

ESCORES OCUPACIONAIS BASEADOS EM TAREFAS: UMA HARMONIZAÇÃO PARA CBO E COD

Nicolas Gobetti Theodorovicz¹

Juan Pablo Gaio Cubilla²

Guilherme Marques Moura³

RESUMO

Apesar da difusão da abordagem baseada em tarefas, aplicações brasileiras ainda enfrentam dificuldades para combinar medidas atuais de conteúdo ocupacional com a Classificação Brasileira de Ocupações de 2002 (CBO 2002) e a Classificação de Ocupações para Pesquisas Domiciliares (COD) em um mesmo quadro documentado. Este artigo enfrenta esse problema ao construir bases de escores de tarefas a partir de 16 itens da O*NET 30.1, selecionados conforme a operacionalização de Acemoglu e Autor (2011). Os itens são organizados em cinco dimensões, sendo cognitiva analítica, cognitiva interpessoal, cognitiva rotineira, manual rotineira e manual não rotineira, e transportados por cadeias explícitas de correspondências ocupacionais. A CBO harmonizada contém 1.728 ocupações e a CBO completa, 2.694, após 966 imputações hierárquicas. A COD harmonizada contém 415 dos 437 códigos classificáveis da lista oficial; a COD completa acrescenta 22 códigos, oito por doadores selecionados por proximidade de conteúdo e 14 por medianas hierárquicas. A validação convergente em 414 ocupações ISCO-08, contra escores independentes derivados de descrições de tarefas, produz correlações de Spearman entre 0,444 e 0,717. No teste por blocos da CBO, a concordância da dimensão predominante cai de 87,9%, quando a ocupação-alvo é estimada com as demais ocupações de sua família, para 56,4%, quando restam apenas doadores do mesmo grande grupo. O resultado são duas bases compatíveis com as classificações brasileiras, acompanhadas de regras explícitas de construção, validação e uso.

Palavras-chave: Escores ocupacionais. Tarefas ocupacionais. Crosswalk. O*NET.

¹ Aluno do 3º período do curso de Ciências Econômicas da FAE Centro Universitário. Voluntário do Programa de Apoio à Iniciação Científica (PAIC 2024-2025). *E-mail:* nicolas.theodorovicz@mail.fae.edu

² Aluno do 7º período do curso de Direito da FAE Centro Universitário. Bolsista do Programa de Apoio à Iniciação Científica (PAIC 2024-2025). *E-mail:* juan.cubilla@mail.fae.edu

³ Orientador da Pesquisa. Doutor em Desenvolvimento Econômico (UFPR). Professor da FAE Centro Universitário. *E-mail:* guilherme.moura@fae.edu

INTRODUÇÃO

A mudança tecnológica tem alterado a composição do emprego e a demanda por diferentes tipos de trabalho. Uma das principais contribuições da literatura econômica recente foi mostrar que esses efeitos não dependem apenas do nível de qualificação dos trabalhadores, mas também do tipo de tarefa desempenhada em cada ocupação. Nesse sentido, ocupações com maior intensidade de tarefas rotineiras tendem a ser mais suscetíveis à automação, enquanto ocupações baseadas em resolução de problemas, interação interpessoal, criatividade ou adaptação a contextos variados apresentam maior conteúdo não rotineiro.

A abordagem baseada em tarefas, consolidada a partir de Autor, Levy e Murnane (2003) e sistematizada por Acemoglu e Autor (2011), permitiu transformar descrições qualitativas das ocupações em medidas quantitativas de conteúdo do trabalho. Essa literatura possibilita analisar a relação entre tecnologia, emprego, salários e composição ocupacional por meio de escores que distinguem dimensões cognitivas, manuais, rotineiras e não rotineiras.

