



## **Análise de clusters para tipologias do Quadro Brasileiro de Qualificações: evidências para o mercado de trabalho formal em Minas Gerais (2021)<sup>1</sup>**

**Ingrid Machado Mendonça**

Mestranda do Programa de Pós-Graduação em Economia do Centro de Desenvolvimento e Planejamento Regional (Cedeplar/UFMG)

E-mail: [ingridmachadomendonca@gmail.com](mailto:ingridmachadomendonca@gmail.com)

**Felipe Pureza Cardoso**

Doutorando do Programa de Pós-Graduação em Economia do Centro de Desenvolvimento e Planejamento Regional (Cedeplar/UFMG)

E-mail: [felppureza@gmail.com](mailto:felppureza@gmail.com)

**Matheus Ferreira dos Reis**

Mestrando do Programa de Pós-Graduação em Economia do Centro de Desenvolvimento e Planejamento Regional (Cedeplar/UFMG)

E-mail: [theusfe@egresso.ufg.br](mailto:theusfe@egresso.ufg.br)

**Resumo:** Este trabalho apresenta uma classificação das ocupações da Classificação Brasileira de Ocupações (CBO) a partir das competências exigidas para a execução das tarefas, com base no Quadro Brasileiro de Qualificações (QBQ). Especificamente, busca-se desenvolver uma nova tipologia ocupacional que considera os conhecimentos, habilidades e atitudes associados às ocupações, além de verificar de que forma as diferenças salariais por gênero se manifestam dentro dos grupos formados. Para o exercício de classificação, é empregado o método de fuzzy cluster, que resulta na criação de seis categorias ocupacionais distintas, considerando as variações nos níveis de qualificação exigidos em cada uma. Em seguida, aplica-se o método de decomposição de Oaxaca-Blinder aos dados do mercado de trabalho formal de Minas Gerais referentes a 2021 para investigar o papel das competências nas desigualdades salariais. Os resultados indicam a existência de uma vantagem salarial para os homens, atribuída principalmente ao componente associado à discriminação. Ademais, a classificação proposta permite observar que as características produtivas dos homens contribuem para ampliar o já existente gap de rendimentos por gênero em ocupações do tipo “mista” em relação às atitudes e naquelas mais intensivas em conhecimento STEM.

**Palavras-chaves:** I Quadro Brasileiro de Qualificações; fuzzy cluster; classificação ocupacional.

---

<sup>1</sup> Este trabalho, assim como outras investigações relacionadas, foi e continua a ser desenvolvido no âmbito do Grupo de Pesquisas em Economia e Demografia da Estratificação Social (PEDES) do Cedeplar/UFMG.

**Abstract:** *This paper presents a classification of occupations from the Classificação Brasileira de Ocupações (CBO), grounded in the skills required to perform tasks as outlined in the Quadro Brasileiro de Qualificações (QBQ). It aims to develop a new occupational typology that considers the knowledge, skills, and attitudes associated with occupations, and to examine how gender wage disparities emerge within the resulting groups. The fuzzy cluster method is used to generate six distinct occupational categories, reflecting variations in qualification levels. Subsequently, the Oaxaca-Blinder decomposition is applied to data from the formal labor market in Minas Gerais for the year 2021 to investigate the role of competencies in wage inequalities. The results reveal a wage advantage for men, mainly attributed to the discrimination component. Furthermore, the proposed classification highlights how men's productive characteristics contribute to widening the existing gender wage gap in occupations classified as “mixed” in terms of attitudes and in those more intensive in STEM-related knowledge.*

**Keywords:** *Quadro Brasileiro de Qualificações; fuzzy cluster; occupational classification.*

**Classificação JEL:** J01, J71

## 1. INTRODUÇÃO

As ocupações têm recebido destaque como sendo uma unidade de análise importante para compreensão de relações entre agentes no mercado de trabalho. Estudos como os de Silva e Vaz (2022) e Lordan e Pischke (2022), por exemplo, utilizam dessa unidade para evidenciar o papel da segmentação ocupacional, de modo a explicar o *gap* salarial por gênero em postos de trabalho em que a participação é tipicamente feminina. Outra abordagem desenvolvida na literatura recente de mercado de trabalho concebe o conceito de ocupações como um conjunto de habilidades e tarefas relacionadas, o que segundo Yamaguchi (2012), fornece uma interpretação clara sobre como e por que as competências são recompensadas de forma diferente entre as ocupações. Além disso, a perspectiva das tarefas permite compreender questões como o tipo de habilidades que são demandas por empregadores, o porquê de elas serem exigidas e como essas relações evoluem ao longo do tempo, conforme explicado em Autor e Handel (2013).

Nesse sentido, a literatura internacional tem demonstrado um maior interesse em relacionar as habilidades e tarefas ocupacionais com o desempenho dos trabalhadores. Spitz-Oener (2006) e Koomen e Backes-Gellner (2022), por exemplo, relacionam a influência das mudanças tecnológicas nos tipos de tarefas que são executadas. Já Gathmann e Schönberg (2010) e Cortes e Gallipoli (2018) investigam como esses pontos estão ligados à

mobilidade laboral. No âmbito nacional, esse assunto tem sido abordado em trabalhos como o de Reis (2016), Júnior e Funchal (2016), Reis (2018), Cardoso, Hartmann e Kealtenberg (2023) e Monsueto, Cardoso e Oliveira (2024).

No Brasil, no entanto, as análises que utilizam esse nível de detalhamento das ocupações ainda são incipientes, uma vez que até 2022 não existiam descritores nacionais para os postos de trabalho. Com o lançamento do Quadro Brasileiro de Qualificações (QBQ) naquele ano, fruto de uma parceria entre a Fundação Instituto de Pesquisas Econômicas (FIPE) em parceria com o Ministério do Trabalho e Emprego, novas investigações foram viabilizadas. Conforme descrito pelos órgãos, ele é um conjunto de informações que descrevem o perfil de mais de 2.700 ocupações registradas na Classificação Brasileira de Ocupações (CBO)<sup>2</sup> a partir dos conhecimentos, habilidades e atitudes necessários para o desempenho das tarefas necessárias do posto de trabalho.

Desse modo, este trabalho busca explorar esse novo conjunto de dados para propor uma classificação das ocupações e contribuir com a discussão acerca do tema a partir da aplicação de uma metodologia de agrupamento difuso, que permite considerar as particularidades desses dados. Especificamente, busca-se desenvolver uma nova tipologia ocupacional que considera os conhecimentos, habilidades e atitudes exigidos para a execução das tarefas, conforme fornecido pelo QBQ, e verificar de que forma as diferenças salariais se manifestam dentro dos grupos, considerando o gênero do trabalhador. A hipótese é de que haja um *gap* salarial que favorece os homens, conforme identificado em estudos como o de Stier e Yaish (2014) e Silva e Vaz (2022), mas que exista diferentes padrões de discriminação baseadas nas características dos agrupamentos de ocupações.

Para isso, é utilizada a técnica de *fuzzy cluster* e, a partir da classificação proposta, pretende-se analisar como os trabalhadores são alocados nesses grupos ocupacionais. Em seguida, é realizado um exercício de decomposição do diferencial salarial por gênero entre os agrupamentos gerados, a partir do método de Oaxaca-Blinder (Oaxaca, 1973; Blinder,

---

<sup>2</sup> Pelo fato de as características das ocupações estarem atreladas à CBO, é possível retroagir essas informações para base de dados anterior a 2022.

1973). Nessa etapa, os dados do QBQ são combinados às informações de vínculos contratuais da Relação Anual de Informações Sociais para Minas Gerais no ano de 2021. Com isso, tem-se o intuito de desenvolver novas classificações baseadas nas competências exigidas pelas ocupações, utilizando-as como um novo ferramental para diversos tipos de estudos laborais.

Além desta introdução, o artigo está organizado em outras quatro seções. A seguir, uma revisão de literatura é apresentada, abordando algumas discussões acerca do impacto das habilidades sobre a estrutura salarial. Mais adiante, a seção metodológica contém as descrições dos dados e técnicas utilizados, seguido pelos principais resultados obtidos e, por fim, as considerações finais do trabalho.

## 2. REVISÃO DE LITERATURA

O tema de tarefas no mercado de trabalho, de forma geral, está relacionado à forma com que a popularização de novas técnicas tem impactado as relações laborais, principalmente no que diz respeito à demanda do mercado por maiores níveis de qualificação. Spitz-Oener (2006), com dados da *Survey of the Working Population on Qualification and Working Conditions in Germany* verifica, no contexto de crescente aumento de oferta de mão de obra mais qualificada e com maiores retornos para a escolaridade, se os requisitos de habilidades no local de trabalho também vêm aumentando e se essas mudanças estão relacionadas às transformações tecnológicas.

