos de cada ocupação, utiliza a comparação entre o nível de escolaridade dos indivíduos e aquele requerido como o máximo para trabalhar em determinada ocupação, conforme a CBO. Dessa forma, para estimar o efeito dos *mismatches* vertical e horizontal sobre os salários, são adicionadas dummies no modelo pooled OLS conforme a Equação 1 a seguir:

$$\ln W_{it} = \sum_{k=1}^n \beta_1 MV_{it} + \beta_2 MH_{it} + \beta_3 MVH_{it} + \beta_4 \Gamma_{it} + \varepsilon_{it} \quad (1)$$

em que $\ln W_{it}$: logaritmo natural dos salários em reais do indivíduo i no tempo t ; MV_{it} : variável dummy igual a 1 se o indivíduo i está em *mismatch* vertical no tempo t ; MH_{it} : variável dummy igual a 1 se o indivíduo i está em *mismatch* horizontal no tempo t ; MVH_{it} :



variável dummy igual a 1 se o indivíduo está em *mismatch* vertical e horizontal no tempo t , Γ_{it} : vetor de variáveis observáveis; ε_{it} : termo de erro.

Conforme Ashenfelter, Harmon e Oosterbeek (1999), a escolaridade dos indivíduos pode estar correlacionada com o termo de erro causando o problema da endogeneidade e, conseqüentemente, estimações inconsistentes. Dessa forma, para mitigar o problema da heterogeneidade não observada entre os indivíduos, o presente artigo utiliza o painel de dados com efeitos fixos conforme a Equação 2 a seguir.

$$\ln W_{it} = \sum_{k=1}^n \beta_1 MV_{it} + \beta_2 MH_{it} + \beta_3 MVH_{it} + \beta_4 \Gamma_{it} + c_i + \varepsilon_{it} \quad (2)$$

em que c_i : termo de características não observáveis.

A principal vantagem das estimações utilizando efeitos fixos é a possibilidade de eliminação do efeito das características individuais que não variam com o tempo. As características não observáveis do indivíduo podem afetar tanto os salários quanto o *mismatch*. Entre essas características estão as habilidades não observáveis do indivíduo que, se consideradas fixas ao longo do tempo, não influenciam o resultado obtido nos salários e/ou no *mismatch*. Após as estimações dos modelos pooled, de efeitos aleatórios e de efeitos fixos, foi realizado o Teste de Hausman que indicou que a heterogeneidade não observada possui um comportamento fixo em todos os casos, fazendo com que os resultados obtidos a partir do modelo com efeitos fixos sejam preferíveis aos obtidos a partir do modelo com efeitos aleatórios.

4. Resultados e discussões

A seção de resultados e discussão a seguir traz as estatísticas descritivas, os resultados das estimações do painel com efeitos fixos e a análise de heterogeneidade.

4.1. Estatísticas descritivas

A Tabela 1 contém as estatísticas descritivas e os testes t das amostras separadas entre mulheres e homens. Pode-se observar que as mulheres ganham aproximadamente



80% do salário dos homens. Além disso, a maior parte dos indivíduos estão concentrados em ocupações na Região Sul e apresentam valores semelhantes em relação à média de idade (35,5 anos para as mulheres e 35,8 anos para os homens); à média de horas contratadas no mercado de trabalho (34,7 horas para as mulheres e 36 horas para os homens) e à média de tempo empregado(a) (58 meses para as mulheres e 59 meses para os homens).

Tabela 1 - Estatísticas descritivas e teste t

Variáveis	Mulheres		Homens		Teste t
	Média	Desvio-Padrão	Média	Desvio-Padrão	
Salário	6.522	7.151	8.168	5.945	41,202***
<i>Mismatch</i> vertical	0,3455	0,4795	0,3582	0,4755	4,414***
<i>Mismatch</i> horizontal	0,8193	0,3795	0,8256	0,3848	2,721***
<i>Mismatch</i> vertical e horizontal	0,2754	0,4467	0,2919	0,4546	6,0826***
Idade	35,4785	8,4871	35,8237	8,6602	6,672***
Sul	0,9180	0,2880	0,9087	0,2744	-5,455***
Sudeste	0,0596	0,2364	0,0594	0,2367	-0.114
Centro-Oeste	0,0160	0,1499	0,0230	0,1256	8,282***
Nordeste	0,0039	0,0660	0,0044	0,0627	1,0980
Norte	0,0024	0,0670	0,0045	0,0494	5,740***
Horas contratadas	34,6896	10,3302	35,9537	10,6977	19,933***
Tempo empregado(a)	58,1967	67,7217	59,3674	66,6510	2,884***
Agrárias	0,0485	0,2565	0,0708	0,2147	15,562***
Biológicas	0,0529	0,1732	0,0309	0,2238	-18,326***
Engenharias	0,0634	0,3956	0,1943	0,2436	65,0340***
Exatas e Terra	0,1013	0,3895	0,1865	0,3017	40,1830***
Humanas	0,1821	0,2888	0,0918	0,3859	-44,277***
Linguística, Letras e Artes	0,0700	0,1922	0,0384	0,2552	-23,378***
Saúde	0,2576	0,3406	0,1339	0,4373	-52,700***
Sociais Aplicadas	0,2243	0,4349	0,2533	0,4171	11,247***
Total de observações		7879		6974	