No caso brasileiro, a aplicação direta dessa abordagem encontra uma limitação importante. Bases como o Dictionary of Occupational Titles e a O*NET são estruturadas a partir de classificações ocupacionais dos Estados Unidos e não possuem correspondência direta com as taxonomias brasileiras. As bases nacionais utilizam classificações próprias, como a Classificação Brasileira de Ocupações de 2002 (CBO 2002) e a Classificação de Ocupações para Pesquisas Domiciliares (COD). O uso de escores norte-americanos exige, portanto, um procedimento prévio de harmonização entre classificações ocupacionais distintas.

Diante desse problema, este artigo tem como objetivo construir, validar e documentar bases de escores ocupacionais baseados em tarefas para a CBO 2002 e a COD. Parte-se de informações da O*NET organizadas segundo cinco categorias da literatura de mudança tecnológica baseada em tarefas e transportadas por cadeias de correspondências ocupacionais auditáveis. A contribuição específica é oferecer uma harmonização atualizada e simultânea para duas taxonomias brasileiras, com regras explícitas de ponderação e imputação, diagnóstico de cobertura, testes de sensibilidade e validação convergente em nível ocupacional.

Além desta introdução, o artigo está dividido em mais quatro seções. A segunda apresenta o referencial teórico sobre tarefas, tecnologia e escores ocupacionais. A terceira descreve a metodologia e as bases de dados utilizadas. A quarta apresenta a validação técnica e as características das bases finais. Por fim, a quinta seção reúne as considerações finais.

1 FUNDAMENTAÇÃO TEÓRICA

A literatura sobre mudança tecnológica e mercado de trabalho passou por uma inflexão importante ao deslocar o foco exclusivo de grupos educacionais ou níveis agregados de qualificação para o conteúdo efetivo das tarefas realizadas em cada ocupação. Nesse contexto, Autor, Levy e Murnane (2003) desenvolveram uma abordagem baseada em tarefas para investigar como a difusão de computadores altera a demanda por diferentes tipos de trabalho. Em vez de tratar a tecnologia como substituta homogênea de trabalhadores pouco qualificados e complementar homogênea de trabalhadores qualificados, os autores argumentam que seu efeito depende da natureza das tarefas desempenhadas.

A contribuição central de Autor, Levy e Murnane (2003) consiste em distinguir tarefas rotineiras e não rotineiras. Tarefas rotineiras são aquelas descritas por regras explícitas, procedimentos repetitivos ou padrões codificáveis, tornando-se mais suscetíveis à substituição por capital computacional. Já tarefas não rotineiras dependem de resolução de problemas, interação interpessoal, criatividade, julgamento ou habilidades manuais pouco padronizáveis, sendo menos diretamente substituíveis pela automação.

Essa abordagem foi posteriormente sistematizada por Acemoglu e Autor (2011), que destacam a importância de compreender o mercado de trabalho a partir da interação entre habilidades, tarefas e tecnologias. Para os autores, trabalhadores não são remunerados apenas por escolaridade ou qualificação abstrata, mas também pelas tarefas que executam e pela forma como essas tarefas interagem com novas tecnologias. Assim, a abordagem baseada em tarefas permite analisar com maior precisão os efeitos heterogêneos da mudança tecnológica sobre emprego, salários e composição ocupacional.

A partir dessa estrutura, Autor, Levy e Murnane (2003) propõem uma mensuração empírica do conteúdo das ocupações com base no Dictionary of Occupational Titles (DOT), classificando ocupações segundo dimensões cognitivas, manuais, rotineiras e não rotineiras.

Essa estratégia permitiu transformar características qualitativas do trabalho em medidas quantitativas comparáveis entre ocupações. No presente artigo, essa tradição é utilizada como referência conceitual para a construção de escores ocupacionais aplicados ao caso brasileiro, aproximando a literatura internacional baseada em tarefas das classificações ocupacionais utilizadas no país.