Com dados da Alemanha Ocidental, a autora classifica as tarefas entre analítica não-rotineira, interativa não-rotineira, rotina cognitiva, manual rotineira e manual não-rotineira e desenvolve uma proporção de cada um dos tipos em relação ao total de atividades desenvolvidas pelos trabalhadores em cada etapa temporal considerada. Utilizando de uma série de análises de regressão, a autora conclui que as ocupações têm experimentado mudanças em direção a ter mais atividades analíticas e interativas e se distanciando de tarefas cognitivas e manuais, o que seria uma evidência de que a adoção de computadores no ambiente de trabalho exige o uso de habilidades mais complexas do que era demandado

anteriormente (Spitz-Oener, 2006).

A análise de Koomen e Backes-Gellner (2022), da mesma forma, relaciona as mudanças nas tarefas exigidas pelas ocupações com o desenvolvimento tecnológico. Nesse estudo, é investigado o papel das tarefas nas desigualdades salariais sob a hipótese de que as mudanças tecnológicas recentes tendem a substituir atividades rotineiras, o que leva a uma redução de rendimentos e da empregabilidade das ocupações intensivas nesse tipo de atividades. Para isso, os autores classificam as ocupações de acordo com a sua intensidade em tarefas abstratas, interativas, rotineiras e manuais usando informações laborais de homens nativos na Alemanha Ocidental que ganham pelo menos 12 euros por dia nos períodos de 1978/79, 1986/86, 1991/92, 1998/99 e 2005/06 também com dados da *Survey of the Working Population on Qualification and Working Conditions in Germany*.

Através de decomposições RIF (*recentered influence functions*), os autores verificam que o conteúdo das ocupações varia ao longo do tempo, o que pode implicar em mudanças nas estruturas dos empregos dado um avanço tecnológico, além de averiguar que entre as décadas de 1970 e 1990, os aspectos de educação e de experiência do trabalhador eram o suficiente para explicar as diferenças salariais. A partir de 1990, a composição de tarefas começou a ter um maior efeito, com as mudanças em retornos nas ocupações rotineiras influenciando no aumento do *gap* na parte inferior da distribuição de rendimentos e os retornos nas cognitivas impactando nos diferenciais na parte mais alta (Koomen e Backes-Gellner, 2022).

Os efeitos das competências relacionadas às ocupações na composição da estrutura salarial e na forma com que os trabalhadores transitam no mercado de trabalho também são tratados em estudos como de Gathemann e Schönberg (2010) e Cortes e Gallipoli (2018). Nesse primeiro, por exemplo, os autores utilizam dados do *German Qualification and Career Survey* para a análise das tarefas relacionadas às ocupações, combinados com o painel de dados de emprego alemão, e encontram evidências de que os indivíduos tendem a mudar para ocupações que exigem tarefas semelhantes e que a distância dessas mudanças diminui conforme se acumula experiência. Segundo Cortes e Gallipoli (2018), essa literatura

demonstra que se o capital humano for específico para uma tarefa, deverá ser parcialmente transferível para ocupações em que seja desempenhado um conjunto semelhante de atividades. Por meio dessa perspectiva, os autores desenvolvem uma estrutura que permite verificar os diferentes custos da mobilidade entre pares de ocupações de acordo com a extensão da sobreposição de suas tarefas.

No Brasil, algumas investigações utilizam a abordagem de ocupações baseadas em tarefas para analisar, além das transformações da demanda por tipos de atividades no mercado de trabalho e mobilidade ocupacional, questões como *mismatch*, ou mesmo para descrever as ocupações a partir dessas tipologias (Reis, 2016; Júnior e Funchal, 2016; Cardoso, Hartmann e Kealtenberg, 2023; Monsueto, Cardoso e Oliveira, 2024). Reis (2018), por exemplo, aplica as informações de tarefas da Classificação Brasileira de Ocupações (CBO) à proposição de uma medida que representa a distância entre as competências adquiridas com a formação superior àquelas requeridas na ocupação do trabalhador. Combinando essa medida aos dados do Censo Demográfico de 2010 e a partir do método de mínimos quadrados ordinários (MQO), o autor demonstra que a penalidade aos indivíduos classificados com *mismatch* ocupacional, em termos de ganhos do trabalho por hora, é maior conforme aumenta essa distância.

No contexto brasileiro, todavia, aplicações relacionadas ao tema ainda são incipientes devido à falta de informações detalhadas sobre as atividades desempenhadas nas diversas ocupações (Reis, 2016). De modo geral, os pesquisadores recorrem à Classificação Brasileira de Ocupações (CBO) em conjunto com as tipologias sugeridas por dicionários internacionais, já que somente em 2019 foi iniciado o projeto para a definição dos descritores nacionais específicos para as ocupações. O projeto “Estudo, Elaboração e Implantação do QBQ - Quadro Brasileiro de Qualificações”, proposto pela Coordenação da Classificação Brasileira de Ocupações (CCBO), divulgou a primeira versão desse ferramental em 2022, viabilizando, portanto, novas investigações mais abrangentes das qualificações exigidas e das transformações profissionais no Brasil. Desse modo, este trabalho tem por objetivo desenvolver uma classificação ocupacional a partir desse novo conjunto de dados e

contribuir com a discussão acerca do tema.

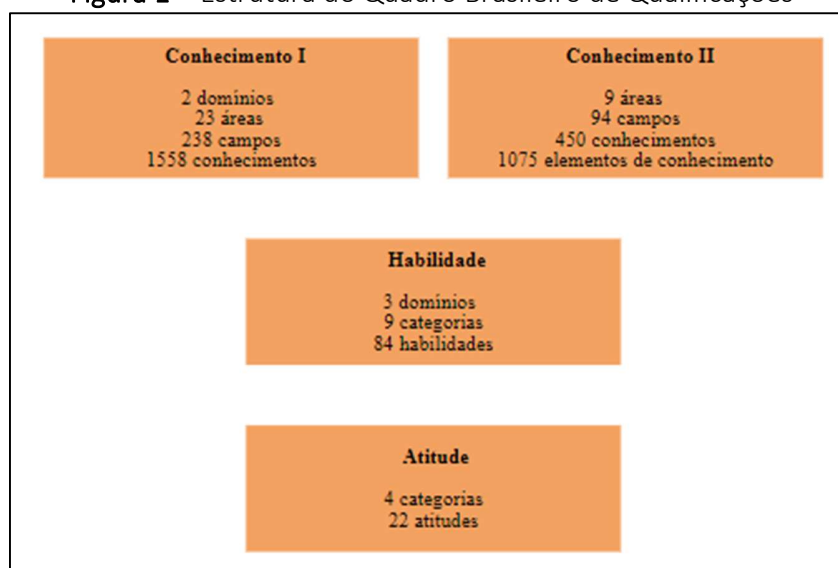
### **3. METODOLOGIA**

#### **3.1. Bases de dados**

Para atingir os objetivos desta análise, são utilizadas informações provenientes de duas fontes de dados. A primeira é o Quadro Brasileiro de Qualificações (QBQ) que, de acordo com o Ministério do Trabalho e Emprego (MTE), é uma ferramenta de análise do mercado de trabalho que mapeia conhecimentos, habilidades e atitudes para cada uma das ocupações listadas na Classificação Brasileira de Ocupações (CBO). Com base nessas informações, o QBQ permite definir os perfis ocupacionais e os indicadores associados ao preparo necessário para o desempenho do trabalhador em cada uma das ocupações oficialmente reconhecidas no mercado de trabalho brasileiro, abrangendo cerca de 2.700 ocupações constantes na CBO. Apesar de ter sido lançada somente em 2022, essa ferramenta apresenta indicadores que não variam ao longo do tempo, o que possibilita sua aplicação em análises de diferentes contextos e períodos.

A Figura 1 apresenta a estrutura das terminologias do QBQ utilizadas neste trabalho. A primeira terminologia selecionada, de conhecimento, define o conjunto de informações, fatos, teorias, práticas e princípios necessários para o exercício de uma ocupação ou para a obtenção de uma qualificação profissional. Nessa terminologia, as ocupações são divididas em dois grupos com base no nível de qualificação exigido. O primeiro grupo representa as ocupações que requerem programas de educação profissional e tecnológica, enquanto o segundo grupo exige um nível de educação superior, incluindo graduação (exceto formação tecnológica) e pós-graduação que abrange programas de mestrado e doutorado, bem como cursos de especialização e aperfeiçoamento.

**Figura 1** – Estrutura do Quadro Brasileiro de Qualificações



Fonte: Elaboração própria. Secretaria do Trabalho – MTE.