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

A área da Saúde é a que concentra a maior parte das mulheres (26%), seguida pelas áreas de Sociais Aplicadas (23%), Humanas (18%), Exatas e Terra (10%), Linguística, Letras e Artes (7%), Engenharias (6%), Biológicas (5%), e Agrárias (5%). Os homens, por



sua vez, estão mais concentrados na área de Sociais Aplicadas (25%), seguida pelas áreas de Engenharias (19%), Exatas e Terra (19%), Saúde (13%), Humanas (9%), Agrárias (7%), Linguística, Letras e Artes (4%) e Biológicas (4%). Dessa forma, pode-se perceber que as áreas com as maiores diferenças entre a concentração de mulheres e homens são a Saúde e as Humanas em favor das mulheres e as áreas das Engenharias e Exatas e Terra em favor dos homens. Vale ressaltar que esses valores apontam para a sub-representação de mulheres nas áreas STEM (ciência, tecnologia, engenharia e matemática), em conformidade com a literatura existente (BARONE E ASSIRELLI, 2018; SAHOO E KLASSEN, 2018; GONZÁLEZ-PÉREZ, CABO, SÁINZ, 2020; BUCKLEY, FARRELL, TYNDALL, 2021).

Em relação aos *mismatches*, pode-se notar que a proporção de indivíduos com *mismatch* vertical e horizontal é semelhante entre mulheres e homens: aproximadamente 35% das mulheres e 36% dos homens estão em *mismatch* vertical e 82% das mulheres e 83% dos homens em *mismatch* horizontal. O mesmo acontece para os indivíduos que estão tanto em *mismatch* vertical quanto em horizontal: 27% das mulheres e 30% dos homens. É possível notar que a tendência de os homens se tornarem mais sobre-educados do que as mulheres a partir de 2010, conforme exposto em Reis (2021), também se reflete nos dados analisados neste artigo.

4.2. Resultados das estimações do painel de dados

A Tabela 2 a seguir apresenta os resultados das estimações do painel de dados. Primeiramente, foram estimados os coeficientes considerando todos os indivíduos. Depois, as estimações foram feitas separadamente para mulheres e homens. As colunas 1 consideram apenas a variável de *mismatch* vertical, as colunas 2 consideram apenas o horizontal, e as colunas 3 consideram as variáveis dos dois tipos de *mismatch* concomitantemente com uma variável de interação entre os *mismatches*. Todas as especificações comparam os indivíduos em *mismatch* vertical (sobre-educados) e/ou *mismatch* horizontal àqueles que possuem a mesma quantidade de anos de estudo formal, mas que estão alocados em ocupações mais adequadas aos seus níveis de escolaridade



(vertical) e às suas áreas de formação (horizontal). O período das análises vai do ano 2009 ao ano de 2021.

Uma vez que o Teste de Hausman apontou para os modelos de efeitos fixos como os mais adequados em todos os casos, apenas estes estão reportados na Tabela 36 a seguir. Os resultados das estimações com os modelos OLS *pooled* e de efeitos variáveis podem ser encontrados nas tabelas em anexo.

4.2. Análise de heterogeneidade

De acordo com Mahuteau *et al.* (2015), é possível que o *mismatch* educacional horizontal seja mais provável entre os indivíduos que concluíram uma formação superior em uma área que fornece habilidades mais gerais e menos provável entre aqueles de áreas que oferecem habilidades mais específicas para uma ocupação. Isso porque, diferentemente do *mismatch* vertical, o *mismatch* horizontal é mais difícil preservar parte do capital humano específico que os indivíduos obtêm de determinadas áreas. No entanto, os efeitos salariais podem ser maiores no último caso para aqueles que não conseguem encontrar uma correspondência adequada. Sendo assim, na tentativa de investigar acerca das heterogeneidades, os coeficientes foram estimados separadamente para cada uma das oito áreas, conforme Tabela 3.