A hipótese de mudança tecnológica enviesada contra tarefas rotineiras encontrou desdobramentos na literatura de polarização ocupacional. Autor e Dorn (2013) associam

a queda do custo de automatizar tarefas codificáveis ao deslocamento de trabalhadores para serviços manuais não rotineiros e ao crescimento relativo das extremidades da distribuição ocupacional. Em evidência para 16 países europeus, Goos, Manning e Salomons (2014) mostram que a tecnologia enviesada contra a rotina e a possibilidade de deslocamento internacional de tarefas ajudam a explicar a polarização do emprego. Esses resultados reforçam que escolaridade e salário, isoladamente, não descrevem o mecanismo tecnológico: é necessário mensurar o conteúdo de tarefas das ocupações.

A mensuração, contudo, não é automática. O O*NET oferece muitas escalas potencialmente sobrepostas, e a escolha dos itens determina a validade de construto. Acemoglu e Autor (2011) enfrentam esse problema com um conjunto parcimonioso de atividades, habilidades e contextos de trabalho para operacionalizar cinco dimensões. Em estratégia independente da O*NET, Mihaylov e Tijdens (2019) classificam 3.264 descrições de tarefas da ISCO-08 e produzem escores para 427 ocupações a quatro dígitos. A existência dessa medida alternativa permite testar se diferentes fontes e procedimentos recuperam ordenações ocupacionais semelhantes.

No Brasil, Maciente (2014) compatibiliza atributos da O*NET com a CBO e sintetiza um amplo conjunto de variáveis em dimensões de competências, habilidades e conhecimentos, demonstrando o potencial analítico da transferência de informação ocupacional. Em perspectiva comparativa, Mahutga, Curran e Roberts (2018) recodificam 23 classificações nacionais, entre elas, a CBO-Domiciliar, para a ISCO-88, a fim de associar medidas de intensidade de rotina e de possibilidade de deslocamento internacional das tarefas. Esses trabalhos também mostram que os resultados dependem da fonte, da taxonomia de destino e das regras de correspondência.

Trabalhos mais recentes exploram outras estratégias de mensuração no país. Firpo *et al.* (2023) contrastam medidas de tarefas baseadas em informações nacionais e norte-americanas ao analisar emprego e desigualdade no Brasil. Adamczyk, Ehrl e Monasterio (2024), por sua vez, extraem medidas de competências das descrições da CBO com técnicas de aprendizado de máquina e processamento de linguagem natural e as aplicam aos vínculos formais de 2003 a 2018. A pluralidade desses métodos reforça a importância de tornar explícitas as escolhas de fonte, classificação e agregação.

O presente artigo se insere nessa literatura sem propor uma nova teoria das tarefas nem reivindicar a primeira mensuração brasileira. Sua contribuição é preservar o mapeamento de 16 itens usado por Acemoglu e Autor (2011), atualizá-lo com a O*NET 30.1 e documentar duas rotas de transporte até a CBO 2002 e a COD. A cobertura de cada etapa, o efeito do preenchimento, a sensibilidade às regras de ponderação e a convergência com uma medida externa são examinados separadamente.

2 METODOLOGIA E BASE DE DADOS

A construção das harmonizações parte do enquadramento conceitual de Autor, Levy e Murnane (2003), originalmente baseado no Dictionary of Occupational Titles (DOT). Como a fonte utilizada aqui é a O*NET, a operacionalização empírica das cinco dimensões segue especificamente o apêndice de dados de Acemoglu e Autor (2011), que define um conjunto parcimonioso de 16 itens da O*NET. Assim, o estudo reproduz a seleção e o sentido dos itens daquele mapeamento, mas adapta as etapas posteriores de padronização e harmonização às finalidades da base brasileira.

A fonte de atributos ocupacionais é a Occupational Information Network (O*NET), sucessora do DOT e desenvolvida pelo National Center for O*NET Development sob patrocínio da Employment and Training Administration do Departamento do Trabalho dos Estados Unidos. A primeira etapa da harmonização transporta os escores da taxonomia O*NET-SOC 2019 até a ISCO-08.