A terminologia de habilidade refere-se à competência de aplicar conhecimentos e utilizar recursos adquiridos para realizar tarefas e solucionar problemas. Por último, a dimensão de atitude está relacionada à capacidade de desempenhar tarefas e resolver problemas com diferentes níveis de autonomia e responsabilidade. Com o objetivo de definir os níveis de qualificação de cada ocupação avaliada, o QBQ organiza as terminologias de maneira detalhada e hierárquica, como ilustrado na Figura 1. No campo das habilidades, por exemplo, as ocupações são classificadas em três domínios: habilidades cognitivas, habilidades práticas e habilidades físicas, psicomotoras e sensoriais. Cada domínio é subdividido em três categorias, que, por sua vez, são divididas no nível mais detalhado, as habilidades, totalizando 84.

Além disso, o QBQ define as variáveis de profundidade, frequência e importância para determinar, respectivamente, o grau de complexidade, a frequência de uso e o peso desses itens na execução das tarefas. Essas variáveis são classificadas em uma escala de 1 a 5, em que 1 representa muito pequeno e 5 representa muito grande e auxiliam na classificação das ocupações em diferentes níveis de qualificação. No caso do conhecimento, todas as três variáveis são consideradas. Para habilidades, são levadas em conta a frequência



e a importância. Já em atitudes, é considerada apenas a importância.

A segunda fonte de dados utilizada é a Relação Anual de Informações Sociais (RAIS), também organizada pelo Ministério do Trabalho e Emprego. A RAIS tem como objetivo compilar e fornecer informações detalhadas sobre os estabelecimentos formais e os vínculos empregatícios celetistas e estatutários registrados no país e, neste estudo, é empregada na aplicação prática das tipologias ocupacionais derivadas do Quadro Brasileiro de Qualificações (QBQ). Foram selecionados especificamente os dados referentes ao estado de Minas Gerais no ano de 2021. O recorte regional justifica-se pelo porte e diversidade econômica do estado, que oferece uma boa aproximação da realidade brasileira em termos de estrutura produtiva e distribuição territorial. Já o recorte temporal foi escolhido por corresponder à base mais atualizada disponível aos pesquisadores no momento da realização deste trabalho.

A amostra foi restrita aos contratos de trabalho ativos em 31 de dezembro de 2021, para trabalhadores em idade ativa, entre 18 e 65 anos. Além disso, foram considerados apenas vínculos com jornada semanal igual ou superior a 20 horas, a fim de garantir a análise de vínculos mais estáveis e evitar possíveis vieses advindos de empregos temporários. Excluíram-se também os vínculos com salários zerados em dezembro, uma vez que esses registros podem indicar vínculos formais inativos, suspensos ou inconsistências nos dados, o que comprometeria a análise das condições reais de trabalho. A amostra final corresponde, portanto, às informações de 4.691.738 vínculos empregatícios. Ademais, para combinar as informações sobre as qualificações, conhecimentos, habilidades e atitudes exigidas para o desempenho das ocupações analisadas pelo QBQ com os dados contratuais disponíveis na RAIS, são utilizados os códigos de 6 dígitos da CBO.

### 3.2. Estratégia empírica

#### 3.2.1. Fuzzy Cluster

Dentre as opções metodológicas adotadas para a redução da dimensionalidade de dados, as técnicas de análise de *cluster* aparecem como alternativa para a identificação de grupos e classificação de objetos. Neste caso, essa metodologia é empregada a fim de se

gerar tipologias ocupacionais, a partir do agrupamento de postos de trabalho semelhantes em termos dos tipos de conhecimento, habilidades e atitudes que exigem. Para isso, são utilizadas as proporções calculadas das terminologias do Quadro Brasileiro de Qualificações no nível mais agregado de detalhamento, conforme apresenta o Quadro 1, ponderadas pelos indicadores de grau de profundidade, nível de importância e frequência de cada elemento, bem como pela diversidade de tarefas associadas às ocupações.

Para o caso de Conhecimento II, o nível mais agregado de informações corresponde às nove áreas de conhecimento, que foram previamente classificadas em *STEM*, *Shape* e mista para facilitar a análise e a interpretação dos dados. A primeira terminologia, é uma sigla em inglês para "*Science, Technology, Engineering, and Maths*", em português, ciência, tecnologia, engenharia e matemática. A segunda, *Shape*, abrange as áreas remanescentes, enquanto na terceira classificação, mista, são incluídas áreas com características intermediárias às demais.

**Quadro 1** – Categorias consideradas para o agrupamento das ocupações

Terminologia	Variáveis (%)
Conhecimentos I	Formação geral e/ou transversal; Formação específica ou técnica-profissional.
Conhecimentos II	9 Áreas divididas em: <i>STEM</i> <i>Shape</i> Misto
Habilidades	Habilidades cognitivas; Habilidades práticas; Habilidades físicas, psicomotoras e sensoriais.
Atitudes	Trabalho sob supervisão direta; Autonomia (ausência de supervisão direta) no próprio trabalho; Supervisão do trabalho de outros; Avaliação de trabalho ou atividade.

Fonte: Elaboração própria. Secretaria do Trabalho – MTE.

São realizados exercícios separadamente para as ocupações listadas nas categorias de Conhecimentos I e Conhecimentos II. Isso ocorre por duas razões: a primeira por, conceitualmente, esses grupos já serem heterogêneos entre si, por exigirem níveis de

qualificação distintos e bem definidos. Em segundo lugar, por motivos metodológicos, essa divisão é necessária porque os grupos não compartilham as mesmas variáveis e, para a viabilidade da análise proposta, é crucial garantir a ausência de dados faltantes. Ademais, durante a execução, uma categoria de cada terminologia é omitida para evitar a redundância das informações.

A partir disso, é empregada a técnica de *Fuzzy Cluster* que se diferencia dos usuais por fugir da lógica rígida e binária de pertencimento ou não pertencimento ao grupo indicado. Segundo Everitt *et al.* (2011), ao invés disso, os objetos possuem uma função de associação que indica o grau de pertencimento em todos ou alguns dos *clusters*. Isto é, as ocupações podem ser representadas em um ou mais grupos, a depender da proximidade de suas características.

De acordo com Simões (2005), essa metodologia deve receber a devida atenção quando as imprecisões do conjunto de informações não derivam, necessariamente, da aleatoriedade advinda de processos estocásticos, mas sim da ausência de fronteiras bem definidas entre as classes. Dada a natureza dos dados do QBQ, em que diferentes postos de trabalho compartilham conhecimentos, habilidades e atitudes em comum, considera-se que essa fronteira não é clara e, em certa medida, são competências transferíveis entre as ocupações. Portanto, esse método de agrupamento parece mais adequado para a criação desta taxonomia. No estudo de Suleman, Suleman e Reis (2016), os autores comparam a eficiência do agrupamento por essa metodologia com a de *clusters* hierárquicos para a redução de dados discretos de habilidades e demonstram, empiricamente, que o método de *Fuzzy* é mais eficiente, propõe menos grupos e é mais fácil de interpretar.

Nesse tipo de análise, assume-se que o número de subconjuntos é conhecido, e a função de associação de cada objeto a cada agrupamento é estimada através de um método iterativo, não necessariamente obedecendo as regras da teoria da probabilidade. Apesar disso, uma vez encontradas, as associações podem ser escalonadas entre zero e um e, então, podem ser interpretadas como probabilidades de pertencimento aos *clusters*. Especificamente para o método FANNY, descrito por Kaufman e Rousseeuw (2005) e

adotado neste estudo, para cada objeto  $i$  e cada *cluster*  $t$ , existe uma associação  $u_{it}$  que indica esse grau de pertinência do objeto ao grupo, que são obtidas a partir da minimização da função  $f$ :

$$f = \sum_{t=1}^k \{ \sum_{i,j=1}^n u_{it}^2 u_{jt}^2 d(x_i x_j) / 2 \sum_{j=1}^n u_{jt}^2 \} \quad (1)$$

Em que  $d(x_i x_j)$  representa as distâncias conhecidas entre os objetos  $i$  e  $j$  enquanto o pertencimento  $u_{it}$  é desconhecido e deve obedecer às condições:

$$\begin{aligned} u_{it} &\geq 0 \text{ para todo } i = 1, \dots, n; t = 1, \dots, k \\ \sum_{t=1}^g u_{it} &= 1 \text{ para } i = 1, \dots, n \end{aligned} \quad (2)$$

Que demonstram que as associações entre cada objeto e os agrupamentos devem ser nulas ou positivas e a soma dessas é constante e igual a 1.