**Tabela 2 - Resultados das estimações do painel de dados com efeitos fixos**

Variáveis	Todos			Mulheres			Homens		
	1	2	3	1	2	3	1	2	3
<i>Mismatch</i> vertical	-0.142*** (0.006)		-0.263*** (0.013)	-0.155*** (0.008)		-0.277*** (0.018)	-0.127*** (0.009)		-0.250*** (0.020)
<i>Mismatch</i> horizontal		-0.029*** (0.007)	-0.105*** (0.009)		-0.029*** (0.009)	-0.110*** (0.012)		-0.030*** (0.010)	-0.101*** (0.013)
<i>Mismatch</i> vertical x <i>Mismatch</i> horizontal			0.134*** (0.014)			0.135*** (0.019)			0.136*** (0.021)
Controles	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
Observações	108.799	108.799	108.799	57.980	57.980	57.980	50.819	50.819	50.819
R ²	0.101	0.096	0.102	0.110	0.104	0.112	0.091	0.087	0.093
R ² Ajustado	-0.041	-0.047	-0.039	-0.029	-0.037	-0.028	-0.053	-0.058	-0.051

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Notas: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.
Estatística F signficante a 1% em todos os casos

**Tabela 3 - Resultados das estimações por gênero e área do conhecimento**

Áreas	Casos	Mismatch vertical	Mismatch horizontal	Mismatch vertical e horizontal
Agrárias	Todos	-0.233*** (0.027)	-0.008 (0.025)	-0.080 (0.057)
	Mulheres	-0.247*** (0.041)	0.007 (0.035)	-0.172*** (0.036)
	Homens	-0.221*** (0.036)	-0.022 (0.035)	-0.186*** (0.033)
Biológicas	Todos	-0.019 (0.031)	0.024 (0.030)	0.025 (0.029)
	Mulheres	-0.025 (0.037)	0.007 (0.039)	0.008 (0.036)
	Homens	0.039 (0.059)	0.045 (0.050)	0.083* (0.049)
Engenharias	Todos	-0.148*** (0.014)	-0.171*** (0.019)	-0.131*** (0.014)
	Mulheres	-0.184*** (0.029)	-0.196*** (0.037)	-0.175*** (0.028)
	Homens	-0.134*** (0.016)	-0.157*** (0.022)	-0.113*** (0.016)
Exatas e Terra	Todos	-0.124*** (0.015)	0.027 (0.018)	-0.124*** (0.015)
	Mulheres	-0.114*** (0.025)	0.011 (0.028)	-0.114*** (0.025)
	Homens	-0.134*** (0.019)	0.041* (0.023)	-0.134*** (0.019)
Humanas	Todos	-0.161*** (0.017)	-0.090*** (0.018)	-0.144*** (0.017)
	Mulheres	-0.215*** (0.020)	-0.121*** (0.020)	-0.184*** (0.020)
	Homens	-0.036 (0.032)	0.031 (0.041)	-0.052* (0.032)
Linguística, Letras e Artes	Todos	-0.126*** (0.029)	-0.079*** (0.023)	-0.082*** (0.025)
	Mulheres	-0.161*** (0.034)	-0.037 (0.028)	-0.104*** (0.031)
	Homens	-0.036 (0.052)	-0.156*** (0.040)	-0.037 (0.042)
Saúde	Todos	-0.166*** (0.016)	-0.353*** (0.022)	-0.166*** (0.016)
	Mulheres	-0.209*** (0.018)	-0.325*** (0.032)	-0.209*** (0.018)
	Homens	-0.054 (0.033)	-0.390*** (0.032)	-0.054 (0.033)
Sociais Aplicadas	Todos	-0.112*** (0.011)	0.121*** (0.012)	-0.024** (0.011)
	Mulheres	-0.086*** (0.014)	0.106*** (0.016)	-0.003 (0.015)
	Homens	-0.143*** (0.016)	0.140*** (0.019)	-0.047*** (0.017)

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Notas: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.



Em relação aos efeitos nos salários, pode-se observar que, para todos os indivíduos, as áreas caracterizadas por habilidades mais específicas, como a Saúde e Engenharias, apresentam efeitos negativos significativamente mais acentuados no que diz respeito ao *mismatch* horizontal em comparação às demais áreas de conhecimentos menos específicos, como Humanas, Linguística; Letras e Artes e Sociais Aplicadas, por exemplo, em conformidade com Mahuteau *et al.* (2015). O mesmo padrão, contudo, não aparece quando se analisam separadamente cada gênero. Enquanto as mulheres são mais penalizadas do que os homens por estarem em *mismatch* horizontal nas áreas de Engenharias e Humanas, os homens são mais penalizados nas áreas de Linguística Letras e Artes, Saúde e Sociais Aplicadas.