A taxonomia O*NET-SOC é fortemente alinhada à Standard Occupational Classification (SOC), padrão estatístico federal de classificação ocupacional dos Estados Unidos. A SOC é coordenada pelo Office of Management and Budget, com apoio do Standard Occupational Classification Policy Committee, e divulgada pelo U.S. Bureau of Labor Statistics. A International Standard Classification of Occupations (ISCO), por sua vez, é elaborada pela International Labour Organization para permitir comparabilidade internacional de dados ocupacionais.

A construção dos escores utilizou informações da O*NET sobre habilidades, atividades e contexto de trabalho. A base contempla cinco dimensões de tarefa: cognitiva analítica não rotineira, cognitiva interpessoal não rotineira, cognitiva rotineira, manual rotineira e manual não rotineira. A TAB. 1 apresenta os 16 itens. Na dimensão cognitiva rotineira, o item atualmente denominado “Determine Tasks, Priorities and Goals” corresponde a “Structured versus Unstructured Work” no apêndice de Acemoglu e Autor (2011). Como valores altos indicam maior autonomia, seu sinal é invertido para que valores maiores representem maior estruturação e rotina.

TABELA 1 – Categorias da Base O*NET

Categoria	Variável
<i>Nonroutine Cognitive Analytical</i> (Cognitivo analítico não rotineiro)	<i>Analyzing Data or Information</i> (Analisar dados ou informações); <i>Thinking Creatively</i> (Pensar criativamente); <i>Interpreting the Meaning of Information for Others</i> (Interpretar o significado das informações para os outros).
<i>Nonroutine Cognitive Interpersonal</i> (Cognitivo interpessoal não rotineiro)	<i>Establishing and Maintaining Interpersonal Relationships</i> (Estabelecer e manter relações interpessoais); <i>Guiding, Directing, and Motivating Subordinates</i> (Orientar, dirigir e motivar subordinados); <i>Coaching and Developing Others</i> (Treinar e desenvolver outras pessoas).
<i>Routine Cognitive</i> (Cognitivo rotineiro)	<i>Determine Tasks, Priorities and Goals</i> (Determinar tarefas, prioridades e objetivos): com sinal invertido; <i>Importance of Repeating Same Tasks</i> (Importância de repetir as mesmas tarefas); <i>Importance of Being Exact or Accurate</i> (Importância de ser exato ou preciso).
<i>Routine Manual</i> (Manual rotineiro)	<i>Pace Determined by Speed of Equipment</i> (Ritmo determinado pela velocidade do equipamento); <i>Controlling Machines and Processes</i> (Controlar máquinas e processos); <i>Spend Time Making Repetitive Motions</i> (Passar tempo realizando movimentos repetitivos).
<i>Nonroutine Manual</i> (Manual não rotineiro)	<i>Operating Vehicles, Mechanized Devices, or Equipment</i> (Operar veículos, dispositivos mecanizados ou equipamentos); <i>Spend Time Using Your Hands to Handle, Control, or Feel Objects, Tools, or Controls</i> (Passar tempo usando as mãos para manusear, controlar ou sentir objetos, ferramentas ou controles); <i>Manual Dexterity</i> (Destreza manual); <i>Spatial Orientation</i> (Orientação espacial).

FONTE: Adaptado de Acemoglu e Autor (2011) e O*NET 30.1 Database (2025)

A justificativa dos itens é simultaneamente teórica e de comparabilidade. Os três componentes analíticos representam processamento, criação e interpretação de informação; os três interpessoais captam relações, direção e desenvolvimento de pessoas; os itens cognitivos rotineiros medem estrutura, repetição e exatidão; os manuais rotineiros captam ritmo imposto por equipamentos, controle de processos e

movimentos repetitivos; e os manuais não rotineiros medem operação, manipulação, destreza e orientação espacial. A reprodução do conjunto definido por Acemoglu e Autor (2011), em vez de uma seleção ad hoc por rótulo, reduz sobreposição entre construtos e assegura continuidade com a literatura.