Para a determinação do número de *clusters*, foi utilizado a média do tamanho das silhuetas proposta por Kaufman e Rousseeuw (1990). Nesse, para cada objeto  $i$  é definido um índice  $s(i) \in [-1, 1]$  que compara a distância do objeto ao *cluster* que pertence com a heterogeneidade desse agrupamento. Quando  $s(i)$  assume um valor próximo de 1, a heterogeneidade do cluster do objeto  $i$  é menor que sua distância e, portanto, o objeto está bem classificado. A relação oposta ocorre quando o indicador se aproxima de -1, mas não é clara a classificação do objeto quando o indicador se aproxima de zero, de modo que não é possível definir se esse estaria mais bem alocado no grupo atual ou no imediatamente vizinho. Essas representações foram comparadas, possibilitando a escolha do número de grupos que maximiza a qualidade das classificações.

### 3.2.2. Decomposição de Oaxaca

O segundo exercício deste artigo é uma decomposição de Oaxaca-Blinder (Oaxaca, 1973; Blinder, 1973), utilizada para compreender como se dão os diferenciais de remuneração entre homens e mulheres, ao se considerar as tipologias ocupacionais obtidas a partir da aplicação do *Fuzzy Cluster*. O método divide a diferença total entre os grupos sociais em dois componentes: um explicado pelas diferenças nas características observáveis

— como escolaridade, experiência e tipo de ocupação — e outro componente não explicado, que pode ser atribuído a fatores não observáveis. Esses podem incluir habilidades não mensuradas, redes de contato, preferências individuais, responsabilidades familiares, estilos de negociação salarial ou ainda práticas discriminatórias de gênero por parte dos empregadores. O último componente é comumente interpretado, na literatura de economia do trabalho, como um “efeito discriminatório”.

As variáveis utilizadas para calcular o salário-hora utilizada no exercício de decomposição é a remuneração de dezembro e a quantidade de horas contratuais. Como variáveis explicativas, são incluídas *dummies* de faixa etária (18 a 29 anos e 30 a 50 anos, com a variável de 51 anos ou mais sendo a referência), binárias referentes ao grau de escolaridade do trabalhador (ensino fundamental completo, ensino médio completo e ensino superior completo, com o ensino fundamental incompleto de referência), o setor de atividade (comércio, construção civil, serviços distributivos, serviços produtivos, serviços sociais, serviços pessoais, administração pública, atividades agrícolas e demais setores, com a indústria e extrativismo mineral de categoria de referência), o tamanho do estabelecimento (médio-pequeno porte, médio porte, médio-grande porte e grande porte, com o pequeno porte de referência) e o tempo de emprego com o seu quadrado.

Dessa forma, o salário-hora para a amostra de homens e mulheres no mercado de trabalho formal em Minas Gerais é representado, respectivamente, pelas seguintes equações:

$$W_{h,i} = x'_{h,i}b_h + \varepsilon_{h,i} \quad (3)$$

$$W_{m,i} = x'_{m,i}b_m + \varepsilon_{m,i} \quad (4)$$

Onde  $x'_g$  é um vetor de características observáveis,  $b_g$  são os coeficientes estimados,  $W_g$  representa a remuneração média por hora e  $\varepsilon_g$  o erro estocástico para todos os indivíduos no grupo  $g$  (onde  $g = h, m$ ). Desse modo, a diferença salarial média geral pode ser definida como:

$$E[W_{h,i}|x_{h,i}] - E[W_{m,i}|x_{m,i}] = x'_{h,i}b_h - x'_{m,i}b_m \quad (5)$$

Para decompor os efeitos é somado e subtraído o salário médio contrafactual que os

trabalhadores homens receberiam sob a estrutura salarial das mulheres,  $x'_{h,i}b_m$ , a expressão anterior se torna:

$$E[W_{h,i}|x_{h,i}] - E[W_{m,i}|x_{m,i}] = x'_{h,i}(b_h - b_m) - (x_{h,i} - x_{m,i})'b_m \quad (6)$$

O segundo termo da equação anterior representa o componente explicado, que se deve às diferenças médias salariais nas características individuais dos trabalhadores entre homens e mulheres (por exemplo, educação, experiência, setor de atividade). Por outro lado, o primeiro termo representa o componente não explicado, ou seja, a parte do diferencial salarial que persiste mesmo quando consideramos indivíduos com características similares. Em termos mais detalhados, o componente não explicado avalia se há alguma diferença salarial entre homens e mulheres, assumindo que as mulheres possuam as mesmas qualificações, experiência, educação que os homens.

#### 4. RESULTADOS

O objetivo desta seção é apresentar os resultados obtidos a partir das metodologias de *cluster* e decomposição descritas anteriormente. A seção está dividida em duas subseções, em que a primeira aborda os resultados do *Fuzzy Cluster* e a caracterização dos grupos formados, enquanto a segunda apresenta as estimativas dos diferenciais de rendimento entre homens e mulheres, utilizando a decomposição de Oaxaca-Blinder.

##### 4.1. Clusters Ocupacionais – Conhecimento I

A metodologia de agrupamento foi inicialmente aplicada para classificar as ocupações incluídas na categoria de Conhecimento I, com a análise do tamanho médio das silhuetas indicando que o número ideal de *clusters* é três. A Figura 2, que representa cerca de 66,7% da variância das variáveis utilizadas na construção dos *clusters*, ilustra a organização desses grupos. Nessa figura, a alocação das ocupações em cada agrupamento depende da maior probabilidade de pertencimento, conforme calculado pela metodologia de *Fuzzy*. No entanto, ela evidencia que algumas ocupações, embora tenham maior grau de pertencimento a um grupo, podem estar próximas a outro *cluster* vizinho, o que justifica a

adoção do método e o uso dessas probabilidades como ponderadores na classificação das ocupações em um ou mais grupos.

**Figura 2** – Nuvem de pontos para as ocupações de Conhecimento 1



Fonte: elaboração própria baseada nos dados do Quadro Brasileiro de Qualificações.

Os grupos, portanto, foram definidos a partir da probabilidade de cada ocupação pertencer a um dos *clusters* gerados pela análise *fuzzy*. Com base nessa classificação, a Tabela 1 apresenta as proporções (%) das variáveis do Quadro Brasileiro de Qualificações (QBQ) em cada agrupamento, destacando as características que sustentam sua composição. Verifica-se que as dimensões de conhecimento e habilidades apresentam padrões relativamente semelhantes entre os grupos. A proporção de conhecimentos específicos, por exemplo, varia entre 66,4% e 72%, e sugere que a exigência dessa dimensão é uma característica comum das ocupações dentro desses conjuntos, todavia, mostra-se menos determinante na diferenciação entre os grupos formados. No que se refere às habilidades, o terceiro *cluster* se diferencia dos demais por apresentar maior presença de habilidades físicas (52,3%), em contraste com as dimensões prática e cognitiva. Nos dois outros grupos, os valores dessas três categorias são mais equilibrados.

Em contrapartida, a dimensão de atitudes, relacionada à capacidade de desempenhar tarefas e resolver problemas com diferentes níveis de autonomia e responsabilidade, apresenta padrões mais distintos entre os clusters analisados. Em todos os grupos, observa-se uma presença significativa das atitudes de avaliação, em que os percentuais variam de 45,4% a 50,1%. Essa categoria abrange desde a capacidade de avaliar o próprio desempenho, com ou sem orientação, até de avaliar o desempenho de outros trabalhadores, de setor de empresa ou instituição como um todo, podendo envolver autodesenvolvimento, gestão de desenvolvimento profissional de terceiros e processos de melhoria institucionais.

**Tabela 1** – Proporção (%) das variáveis do Quadro Brasileiro de Qualificações (QBQ) por cluster de Conhecimento I

Tipologia	Variáveis	Clusters de Conhecimento I		
		Misto em atitudes	Intensivo em atitudes autônomas	Intensivo em atitudes sob supervisão
Conhecimento I	Específicos	67,5	66,4	72,0
	Geral	32,5	33,6	28,0
Habilidades	Práticas	30,7	24,8	22,2
	Cognitivas	38,6	32,6	25,5
	Física	30,7	42,6	52,3
Atitudes	Sob supervisão	3,5	0,0	54,6
	Autônoma	23,8	49,9	0,0
	De supervisão	22,8	0,0	0,0
	De avaliação	49,9	50,1	45,4

Fonte: elaboração própria baseada nos dados do Quadro Brasileiro de Qualificações.

No primeiro *cluster*, a avaliação é predominante, com 49,9% das atitudes observadas, e se associa a uma distribuição relativamente equilibrada entre atitudes autônomas (23,8%) e de supervisão (22,8%), com baixa participação de atitudes sob supervisão. Como definição, as atitudes de supervisão referem-se à responsabilidade por coordenar e supervisionar o trabalho de outros profissionais, diferenciando-se das atitudes sob supervisão, que envolvem atuação sob supervisão direta, com pouca ou nenhuma autonomia. Já o segundo *cluster* se diferencia por uma ênfase nas atitudes autônomas, que representam 49,9% das ocupações, além de manter um peso relevante na dimensão de avaliação. As atitudes autônomas estão



associadas à execução de tarefas sem supervisão direta, em contextos de trabalho com ou sem mudanças previstas, ou até em projetos técnicos complexos que demandam abordagens estratégicas e inovadoras, incluindo atividades de pesquisa.