Em relação àqueles que estão em situação de *mismatch* vertical e *mismatch* horizontal, as mulheres são mais penalizadas na maior parte dos casos: Agrárias, Humanas, Engenharias, Linguística, Letras e Artes e Saúde. Os homens, por sua vez, são mais penalizados nas áreas de Exatas e Terra e Sociais Aplicadas.

A principal vantagem das estimações utilizando efeitos fixos é a possibilidade de eliminação do efeito das características individuais que não variam com o tempo. Dessa forma, os coeficientes podem ser interpretados como o efeito médio nos salários devido a mudanças no status de *mismatch* vertical e/ou horizontal dos indivíduos ao longo do tempo (MARIONI, 2020).

A partir dos resultados reportados na Tabela 36, pode-se notar que estar em *mismatch* vertical e horizontal afeta negativamente os salários em todos os casos, em consonância com os resultados encontrados por Bauer (2002), Korpi e Tåhlin (2009), Marioni (2020), Mahuteau *et al.* (2015). Considerando todos os indivíduos, as colunas 1 e 2 mostram que o *mismatch* vertical diminui o salário em aproximadamente 14% enquanto o efeito do horizontal é uma diminuição de cerca de 3%. Já os indivíduos que estão tanto em *mismatch* vertical quando em horizontal têm seus salários afetados negativamente em cerca de 11%. Vale ressaltar que, conforme Bauer (2002), os resultados estimados por efeitos fixos são significativamente menores daqueles estimados pelo modelo *pooled* (Tabelas no apêndice), sustentando a importância do controle da heterogeneidade não observada dos indivíduos e a cautela ao se interpretar os resultados de estimações que não tiveram a possibilidade de fazê-lo.

Analisando separadamente para mulheres e homens, é possível observar que os resultados são semelhantes, tanto na direção do sinal quanto em termos de magnitude: enquanto o *mismatch* vertical afeta negativamente o salário dos homens em cerca de 13% e o horizontal 3%, o efeito é ligeiramente pior para as mulheres: aproximadamente 15% e 3%, respectivamente, em conformidade com os resultados encontrados por Mahuteau *et al.* (2015). As mulheres também são ligeiramente mais penalizadas do que os homens por estarem em ambos os *mismatches*: enquanto seus salários são impactados negativamente em aproximadamente 12%, o valor para os homens é de cerca de 9%.

5. Considerações finais

Os últimos anos foram marcados por um aumento da média de anos de estudo e da participação de indivíduos com ensino superior na força de trabalho no Brasil. Entretanto, quando a demanda por mão de obra qualificada não acompanha o aumento da sua oferta, surge o problema do *mismatch* educacional, ou seja, os indivíduos estão desempenhando funções em ocupações que não requerem o nível de escolaridade que possuem (*mismatch* vertical) ou, então, estão em ocupações não relacionadas com as suas formações acadêmicas (*mismatch* horizontal).

Apesar do *mismatch* educacional ser um problema atual no Brasil, caracterizado por ser um país em desenvolvimento e apresentar uma tendência crescente do número de indivíduos, especialmente as mulheres, com ensino superior completo, poucos ainda são os estudos que objetivam analisá-lo com foco em questões de gênero. Nesse sentido, o presente artigo buscou auxiliar no preenchimento dessa lacuna e se debruçou na investigação da hipótese que as mulheres são mais penalizadas em termos salariais no mercado de trabalho formal por estarem em situação de *mismatch* educacional quando comparadas aos homens.

A partir de um painel de dados que integrou informações de egressos da Universidade Federal do Paraná, como área de formação, com suas informações sobre ocupação e remuneração no mercado de trabalho da base de dados da RAIS identificada,

os resultados encontrados apontam que os *mismatches* vertical e horizontal afetam negativamente o salário dos indivíduos.

Quando os gêneros são analisados separadamente, observa-se que os resultados são semelhantes, tanto na direção do sinal quanto em termos de magnitude, com as mulheres sendo ligeiramente mais penalizadas em comparação aos homens, tanto no *mismatch* vertical quanto no horizontal, corroborando os resultados encontrados por Mahuteau *et al.* (2015) e Salinas-Jiménez, Rahona-López, Murillo-Huertas, (2013).

Por fim, quando as análises são feitas separadamente para cada área, os resultados indicam que áreas caracterizadas por habilidades mais específicas, como a Saúde e Engenharias, apresentam efeitos negativos significativamente mais acentuados no que diz respeito ao *mismatch* horizontal em comparação às demais áreas de habilidades menos específicos, como Humanas, Linguística; Letras e Artes e Sociais Aplicadas. Porém, esse padrão não se repete quando as análises são feitas separadamente entre mulheres e homens. Essa é uma das contribuições centrais do presente artigo, posto que os resultados separados por áreas do conhecimento de indivíduos que foram acompanhados ao longo do tempo não foram previamente encontrados na literatura. Ademais, vale ressaltar que outras alternativas de análises, principalmente em relação ao cálculo *mismatch* horizontal, poderiam tornar os resultados mais robustos. Uma delas diz respeito à autoavaliação, na qual os próprios indivíduos determinam, a partir de uma análise subjetiva, se exercem funções no mercado de trabalho relacionadas as suas formações acadêmicas.