Acemoglu e Autor (2011) empregam pesos de oferta de trabalho dos Estados Unidos para padronizar e colapsar ocupações em suas aplicações históricas. Neste artigo, as médias e os desvios-padrão são calculados sem ponderação por emprego no universo ocupacional da O*NET. Trata-se de uma adaptação deliberada: o objetivo é produzir escores ocupacionais neutros a país e ano, que possam ser posteriormente combinados com pesos de emprego escolhidos pelo usuário. Portanto, a estratégia reproduz o mapeamento dos itens, mas não constitui réplica literal de todas as ponderações do estudo de 2011.

Como a presente pesquisa utiliza O*NET, os escores foram construídos com itens do O*NET que correspondem a essas mesmas categorias conceituais. Seja, o valor bruto do item j para a ocupação o no O*NET. O primeiro passo consiste em padronizar cada item entre ocupações, por meio de um z-score:

$$z_{oj} = \frac{x_{oj} - \mu_j}{\sigma_j} \quad (1)$$

Em que μ_j e σ_j representam, respectivamente, a média e o desvio-padrão populacional do item j no conjunto das ocupações O*NET-SOC observadas. Essa padronização é necessária porque os itens pertencem a escalas diferentes e, portanto, precisam ser trazidos para uma métrica comparável antes da agregação. Esse procedimento é implementado de item a item, após reorganizar a base em formato largo por ocupação.

Em seguida, para cada dimensão de tarefa k , o escore bruto é construído como a soma dos itens padronizados pertencentes àquele bloco:

$$T_{ok}^{raw} = \sum_{j \in Jk} z_{oj} \quad (2)$$

Onde Jk representa o conjunto de itens associados à dimensão k . Depois disso, o escore composto é novamente padronizado entre ocupações:

$$T_{ok}^z = \frac{T_{ok}^{raw} - \mu_k}{\sigma_k} \quad (3)$$

Em que μ_k e σ_k representam, respectivamente, a média e o desvio-padrão de entre as ocupações O*NET-SOC observadas.

A etapa O*NET-SOC 2019 para SOC2018 foi tratada como uma agregação prévia, não como ponderação por caminhos, pois diferentes códigos O*NET-SOC 2019 podem convergir para um mesmo código SOC 2018. Para evitar super-representação de códigos SOC 2018 com maior número de ocupações-filhas no O*NET, os escores foram agregados no nível SOC 2018 por média simples entre os códigos O*NET-SOC 2019 associados a cada código SOC 2018:

$$\bar{T}_{sk} = \frac{1}{|O(s)|} \sum_{o \in O(s)} T_{ok}^z \quad (4)$$

Em que s representa o código SOC2018, o conjunto de ocupações O*NET-SOC 2019 associadas a esse código, e o o escore padronizado da dimensão para a ocupação. Após essa etapa de agregação, a compatibilização prossegue para SOC2018 → SOC2010 → ISCO-08 através de *crosswalks*⁴ oficiais.

Até o estágio ISCO-08, os procedimentos metodológicos são comuns às duas bases finais produzidas neste estudo. A partir desse ponto, as rotas de harmonização se diferenciam conforme a taxonomia de destino.

2.1 HARMONIZAÇÃO COM A CBO

Em seguida, foi necessário compatibilizar a base com a CBO através da série de *crosswalks* oficiais, apresentados na FIG. 1 abaixo:

FIGURA 1 – Cadeia de harmonização ocupacional entre O*NET-SOC 2019 e CBO 2002



FONTE: Os autores (2026)

Depois de construídos no nível O*NET-SOC 2019, os escores são compatibilizados até a CBO 2002 por meio da cadeia de *crosswalks* descrita acima.