O terceiro *cluster*, por sua vez, apresenta uma clara predominância de atitudes sob supervisão (54,6%), caracterizando ocupações com menor grau de autonomia. Neste grupo, não há registro de atitudes autônomas nem de supervisão, além daquelas relacionadas à avaliação, indicando um padrão organizacional mais verticalizado. Com base nesses padrões de distribuição das atitudes, foram atribuídas denominações descritivas aos *clusters*, como sendo respectivamente: “misto em atitudes”, “intensivo em atitudes autônomas” e “intensivo em atitudes sob supervisão”. Essas denominações não seguem uma tipologia teórica predefinida, sendo construídas a partir da interpretação dos dados observados.

Essas tipologias foram combinadas aos dados do mercado de trabalho formal de Minas Gerais para o ano de 2021, extraídos da RAIS, a partir do código identificador ocupacional de 6 dígitos da Classificação Brasileira de Ocupações (CBO). As informações relativas ao salário-hora médio, à composição por sexo e por faixa etária, bem como ao tempo médio de emprego de cada *cluster*, são resumidas na Tabela 2. Para isso, considera-se o grau de pertencimento de cada ocupação aos agrupamentos como ponderador na análise individual, de forma que a contribuição de cada trabalhador reflita o grau com que sua ocupação se associa a cada grupo. Dessa forma, em relação às informações salariais, o *cluster* misto apresenta uma média ligeiramente superior à dos demais, embora essa diferença não seja muito expressiva, o mesmo ocorre com a maior proporção de mulheres, de pessoas com ensino superior e com o tempo médio de emprego. Já a composição etária dos *clusters* é bastante similar.

**Tabela 2** – Informações médias dos vínculos da RAIS em Minas Gerais (2021) por *cluster* de Conhecimento I

Variáveis	Misto em atitudes	Intensivo em atitudes autônomas	Intensivo em atitudes sob supervisão
Salário-hora (R\$)	16,1	14,4	11,3
Mulheres (%)	47,6	40,9	39,1
18 a 29 anos (%)	32,0	30,1	32,0
30 a 50 anos (%)	54,1	55,1	51,7
51 anos ou mais (%)	13,9	14,8	16,3
Ensino superior (%)	16,9	10,7	5,1
Tempo de emprego médio (meses)	61,3	59,1	52,5

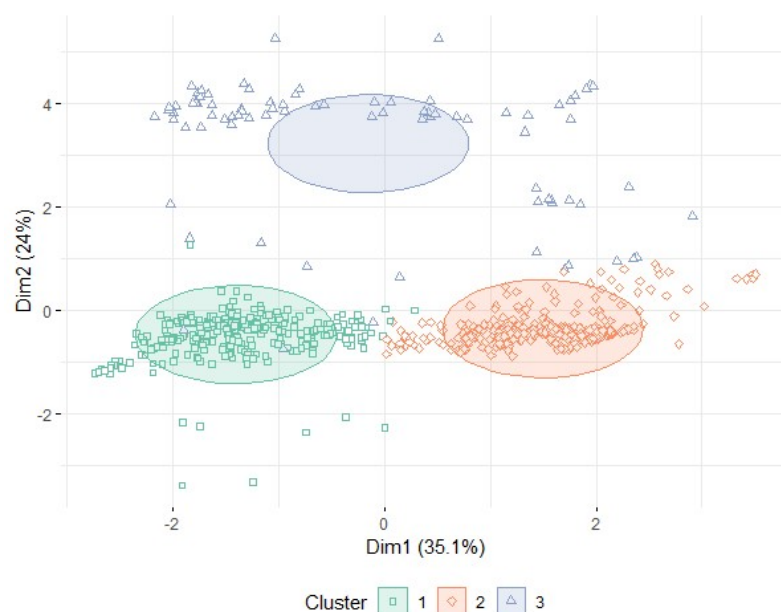
Fonte: elaboração própria baseada nos dados da Relação Anual de Informações Sociais de 2021.

Os valores observados sugerem que ocupações com perfil misto de atitudes tendem a estar associadas a salários-hora ligeiramente mais elevados, do que aquelas intensivas em apenas um tipo de atitude, seja autônoma (R\$14,4) ou sob supervisão (R\$11,3), indicando uma possível vantagem para ocupações com maior equilíbrio entre os tipos de atitude. Dada as estatísticas analisadas na Tabela 2, isso pode ser um reflexo de uma diferença educacional, expressa por um maior percentual de pessoas com ensino superior, e pelo fato de as ocupações mistas possuírem um maior tempo médio de emprego na empresa, o que pode indicar que nessas ocupações há maior formação de capital humano específico.

#### 4.2. Clusters Ocupacionais – Conhecimento II

Aplicou-se também a metodologia *Fuzzy* às ocupações classificadas como de Conhecimento II, que, conforme o QBQ, exigem formação de nível superior. A análise do tamanho médio das silhuetas indicou que a divisão em três agrupamentos é a que melhor maximiza a qualidade da classificação para essa categoria. A distribuição dos clusters pode ser observada no gráfico de pontos da Figura 3, no qual as ocupações são organizadas conforme a maior probabilidade de pertencimento. As variáveis utilizadas explicam 59,1% da variância, e a visualização evidencia a proximidade de algumas ocupações com grupos vizinhos, o que reforça a adequação da abordagem *fuzzy* frente a classificações rígidas.

**Figura 3** – Nuvem de pontos para as ocupações de conhecimento 2



Fonte: elaboração própria baseada nos dados do Quadro Brasileiro de Qualificações.

Da mesma forma, os graus de pertencimento das ocupações permitiram a identificação de três grupos, descritos na Tabela 3 com base nos indicadores utilizados na formação dos *clusters*. Diferentemente dos agrupamentos anteriores, neste caso, a área de conhecimento se destaca como o principal critério de diferenciação entre os *clusters*. O primeiro grupo reúne ocupações com forte predominância de conhecimentos da área *Shape* (78,4%), enquanto o segundo é intensivo em conhecimentos da área *STEM* (74,9%). Já o terceiro apresenta uma composição mais equilibrada entre as duas dimensões, o que sugere que essas ocupações demandam uma abordagem mais multidisciplinar. Com base nesse padrão, os grupos foram denominados, respectivamente, como “intensivo na área *Shape*”, “intensivo na área *STEM*” e “misto em conhecimentos”.

**Tabela 3** – Proporção (%) das variáveis do Quadro Brasileiro de Qualificações (QBQ) observadas em cada cluster de Conhecimento II

Tipologia	Variáveis	Clusters de Conhecimento II		
		Intensivo na área <i>Shape</i>	Intensivo na área <i>STEM</i>	Misto
Conhecimento II	<i>STEM</i>	15,3	74,9	39,8
	<i>Shape</i>	78,4	17,8	52,4
	Misto	6,3	7,3	7,9
Habilidades	Práticas	38,8	31,9	30,3
	Cognitivas	45,2	40,5	46,7
	Física	16,0	27,6	23,0
Atitudes	Sob supervisão	0	0	0
	Autônoma	25,3	26,2	45,9
	De supervisão	25,8	26,2	4,8
	De avaliação	48,9	47,7	49,3

Fonte: elaboração própria baseada nos dados do Quadro Brasileiro de Qualificações.

Em relação às demais variáveis, nota-se que os percentuais associados às habilidades e atitudes apresentam certa semelhança entre os grupos. Em todos os casos, as habilidades cognitivas predominam, compreendendo capacidades como comunicação, raciocínio e trabalho com números. As habilidades práticas, por sua vez, ocupam a segunda posição em relevância e dizem respeito a competências operacionais, organizacionais e interpessoais. Essa categoria se destaca especialmente nas ocupações do grupo intensivo em conhecimento *Shape*, com 38,8%. Já as habilidades físicas, que englobam as capacidades físicas, psicomotoras e sensoriais, exercem influência menor, especialmente nesse agrupamento, porém relevante, na composição de todos os grupos.

No que se refere às atitudes, observa-se que aquelas sob supervisão não contribuem para a definição dos clusters. Em todos eles, predominam as atitudes de avaliação, seguidas por proporções relativamente equilibradas entre atitudes autônomas e de supervisão. A principal distinção está na maior presença de atitudes autônomas no cluster misto, que atinge 45,9%.