Referências

ALLEN, J.; VAN DER WELDEN, R. Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. **Oxford Economic Papers**, 53(3): 434–452, 2001.

ATKINSON, A.; MULIRA, H.-M. The stalactite plot for the detection of multivariate outliers. **Statistics and Computing**, 3(1): 27–35, 1993.

BARONE, C.; ASSIRELLI, G. Gender segregation in higher education: an empirical test of seven explanations. **Higher Education**, 79: 55–78, 2020.

BAUER, T. Educational mismatch and wages: a panel analysis. **Economics of Education Review**, 21: 221–229, 2002.



BOTO-GARCÍA, D.; ESCALONILLA, M. University education, mismatched jobs: are there gender differences in the drivers of overeducation? **Economia Política**, 39: 861–902, 2022.

BUCKLEY, C.; FARRELL, L; TYNDALL, I. *Brief stories of successful female role models in science help counter gender stereotypes regarding intellectual ability among young girls: A pilot study*. **Early Education and Development**, 33(4): 1–12, 2021.

CARVALHO, S.; REIS, M. Evolução da sobre-educação no mercado de trabalho no Brasil entre 2012 e 2022: primeiros resultados. **IPEA – Nota Técnica, Mercado de Trabalho 75**, 2023.

GONZÁLEZ-PÉREZ, S.; CABO, R.; SÁINZ, M. Girls in STEM: is it a female role-model thing? **Frontiers in Psychology**, 11: 1–21, 2020.

KORPI, T.; TÅHLIN, M. Educational mismatch, wages, and wage growth: overeducation in Sweden. **Labour Economics**, 16(2): 183–193, 2009.

LEWIS, T.; BARNETT, V. **Outliers in statistical data**. Chichester: John Wiley & Sons, 1978.

MAHUTEAU, S.; MAVROMARAS, K.; SLOANE, P.; WEI, Z. Horizontal and vertical educational mismatch and wages. **National Institute of Labor Studies, Flinders University – Working Paper**, 2015.

MARIONI, L. Overeducation in the labour market: evidence from Brazil. **Education Economics**, 29(1): 53–72, 2020.

PARK, K. Education-job mismatch and gender wage gap: evidence from recent college graduates in Korea. **Asian Women**, 37(1): 1–24, 2021.

REIS, M. Education mismatch and labor earnings in Brazil. **International Journal of Manpower**, 38(2): 180–197, 2017.

REIS, M. Measuring the mismatch between field of study and occupation using a task-based approach. **Journal for Labour Market Research**, 52(9): 1–15, 2018.

REIS, M. Os ensinos público e privado no Brasil e a incidência de sobre-educação no mercado de trabalho. **Economia Aplicada**, 24(3): 367–392, 2020.

REIS, M. Educação superior e sobre-educação no Brasil entre 1980 e 2010. **IPEA – Texto para Discussão 2655**, 2021.

REIS, M.; FORTES, M. Educação profissional, exigências da ocupação e rendimentos do trabalho no Brasil. **IPEA – Texto para Discussão 2446**, 2019.



SAHOO, S.; KLASSEN, S. Gender segregation in education and its implications for labour market outcomes: evidence from India. **IZA – Institute of Labor Economics**. Discussion paper nº 11660, 2018.

SALINAS-JIMÉNEZ, M.; RAHONA-LÓPEZ, M. MURILLO-HUERTAS, I. Gender wage differentials and educational mismatch: an application to the Spanish case. **Applied Economics**, 45(30): 4226–4235, 2013.

SICHERMAN, N. Overeducation in the labor market. **Journal of Labor Economics**, 9(2): 101–122, 1991.

TSAI, Y. Returns to overeducation: a longitudinal analysis of the US labor market. **Economics of Education Review**, 29(4): 606–617, 2010.

VERDUGO, R.; VERDUGO, N. The impact of surplus schooling on earnings: some additional findings. **Journal of Human Resources**, 24(4): 629–643.

VERHOFSTADT, E.; DE WITTE, H.; OMEY, E. Higher educated workers: better jobs but less satisfied? **International Journal of Manpower**, 28: 135–151, 2007.

WALFISH, S. A review of statistical outlier methods. **Pharmaceutical Technology**, 30(11): 82–88.