Em etapas de compatibilização com correspondência muitos-para-muitos, que ocorre quando não há correspondência unívoca entre classificações, de modo que

⁴ Processo em que se conecta/traduz informações equivalentes em diferentes bases de dados

múltiplos códigos de uma classificação podem estar associados a múltiplos códigos de outra, e vice-versa, adotou-se uma regra simples e transparente: peso igual em cada desdobramento. Assim, quando um código de origem se desdobra em $1/n$ códigos de destino, cada caminho recebe peso $1/n$. Seja p um caminho completo de compatibilização ligando um código SOC2018, após a agregação prévia dos códigos O*NET-SOC 2019, a uma ocupação CBO 2002. O peso total do caminho é:

A regra $1/n$ é transparente, mas não é a única possível. Como teste de sensibilidade, os escores foram recalculados atribuindo peso igual a cada caminho terminal, em vez de dividir o peso em cada ramificação. Entre as 1.728 ocupações da CBO harmonizada, as correlações de Spearman entre as duas regras variaram de 0,862 a 0,981 nas cinco dimensões, e a dimensão predominante coincidiu em 90,7% dos casos. A regra por ramificação permanece como especificação principal porque representa cada desdobramento observado da cadeia. Quando um ramo não encontra correspondência posterior, seu peso não é redistribuído; a média de cada destino é calculada apenas com os pesos que efetivamente chegam a ele. Ponderações por emprego não foram adotadas, pois tornariam os escores dependentes do país, do ano e da fonte de emprego.

$$\omega_p^{CBO} = \omega_{SOC2018 \rightarrow SOC2010} \times \omega_{SOC2010 \rightarrow ISCO08} \times \omega_{ISCO08 \rightarrow ISCO88} \times \omega_{ISCO88 \rightarrow CBO94} \times \omega_{CBO94 \rightarrow CBO2002} \quad (5)$$

O escore final da ocupação brasileira c , na dimensão k , é obtido como média ponderada dos escores provenientes de todos os caminhos válidos que levam a essa ocupação:

$$S_{ck}^{CBO} = \frac{\sum_{p \in P(c)} \omega_p^{CBO} T_{pk}}{\sum_{p \in P(c)} \omega_p^{CBO}} \quad (6)$$

Em que $P(c)$ representa o conjunto de caminhos válidos que chegam à ocupação c , ω_p^{CBO} é o peso total do caminho p , e T_{pk} é o escore da dimensão associado ao código de origem transportado por esse caminho.

Como nem todas as ocupações da CBO 2002 receberam escore diretamente da cadeia de *crosswalks*, a base foi completada por um procedimento de imputação hierárquica com base na própria estrutura da CBO. A regra segue a hierarquia oficial da classificação ocupacional: os 4 primeiros dígitos fazem parte da família, os 3 primeiros dígitos do subgrupo, os 2 primeiros dígitos do subgrupo principal, e o primeiro dígito representa o grande grupo.

Para cada ocupação sem escore harmonizado, procura-se um conjunto de ocupações harmonizadas pertencentes ao mesmo prefixo hierárquico. O procedimento

começa no nível mais próximo disponível (4 dígitos) e, caso não haja ocupações harmonizadas naquele grupo, sobe sucessivamente para 3, 2 e 1 dígito. A imputação é feita exclusivamente com base em ocupações com escore obtido pela harmonização, sem utilizar valores já imputados como doadores.

Formalmente, para cada nível hierárquico define-se o conjunto de ocupações harmonizadas com o mesmo prefixo que a ocupação faltante c . Para cada dimensão k , o valor imputado corresponde à mediana dos escores brutos no conjunto de doadores:

$$\tilde{S}_{ck,\ell}^{\text{raw}} = \text{mediana} \{S_{dk}^{\text{raw}} : d \in D_{\ell}(c)\} \quad (7)$$

A imputação é realizada no nível mais desagregado possível, seguindo a ordem $4 \rightarrow 3 \rightarrow 2 \rightarrow 1$. Se não houver doadores em nenhum nível, a observação permanece sem preenchimento.