A Tabela 4 apresenta as médias ponderadas pelas probabilidades de pertencimento aos *clusters* em relação ao salário-hora, à composição por gênero, faixa etária, escolaridade e ao tempo médio de emprego a partir dos dados da RAIS. Os dados indicam que tanto o

rendimento por hora quanto a distribuição etária são relativamente semelhantes entre os grupos. O salário-hora varia em torno de R\$ 45,00, um valor elevado, mas coerente com o fato de se tratar de ocupações que exigem formação superior. Além disso, observa-se que mais de 65% dos trabalhadores estão na faixa etária entre 30 e 50 anos.

**Tabela 4** – Informações médias dos vínculos da RAIS em Minas Gerais (2021) por *cluster* de Conhecimento II

Variáveis	Intensivo na área <i>Shape</i>	Intensivo na área <i>STEM</i>	Misto
Salário-hora (R\$)	45,1	44,6	45,6
Mulheres (%)	61,4	56,4	72,0
18 a 29 anos (%)	15,0	18,1	10,3
30 a 50 anos (%)	65,9	67,8	65,6
51 anos ou mais (%)	19,1	14,1	27,1
Ensino superior (%)	69,2	84,3	83,4
Tempo de emprego médio (meses)	102,1	77,7	118,1

Fonte: elaboração própria baseada nos dados da Relação Anual de Informações Sociais de 2021.

A composição de gênero e o tempo médio de emprego seguem uma tendência semelhante, com os menores valores observados no grupo mais concentrado em áreas *STEM* (ciência, tecnologia, engenharia e matemática). Esses resultados corroboram estudos prévios (Casad et al., 2019; Jiang, 2021), que destacam a baixa participação feminina em setores como tecnologia, engenharia e ciências. De acordo com um estudo do Banco Mundial, utilizando dados do LinkedIn<sup>3</sup>, as mulheres representam apenas 29,2% do total de empregados nessas áreas. Jiang (2021), por exemplo, encontra que as mulheres formadas em graduações da área *STEM* não atuam nesse tipo de carreira por consequências de suas escolhas educacionais e laborais, sendo mais representativas em cursos menos intensivos em matemática e que tendem a ser reconhecidos como *Shape* no mercado de trabalho, bem como são mais penalizadas por decisões relacionadas ao equilíbrio entre a vida familiar e profissional. Já Casad *et al.* (2019) argumentam que essa sub-representação é consequência de estereótipos, falta de rede de apoio e um ambiente acadêmico menos receptivo.

<sup>3</sup> Informações constantes no *Global Gender Report* de 2023, do Banco Mundial. Acesso em: <https://www.weforum.org/publications/global-gender-gap-report-2023/in-full/gender-gaps-in-the-workforce/>

Além disso, o menor tempo de emprego pode sugerir que a demanda por ocupações desse tipo é maior que a disponibilidade de mão de obra nesse grupo de ocupações, especialmente no que tange às tecnologias digitais. De acordo com um relatório do Banco Mundial<sup>4</sup>, o setor de tecnologia da informação e comunicação observou um crescimento global massivo em sua empregabilidade no mundo, variando de 37 milhões de pessoas em 2000 para 68 milhões em 2022. Esse resultado ilustra uma situação de geração de oportunidades dentro de uma carreira *STEM*, que garante o surgimento de mais opções de emprego. Uma consequência da ampliação da oferta de trabalho na área é o aumento da mobilidade de trabalhadores entre os postos de trabalho, que resulta nos registros de tempo de emprego nas informações administrativas da firma.

Já em relação à educação superior, o *cluster* intensivo na área *Shape* possui o menor valor de pessoas com uma graduação completa, destoando dos resultados dos outros 2 grupos. Ainda que as ocupações do grupo de Conhecimento II do Quadro Brasileiro de Qualificações em tese requerem ensino superior, o percentual mais baixo pode significar uma certa flexibilização dessas ocupações no mercado de trabalho no que diz respeito à formação exigida para a atuação.

#### 4.3. Resultados econométricos

Esta seção descreve uma síntese dos resultados<sup>5</sup> das decomposições propostas na etapa metodológica, com o objetivo de verificar como se comportam os diferenciais por gênero em cada agrupamento ocupacional sugerido. De modo geral, é encontrado um *gap* salarial que beneficia os homens, resultado que corrobora com aqueles encontrados na literatura, a qual destaca o fato de a alocação ocupacional ser fruto de um processo discriminatório (Stier; Yaish, 2014), e de existir uma penalização salarial quando um trabalhador transita de uma ocupação majoritariamente masculina para uma feminina (Silva;

---

<sup>4</sup> Informações constantes no *Digital Progress and Trend Report* de 2023, do Banco Mundial. Acesso em: <https://openknowledge.worldbank.org/server/api/core/bitstreams/95fe55e9-f110-4ba8-933f-e65572e05395/content>

<sup>5</sup> Resultados completos estão disponíveis no Apêndices A.

Vaz, 2022).

Seguindo a mesma lógica adotada na construção da tipologia, as decomposições também são realizadas separadamente para os clusters de ocupações de Conhecimento I e II. Na Tabela 5, são apresentados os resultados da decomposição de Oaxaca-Blinder para as ocupações de Conhecimento I, considerando as probabilidades de pertencimento a cada um dos três *clusters* como pesos na análise. Esses valores são atribuídos aos indivíduos conforme sua ocupação, de forma que a contribuição de cada trabalhador na análise reflita o grau de pertencimento ao grupo. Em todos os casos, observa-se uma diferença salarial que favorece os homens, com o componente não explicado sendo consistentemente maior do que o explicado e apresentando valores relativamente próximos entre os grupos. Isso implica que, nas três situações analisadas, as desigualdades de gênero são mais explicadas por fatores não observáveis, frequentemente associados à discriminação contra as mulheres.

**Tabela 5** – Resultados da decomposição de Oaxaca-Blinder para os *clusters* do Conhecimento I

Logaritmo do salário-hora	Misto em atitudes	Intensivo em atitudes autônomas	Intensivo em atitudes sob supervisão
Homens	2,6048*** (0,00)	2,4982*** (0,00)	2,3072*** (0,00)
Mulheres	2,3796*** (0,00)	2,3517*** (0,00)	2,1681*** (0,00)
Diferença	0,2252*** (0,00)	0,1465*** (0,00)	0,1391*** (0,00)
Componente explicado	0,0020*** (0,00)	-0,0565*** (0,00)	-0,0526*** (0,00)
Componente não explicado	0,2232*** (0,00)	0,2031*** (0,00)	0,1917*** (0,00)

Nota: \* p < 0,1; \*\* p < 0,05; \*\*\* p < 0,01. Fonte: elaboração própria baseada nos dados do Quadro Brasileiro de Qualificações.

Embora os resultados indiquem um padrão consistente de desigualdade em desfavor das mulheres, há uma diferença relevante entre os grupos quanto ao sinal do componente explicado. No caso do *cluster* com competências mistas em atitudes, esse componente assume valor positivo, o que sinaliza uma maior valorização das características produtivas

associadas aos homens nessa classificação — contribuindo para um *gap* salarial de 22,5%. As ocupações com maior probabilidade de pertencimento a esse grupo incluem: supervisor de curtimento, técnico mecânico na manutenção de ferramentas, técnico em petroquímica, mestre de galvanoplastia, cenotécnico e mestre de pintura. Essas funções podem ser consideradas tipicamente masculinas, o que pode indicar uma possível origem dessa diferença, mesmo em um grupo cuja composição por gênero é relativamente equitativa.

A Tabela 6 apresenta os resultados da decomposição de Oaxaca-Blinder para os *clusters* de ocupações classificados como Conhecimento II, utilizando as probabilidades de pertencimento a cada cluster como pesos na análise. Os resultados indicam, de forma geral, uma vantagem salarial para os homens, já que o componente não explicado supera o explicado em todos os casos, sugerindo a presença de discriminação de gênero. No entanto, essa desvalorização do trabalho feminino ocorre com diferentes intensidades entre os grupos. Nos *clusters* compostos por ocupações intensivas em conhecimento do tipo *Shape* e naquelas com conhecimentos mistos, observa-se uma valorização relativamente maior das características produtivas das mulheres, em comparação ao grupo *STEM*, resultando em diferenciais salariais menores, de 5% e 2,7%, respectivamente, a favor dos homens. Já no *cluster STEM*, essa diferença se acentua de forma mais expressiva, com um ganho salarial de 20,4% favorável aos homens, evidenciando um cenário mais marcado de desigualdade de gênero nesse grupo.