Apêndices

Tabela A1– Resultados das estimações com efeitos fixos para toda a base

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.142*** (0.006)			-0.263*** (0.013)
<i>Mismatch</i> horizontal		-0.029*** (0.007)		-0.105*** (0.009)
<i>Mismatch</i> vertical x horizontal			-0.106*** (0.006)	0.134*** (0.014)
Norte	0.169*** (0.058)	0.175*** (0.058)	0.176*** (0.058)	0.167*** (0.058)
Nordeste	0.243*** (0.051)	0.246*** (0.051)	0.241*** (0.051)	0.249*** (0.051)
Centro-Oeste	0.321*** (0.027)	0.310*** (0.028)	0.314*** (0.028)	0.318*** (0.027)
Sudeste	0.213*** (0.016)	0.212*** (0.016)	0.215*** (0.016)	0.214*** (0.016)
Horas contratadas	0.020*** (0.0002)	0.020*** (0.0002)	0.020*** (0.0002)	0.020*** (0.0002)
Tempo empregado(a)	0.002*** (0.0001)	0.002*** (0.0001)	0.002*** (0.0001)	0.002*** (0.0001)
Observações	108,799	108,799	108,799	108,799
R ²	0.101	0.096	0.099	0.102
R ² Ajustado	-0.041	-0.047	-0.043	-0.039
Estatística F	1,508.243***	1,419.930***	1,469.389***	1,190.431***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.



Tabela A2 – Resultados das estimações com efeitos aleatórios para toda a base

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.307*** (0.011)			-0.483*** (0.026)
<i>Mismatch</i> horizontal		-0.075*** (0.013)		-0.196*** (0.018)
<i>Mismatch</i> vertical x horizontal			-0.253*** (0.011)	0.192*** (0.027)
Norte	0.353*** (0.102)	0.372*** (0.105)	0.370*** (0.103)	0.357*** (0.102)
Nordeste	0.313*** (0.091)	0.328*** (0.093)	0.313*** (0.091)	0.323*** (0.090)
Centro-Oeste	0.517*** (0.048)	0.495*** (0.050)	0.507*** (0.048)	0.509*** (0.048)
Sudeste	0.295*** (0.027)	0.294*** (0.028)	0.300*** (0.027)	0.298*** (0.027)
Horas contratadas	0.021*** (0.0004)	0.019*** (0.0004)	0.020*** (0.0004)	0.021*** (0.0004)
Tempo empregado(a)	0.004*** (0.0001)	0.003*** (0.0001)	0.003*** (0.0001)	0.004*** (0.0001)
Constante	7.561*** (0.027)	7.556*** (0.042)	7.539*** (0.041)	7.718*** (0.031)
Observações	108,799	108,799	108,799	108,799
R ²	0.311	0.225	0.287	0.324
R ² Ajustado	0.311	0.224	0.287	0.324
Estatística F	4,330.469***	3,310.935***	4,007.058***	4,523.015***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.

Tabela A3 – Resultados das estimações OLS *pooled* para toda a base

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.578*** (0.005)			-0.759*** (0.011)
<i>Mismatch</i> horizontal		-0.169*** (0.006)		-0.298*** (0.007)
<i>Mismatch</i> vertical x horizontal			-0.541*** (0.005)	0.203*** (0.012)
Norte	0.453*** (0.038)	0.608*** (0.040)	0.495*** (0.038)	0.467*** (0.037)
Nordeste	0.364*** (0.035)	0.468*** (0.037)	0.386*** (0.035)	0.378*** (0.034)
Centro-Oeste	0.633*** (0.016)	0.682*** (0.017)	0.633*** (0.017)	0.620*** (0.016)
Sudeste	0.387*** (0.009)	0.448*** (0.010)	0.414*** (0.009)	0.398*** (0.009)
Horas contratadas	0.026*** (0.0002)	0.020*** (0.0002)	0.024*** (0.0002)	0.027*** (0.0002)
Tempo empregado(a)	0.005*** (0.00003)	0.005*** (0.00003)	0.005*** (0.00003)	0.005*** (0.00003)
Constante	7.515*** (0.008)	7.637*** (0.009)	7.527*** (0.008)	7.704*** (0.009)
Observações	108,799	108,799	108,799	108,799
R ²	0.321	0.234	0.302	0.332
R ² Ajustado	0.321	0.234	0.302	0.331
Estatística F	7,345.949***	4,750.393***	6,721.881***	5,994.333***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.



Tabela A4 – Resultados das estimações com efeitos fixos para as mulheres

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.155*** (0.008)			-0.277*** (0.018)
<i>Mismatch</i> horizontal		-0.029*** (0.009)		-0.110*** (0.012)
<i>Mismatch</i> vertical x horizontal			-0.116*** (0.008)	0.135*** (0.019)
Norte	-0.083 (0.097)	-0.067 (0.098)	-0.066 (0.097)	-0.081 (0.097)
Nordeste	0.204*** (0.067)	0.214*** (0.068)	0.208*** (0.067)	0.207*** (0.067)
Centro-Oeste	0.338*** (0.040)	0.328*** (0.040)	0.333*** (0.040)	0.335*** (0.040)
Sudeste	0.202*** (0.021)	0.199*** (0.021)	0.204*** (0.021)	0.201*** (0.021)
Horas contratadas	0.021*** (0.0003)	0.020*** (0.0003)	0.021*** (0.0003)	0.021*** (0.0003)
Tempo empregado(a)	0.002*** (0.0001)	0.002*** (0.0001)	0.002*** (0.0001)	0.002*** (0.0001)
Observações	57,98	57,98	57,98	57,98
R ²	0.110	0.104	0.108	0.112
R ² Ajustado	-0.029	-0.037	-0.033	-0.028
Estatística F	887.501***	830.574***	862.589***	700.521***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em
entre parênteses estão os erros-padrão.



Tabela A5 – Resultados das estimações com efeitos aleatórios para as mulheres

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.318*** (0.015)			-0.492*** (0.035)
<i>Mismatch</i> horizontal		-0.080*** (0.018)		-0.207*** (0.024)
<i>Mismatch</i> vertical x horizontal			-0.267*** (0.016)	0.187*** (0.037)
Norte	0.160 (0.169)	0.189 (0.174)	0.189 (0.169)	0.174 (0.167)
Nordeste	0.223* (0.121)	0.255** (0.124)	0.235* (0.121)	0.233* (0.120)
Centro-Oeste	0.566*** (0.069)	0.550*** (0.072)	0.562*** (0.070)	0.553*** (0.069)
Sudeste	0.277*** (0.036)	0.275*** (0.037)	0.284*** (0.036)	0.280*** (0.036)
Horas contratadas	0.021*** (0.001)	0.019*** (0.001)	0.020*** (0.001)	0.021*** (0.001)
Tempo empregado(a)	0.004*** (0.0001)	0.003*** (0.0001)	0.004*** (0.0001)	0.004*** (0.0001)
Constante	7.486*** (0.036)	7.486*** (0.045)	7.465*** (0.044)	7.649*** (0.040)
Observações	57,98	57,98	57,98	57,98
R ²	0.306	0.220	0.284	0.321
R ² Ajustado	0.306	0.220	0.284	0.321
Estatística F	2,415.570***	1,859.155***	2,252.581***	2,535.709***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.



Tabela A6 – Resultados das estimações OLS pooled para as mulheres

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.563*** (0.006)			-0.730*** (0.014)
<i>Mismatch</i> horizontal		-0.169*** (0.008)		-0.306*** (0.010)
<i>Mismatch</i> vertical x horizontal			-0.534*** (0.007)	0.180*** (0.016)
Norte	0.335*** (0.059)	0.470*** (0.062)	0.372*** (0.059)	0.343*** (0.058)
Nordeste	0.284*** (0.047)	0.403*** (0.050)	0.326*** (0.048)	0.311*** (0.047)
Centro-Oeste	0.690*** (0.024)	0.779*** (0.026)	0.702*** (0.024)	0.662*** (0.024)
Sudeste	0.351*** (0.012)	0.398*** (0.013)	0.372*** (0.013)	0.361*** (0.012)
Horas contratadas	0.024*** (0.0003)	0.018*** (0.0003)	0.022*** (0.0003)	0.026*** (0.0003)
Tempo empregado(a)	0.005*** (0.00004)	0.005*** (0.00005)	0.005*** (0.00004)	0.005*** (0.00004)
Constante	7.487*** (0.010)	7.629*** (0.012)	7.507*** (0.011)	7.673*** (0.012)
Observações	57,98	57,98	57,98	57,98
R ²	0.317	0.231	0.300	0.329
R ² Ajustado	0.317	0.231	0.300	0.329
Estatística F	3,837.712***	2,488.700***	3,542.966***	3,158.394***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.