**Tabela 6** – Resultados da decomposição de Oaxaca-Blinder para os *clusters* do conhecimento 2

Logaritmo do salário-hora	Intensivo na área <i>Shape</i>	Intensivo na área <i>STEM</i>	Misto
Homens	3,4167*** (0,00)	3,5796*** (0,00)	3,5134*** (0,00)
Mulheres	3,3643*** (0,00)	3,3755*** (0,00)	3,4862*** (0,00)
Diferença	0,0524*** (0,00)	0,2041*** (0,00)	0,0271*** (0,00)
Componente explicado	-0,1620*** (0,00)	-0,0352*** (0,00)	-0,2020*** (0,00)
Componente não explicado	0,2145*** (0,00)	0,2393*** (0,00)	0,2291*** (0,00)

Nota: \*  $p < 0,1$ ; \*\*  $p < 0,05$ ; \*\*\*  $p < 0,01$ . Fonte: elaboração própria baseada nos dados do Quadro Brasileiro de Qualificações.

Esse resultado pode ser atribuído à existência de barreiras à inserção feminina em ocupações que demandam conhecimentos específicos da área *STEM*, conforme demonstrado pela menor presença de mulheres nesse tipo de atividade, como discutido na seção descritiva. Além dessas barreiras de entrada, as mulheres que conseguem acessar essas ocupações continuam enfrentando penalizações salariais. Evidência nesse sentido é apresentada por Zajac et al. (2023), que, a partir de dados do mercado de trabalho da Polônia, mostram que mulheres em carreiras *STEM* recebem, em média, 20% a menos do que os homens no início de suas carreiras, com um aumento progressivo dessa diferença ao longo do tempo. Esses achados reforçam que a discriminação de gênero se manifesta tanto no acesso quanto na progressão do trabalho feminino em áreas com essas exigências de qualificação.

## 5. CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo desenvolver classificações ocupacionais com base nas informações de conhecimentos, habilidades e atitudes do Quadro Brasileiro de Qualificações, para criar uma ferramenta a partir das competências exigidas pelas

ocupações. Para alcançá-lo, foi utilizada a técnica de *Fuzzy Cluster*, que, ao contrário de outras técnicas de agrupamento, propõe o cálculo de uma medida de pertencimento das ocupações a um ou mais grupos, permitindo assim considerar a natureza específica desses dados ocupacionais, nos quais as competências são, em certa medida, transferíveis e, portanto, possuem um grau de proximidade.

O estudo foi realizado separadamente para as ocupações incluídas nas categorias de Conhecimento I e II, que, por definição, apresentam níveis de qualificação distintos e bem estabelecidos. Os resultados apontaram que três agrupamentos maximizam a qualidade das classificações em ambas as categorias. Na primeira, os três grupos se diferenciam de forma expressiva pelos tipos de atitudes exigidas, sendo denominados respectivamente como mistos em atitudes, intensivos em atitudes autônomas e intensivos em atitudes sob supervisão. No caso das ocupações de Conhecimento II, os três grupos são definidos em termos das áreas de conhecimento, sendo chamados de intensivos na área *Shape*, intensivos na área *STEM* e mistos em conhecimento.

Para além disso, foi realizado um exercício de decomposição para investigar como se comportam os diferenciais por gênero entre esses grupos. Através da metodologia de Oaxaca-Blinder, foi identificado uma vantagem salarial para os homens em detrimento às mulheres devido ao fator de discriminação, conforme esperado pela literatura empírica da área. As particularidades verificadas nesta análise correspondem aos resultados do cluster misto em atitude, do Conhecimento I, e intensivo na área *STEM*, no caso de Conhecimento II, em que as características produtivas dos homens parecem ser mais valorizadas.

Em análises futuras, pretende-se examinar como as tipologias sugeridas refletem a segmentação ocupacional por gênero, considerando que os resultados da decomposição indicam um favorecimento da mão de obra masculina em determinadas ocupações. Além disso, outros temas relacionados ao mercado de trabalho também podem ser explorados a partir dos agrupamentos propostos neste estudo, como a análise de *mismatch* ocupacional, investigando a relação entre a formação e a qualificação dos trabalhadores e o grupo ao qual pertencem as ocupações em que atuam, assim como estudos sobre a mobilidade entre esses

postos de trabalho e o impacto desses movimentos intra e inter-grupos.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Autor, D. H., e Handel, M. J. (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of labor Economics*, 31(S1), S59–S96.
- Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436–455.
- Cardoso, B.-H., Kaltenberg, M. e Hartmann, D. (2023). How skills-related local labor markets shape occupational upward mobility prospects. SSRN.
- Casad, B. J., Franks, J. E., Garasky, C. E., Kittleman, M. M., Roesler, A. C., Hall, D. Y. e Petzel, Z. W. (2021). Gender inequality in academia: Problems and solutions for women faculty in stem. *Journal of neuroscience research*, 99(1), 13–23.
- Cortes, G. M., e Gallipoli, G. (2018). The costs of occupational mobility: An aggregate analysis. *Journal of the European Economic Association*, 16(2), 275–315.
- Everitt, B. S., Landau, S., Leese, M. e Stahl, D. (2011). *Cluster analysis* (5th ed.). Chichester: J. Wiley Sons.
- Gathmann, C., e Schönberg, U. (2010). How general is human capital? a task-based approach. *Journal of Labor Economics*, 28(1), 1–49.
- Jiang, X. (2021). Women in stem: Ability, preference, and value. *Labour Economics*, 70, 101991.
- Junior, J. S., e Funchal, B. (2016). Mudanças na demanda por diferentes tipos de atividades no mercado de trabalho brasileiro entre 1985 e 2002 (Nota Técnica). Instituto de Pesquisa Econômica Aplicada (Ipea).
- Kaufman, L., e Rousseeuw, P. J. (1990). *Finding groups in data: an introduction to cluster analysis*. New York: John Wiley & Sons.
- Kaufman, L., e Rousseeuw, P. J. (2005). *Finding groups in data: an introduction to cluster analysis*. Hoboken: John Wiley & Sons.
- Koomen, M., e Backes-Gellner, U. (2022). Occupational tasks and wage inequality in west

- germany: A decomposition analysis. *Labour Economics*, 79, 102284.
- Lordan, G., e Pischke, J.-S. (2022). Does rosie like riveting? male and female occupational choices. *Economica*, 89(353), 110–130.
- Monsueto, S. E., Cardoso, F. P. e Oliveira, A. A. A. d. (2024). Um Índice de similaridade entre ocupações da cbo (Texto para Discussão No. 101). Goiânia: FACE/UFG.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693–709.
- Reis, M. C. (2016). Uma descrição das ocupações no brasil a partir de informações sobre as atividades normalmente desempenhadas pelos trabalhadores.
- Reis, M. C. (2018). Measuring the mismatch between field of study and occupation using a task-based approach. *Journal for labour market research*, 52, 1–15.
- Silva, Y. G. d., e Vaz, D. V. (2022). Por que as ocupações “femininas pagam menos? um estudo longitudinal. *Revista Brasileira de Estudos de População*, 39, e0212.
- Simões, R. F. (2005). Métodos de análise regional e urbana: diagnóstico aplicado ao planejamento (Texto para Discussão No. 259). Belo Horizonte: Cedeplar/UFMG.
- Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of labor economics*, 24(2), 235–270.
- Stier, H., e Yaish, M. (2014). Occupational segregation and gender inequality in job quality: a multi-level approach. *Work, employment and society*, 28(2), 225–246.
- Suleman, A., Suleman, F. e Reis, E. (2016). Fuzzy approach to discrete data reduction: an application in economics for assessing the skill premium. *Journal of Business Economics and Management*, 17(3), 414–429.
- Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics*, 30(1), 1–53.
- Zajac, T., Magda, I., Bozykowski, M., Chłton-Dominczak, A. e Jasinski, M. (2024). Gender pay gaps across stem fields of study. *Studies in Higher Education*, 1–14.

#### **Apêndice A – Resultados expandidos da decomposição de Oaxaca-Blinder**

Tabela A1: Resultados da parte explicada da decomposição de Oaxaca-Blinder para os clusters

Variáveis	Misto em atitudes	Intensivo em atitudes autônomas	Intensivo em atitudes sob-supervisão	Intensivo na área <i>Shape</i>	Intensivo na área STEM	Misto
18 a 29 anos	0,0065*** (0,00)	0,0031*** (0,00)	-0,0002*** (0,00)	-0,0006** (0,00)	0,0064*** (0,00)	-0,0132*** (0,00)
30 a 50 anos	0,0006*** (0,00)	0,0002*** (0,00)	-0,0003*** (0,00)	-0,0000 (0,00)	0,0034*** (0,00)	0,0013*** (0,00)
Fundamental completo	0,0023*** (0,00)	0,0037*** (0,00)	0,0015*** (0,00)	0,0079*** (0,00)	0,0019*** (0,00)	0,0025*** (0,00)
Médio completo	-0,0096*** (0,00)	-0,0095*** (0,00)	-0,0106*** (0,00)	0,0433*** (0,00)	0,0333*** (0,00)	0,0339*** (0,00)
Superior completo	-0,0480*** (0,00)	-0,0582*** (0,00)	-0,0303*** (0,00)	-0,1303*** (0,00)	-0,1274*** (0,00)	-0,1017*** (0,00)
Pequeno-médio porte	-0,0013*** (0,00)	0,0001*** (0,00)	-0,0002*** (0,00)	0,0033*** (0,00)	0,0007*** (0,00)	0,0046*** (0,00)
Médio porte	0,0011*** (0,00)	0,0045*** (0,00)	0,0019*** (0,00)	0,0153*** (0,00)	0,0103*** (0,00)	0,0214*** (0,00)
Grande-médio porte	0,0128*** (0,00)	0,0109*** (0,00)	0,0051*** (0,00)	0,0179*** (0,00)	0,0278*** (0,00)	0,0297*** (0,00)
Grande porte	-0,0002 (0,00)	-0,0200*** (0,00)	-0,0096*** (0,00)	-0,0730*** (0,00)	-0,0793*** (0,00)	-0,1818*** (0,00)
Setor industrial	0,0188*** (0,00)	0,0145*** (0,00)	0,0058*** (0,00)	0,0046*** (0,00)	0,0086*** (0,00)	0,0045*** (0,00)
Setor de civil construção	0,0142*** (0,00)	0,0108*** (0,00)	0,0104*** (0,00)	0,0002*** (0,00)	0,0026*** (0,00)	0,0004*** (0,00)
Setor de serviços distributivos	0,0045*** (0,00)	0,0153*** (0,00)	0,0062*** (0,00)	-0,0001 (0,00)	-0,0007*** (0,00)	-0,0003*** (0,00)
Setor de serviços produtivos	0,0011*** (0,00)	-0,0004*** (0,00)	-0,0018*** (0,00)	0,0063*** (0,00)	0,0043*** (0,00)	0,0091*** (0,00)
Setor de serviços sociais	-0,0011*** (0,00)	-0,0126*** (0,00)	-0,0035*** (0,00)	0,0023*** (0,00)	0,0314*** (0,00)	-0,0002*** (0,00)
Setor de serviços pessoais	0,0007*** (0,00)	0,0006*** (0,00)	0,0001 (0,00)	-0,0000 (0,00)	0,0017*** (0,00)	-0,0073*** (0,00)
Administração pública	-0,0083*** (0,00)	-0,0140*** (0,00)	-0,0122*** (0,00)	-0,0136*** (0,00)	0,0344*** (0,00)	0,0536*** (0,00)
Outros setores	-0,0000 (0,00)	-0,0000*** (0,00)	-0,0000 (0,00)	-0,0001*** (0,00)	-0,0000* (0,00)	-0,0002*** (0,00)
Setor agrícola	0,0019*** (0,00)	0,0004*** (0,00)	0,0009*** (0,00)	0,0015*** (0,00)	-0,0002 (0,00)	0,0023*** (0,00)
Tempo de emprego	0,0090*** (0,00)	-0,0063*** (0,00)	-0,0195*** (0,00)	-0,0705*** (0,00)	0,0122*** (0,00)	-0,0875*** (0,00)
Tempo de emprego ao quadrado	-0,0029*** (0,00)	0,0004*** (0,00)	0,0039*** (0,00)	0,0234*** (0,00)	-0,0065*** (0,00)	0,0269*** (0,00)
Observações	3.961.983	3.961.983	3.961.983	684.171	684.171	684.171

Nota: \* p < 0,1; \*\* p < 0,05; \*\*\* p < 0,01. Erros-padrão entre parênteses. Fonte: elaboração própria baseada nos dados da RAIS de 2021.

Tabela A2: Resultados da parte não-explicada da decomposição de Oaxaca-Blinder para os clusters

Variáveis	Misto em atitudes	Intensivo em atitudes autônomas	Intensivo em atitudes sob-supervisão	Intensivo na área <i>Shape</i>	Intensivo na área <i>STEM</i>	Misto
18 a 29 anos	-0,0475*** (0,00)	-0,0321*** (0,00)	-0,0343*** (0,00)	-0,0262*** (0,00)	-0,0278*** (0,00)	-0,0145*** (0,00)
30 a 50 anos	0,0051*** (0,00)	0,0092*** (0,00)	0,0080*** (0,00)	-0,0523*** (0,00)	-0,0045 (0,01)	-0,0287*** (0,00)
Fundamental completo	0,0028*** (0,00)	0,0069*** (0,00)	0,0044*** (0,00)	-0,0054*** (0,00)	-0,0003 (0,00)	-0,0002 (0,00)
Médio completo	0,0282*** (0,00)	0,0105*** (0,00)	0,0135*** (0,00)	-0,0005 (0,00)	-0,0061** (0,00)	-0,0004 (0,00)
Superior completo	0,0115*** (0,00)	0,0062*** (0,00)	0,0063*** (0,00)	0,0417*** (0,01)	-0,0315 (0,02)	0,1295*** (0,02)
Pequeno-médio porte	0,0012*** (0,00)	0,0004** (0,00)	0,0019*** (0,00)	0,0001 (0,00)	0,0010 (0,00)	-0,0017*** (0,00)
Médio porte	0,0113*** (0,00)	0,0069*** (0,00)	0,0095*** (0,00)	0,0064*** (0,00)	0,0140*** (0,00)	-0,0000 (0,00)
Grande-médio porte	0,0241*** (0,00)	0,0174*** (0,00)	0,0182*** (0,00)	0,0164*** (0,00)	0,0319*** (0,00)	0,0007 (0,00)
Grande porte	0,0715*** (0,00)	0,0468*** (0,00)	0,0482*** (0,00)	0,0486*** (0,00)	0,0890*** (0,01)	0,0284*** (0,01)
Setor industrial	0,0132*** (0,00)	0,0154*** (0,00)	0,0208*** (0,00)	0,0052*** (0,00)	0,0201*** (0,00)	0,0023*** (0,00)
Setor de construção civil	0,0012*** (0,00)	-0,0004*** (0,00)	0,0026*** (0,00)	0,0010*** (0,00)	0,0062*** (0,00)	0,0012*** (0,00)
Setor de serviços distributivos	0,0009*** (0,00)	0,0022*** (0,00)	0,0031*** (0,00)	0,0015*** (0,00)	0,0032*** (0,00)	0,0013*** (0,00)
Setor de serviços produtivos	0,0109*** (0,00)	0,0059*** (0,00)	0,0088*** (0,00)	0,0021* (0,00)	0,0277*** (0,00)	0,0011 (0,00)
Setor de serviços sociais	-0,0069*** (0,00)	-0,0102*** (0,00)	-0,0029*** (0,00)	-0,0057*** (0,00)	0,0428*** (0,00)	-0,0095*** (0,00)
Setor de serviços pessoais	-0,0015*** (0,00)	-0,0019*** (0,00)	0,0009*** (0,00)	-0,0019*** (0,00)	0,0061*** (0,00)	-0,0055*** (0,00)
Administração pública	-0,0289*** (0,00)	-0,0181*** (0,00)	-0,0170*** (0,00)	-0,0360*** (0,00)	0,0488*** (0,00)	-0,0387*** (0,00)
Outros setores	-0,0000 (0,00)	-0,0000*** (0,00)	-0,0000 (0,00)	-0,0004*** (0,00)	0,0001** (0,00)	-0,0001*** (0,00)
Setor agrícola	0,0015*** (0,00)	0,0010*** (0,00)	0,0036*** (0,00)	0,0015*** (0,00)	0,0030*** (0,00)	0,0013*** (0,00)
Tempo de emprego	0,0068*** (0,00)	0,0139*** (0,00)	0,0278*** (0,00)	-0,0967*** (0,01)	-0,0417*** (0,01)	0,0478*** (0,01)
Tempo de emprego ao quadrado	-0,0070*** (0,00)	-0,0106*** (0,00)	-0,0170*** (0,00)	0,0484*** (0,00)	0,0057* (0,00)	0,0063 (0,01)
Constante	0,1247*** (0,00)	0,1339*** (0,00)	0,0853*** (0,00)	0,2668*** (0,02)	0,0518** (0,03)	0,1086*** (0,03)
Observações	3.961.983	3.961.983	3.961.983	684.171	684.171	684.171

Nota: \* p < 0,1; \*\* p < 0,05; \*\*\* p < 0,01. Erros-padrão entre parênteses. Fonte: elaboração própria baseada nos dados da RAIS de 2021.