Tabela A7 – Resultados das estimações com efeitos fixos para os homens

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.127*** (0.009)			-0.250*** (0.020)
<i>Mismatch</i> horizontal		-0.030*** (0.010)		-0.101*** (0.013)
<i>Mismatch</i> vertical x horizontal			-0.094*** (0.008)	0.136*** (0.021)
Norte	0.293*** (0.072)	0.293*** (0.072)	0.295*** (0.072)	0.289*** (0.072)
Nordeste	0.293*** (0.077)	0.288*** (0.077)	0.284*** (0.077)	0.303*** (0.077)
Centro-Oeste	0.315*** (0.038)	0.303*** (0.038)	0.307*** (0.038)	0.313*** (0.038)
Sudeste	0.227*** (0.023)	0.227*** (0.023)	0.227*** (0.023)	0.230*** (0.023)
Horas contratadas	0.019*** (0.0003)	0.019*** (0.0003)	0.019*** (0.0003)	0.020*** (0.0003)
Tempo empregado(a)	0.001*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0001)
Observações	50,819	50,819	50,819	50,819
R ²	0.091	0.087	0.090	0.093
R ² Ajustado	-0.053	-0.058	-0.055	-0.051
Estatística F	630.629***	598.089***	615.993***	498.074***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em
entre parênteses estão os erros-padrão.



Tabela A8 – Resultados das estimações com efeitos aleatórios para os homens

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.300*** (0.016)			-0.485*** (0.039)
<i>Mismatch</i> horizontal		-0.070*** (0.020)		-0.188*** (0.026)
<i>Mismatch</i> vertical x horizontal			-0.242*** (0.017)	0.205*** (0.040)
Norte	0.445*** (0.128)	0.458*** (0.132)	0.454*** (0.129)	0.444*** (0.128)
Nordeste	0.428*** (0.137)	0.424*** (0.141)	0.413*** (0.138)	0.439*** (0.136)
Centro-Oeste	0.479*** (0.067)	0.454*** (0.069)	0.465*** (0.067)	0.474*** (0.066)
Sudeste	0.319*** (0.040)	0.319*** (0.041)	0.322*** (0.040)	0.324*** (0.040)
Horas contratadas	0.020*** (0.001)	0.019*** (0.001)	0.020*** (0.001)	0.021*** (0.001)
Tempo empregado(a)	0.003*** (0.0001)	0.003*** (0.0001)	0.003*** (0.0001)	0.003*** (0.0001)
Constante	7.655*** (0.034)	7.643*** (0.044)	7.630*** (0.044)	7.807*** (0.041)
Observações	50,819	50,819	50,819	50,819
R ²	0.315	0.227	0.291	0.326
R ² Ajustado	0.315	0.227	0.291	0.326
Estatística F	1,974.316***	1,482.074***	1,796.333***	2,044.370***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

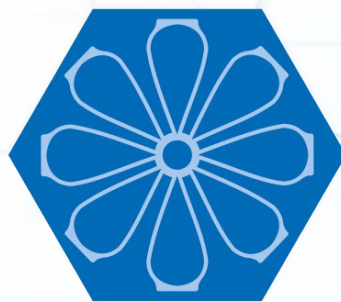
Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.

Tabela A9 – Resultados das estimações OLS *pooled* para os homens

Variáveis	1	2	3	4
<i>Mismatch</i> vertical	-0.592*** (0.007)			-0.773*** (0.016)
<i>Mismatch</i> horizontal		-0.166*** (0.009)		-0.278*** (0.011)
<i>Mismatch</i> vertical x horizontal			-0.549*** (0.007)	0.209*** (0.018)
Norte	0.487*** (0.049)	0.654*** (0.052)	0.530*** (0.050)	0.503*** (0.049)
Nordeste	0.442*** (0.051)	0.531*** (0.054)	0.442*** (0.051)	0.444*** (0.050)
Centro-Oeste	0.561*** (0.022)	0.578*** (0.024)	0.550*** (0.023)	0.560*** (0.022)
Sudeste	0.427*** (0.014)	0.502*** (0.015)	0.458*** (0.014)	0.437*** (0.014)
Horas contratadas	0.026*** (0.0003)	0.021*** (0.0003)	0.025*** (0.0003)	0.027*** (0.0003)
Tempo empregado(a)	0.005*** (0.00005)	0.005*** (0.0001)	0.005*** (0.00005)	0.005*** (0.00005)
Constante	7.584*** (0.012)	7.677*** (0.015)	7.585*** (0.013)	7.770*** (0.014)
Observações	50,819	50,819	50,819	50,819
R ²	0.326	0.236	0.306	0.335
R ² Ajustado	0.326	0.236	0.305	0.335
Estatística F	3,511.961***	2,247.878***	3,193.693***	2,839.907***

Fonte: Elaboração própria a partir de dados da UFPR e da RAIS.

Nota: * Significante a 10%, ** Significante a 5%, *** Significante a 1%, em entre parênteses estão os erros-padrão.



NEDUR



**Núcleo de Estudos em Desenvolvimento Urbano e Regional
Universidade Federal do Paraná**

 Av. Prefeito Lothário Meissner, nº 632 – Setor de Ciências Sociais | UFPR

 www.nedur.ufpr.br

 nedur.ufpr@gmail.com