A imputação é aplicada aos escores compostos antes da transformação final. Para preservar a comparabilidade entre O*NET, CBO e COD, a medida principal reexpressa cada valor, harmonizado ou imputado, com a mesma média e o mesmo desvio padrão do escore composto no universo O*NET-SOC, reutilizando a referência definida na Equação 3. A Equação 8 apresenta, separadamente, uma padronização suplementar calculada após o fechamento da CBO. Essa segunda medida descreve posições relativas apenas dentro da CBO completa e não deve ser comparada diretamente com escores padronizados em outro universo.

$$S_{ck}^z = \frac{S_{ck}^{\text{raw}} - \bar{S}_k^{\text{raw}}}{\text{sd}(S_k^{\text{raw}})} \quad (8)$$

Na Equação 8, a barra sobre S representa a média dos escores brutos finais da dimensão k na CBO completa, e $\text{sd}(S)$ representa o respectivo desvio padrão.

A distinção entre as duas padronizações é substantiva. Na referência comum da O*NET, diferenças entre bases refletem o transporte, a cobertura e a imputação dos escores. Na padronização interna da CBO completa, reflete-se também a mudança de média e dispersão provocada pelo preenchimento. Por isso, as comparações entre CBO harmonizada, CBO completa, COD harmonizada e COD completa utilizam exclusivamente a referência O*NET; a medida da Equação 8 permanece suplementar para análises internas à CBO completa.

Em síntese, a base final CBO foi construída em duas etapas complementares. Na primeira, os escores de tarefas foram gerados no nível O*NET-SOC com base em itens exatos do O*NET, padronizados e agregados em cinco dimensões de tarefas. Na segunda, esses escores foram transportados até a CBO 2002 por meio de *crosswalks*

oficiais e, quando necessário, completados por imputação hierárquica segundo a própria estrutura da CBO. Assim, a base resultante não constitui uma réplica literal da mensuração original de Autor, Levy e Murnane (2003), uma vez que estes trabalharam com o DOT e com distribuições históricas de tarefas, sendo uma adaptação brasileira inspirada nos autores operacionalizada com O*NET contemporâneo e compatibilizações oficiais até a CBO 2002. Uma limitação importante é que a cadeia de *crosswalks* entre O*NET-SOC, SOC, ISCO e CBO contém várias relações muitos-para-muitos, de modo que parte da heterogeneidade ocupacional se perde ao longo da compatibilização. Isso contribui para que diferentes ocupações brasileiras possam compartilhar vetores de escores idênticos.

2.2 HARMONIZAÇÃO COM A COD

Para a construção da base compatibilizada com a Classificação de Ocupações para Pesquisas Domiciliares (COD), adotou-se uma rota direta a partir da ISCO-08. Em vez de prosseguir para ISCO-88, CBO94 e CBO 2002, realizou-se a harmonização final entre ISCO-08 e COD. Essa opção se justifica pelo fato de a COD ter sido elaborada para uso nas pesquisas domiciliares brasileiras com referência direta à ISCO-08, o que torna essa trajetória metodologicamente mais adequada para aplicações em bases domiciliares.

Diferentemente das etapas anteriores da cadeia, para as quais foram identificadas tabelas públicas de correspondência, a etapa final entre ISCO-08 e COD exigiu a construção de uma correspondência própria. O procedimento combinou pareamento estrutural por igualdade de código em quatro dígitos e revisão semântica dos títulos ocupacionais à luz das mudanças documentadas pelo IBGE. Como a planilha contém uma linha por unidade ISCO-08, ela não representa sozinha todos os desdobramentos um-para-muitos específicos da COD. Por essa razão, a cobertura também é calculada tomando como referência a lista oficial de códigos da COD.

Assim, a cadeia de compatibilização da base COD foi dada por O*NET-SOC 2019 → SOC2018 → SOC2010 → ISCO-08 → COD, como mostra a FIG. 2.

FIGURA 2 – Cadeia de harmonização ocupacional entre O*NET-SOC 2019 e COD



FONTE: Os autores (2026)

Em cada etapa com correspondência múltipla, manteve-se a regra de pesos iguais em cada desdobramento.

$$\omega_p^{\text{COD}} = \omega_{\text{SOC2018} \rightarrow \text{SOC2010}} \times \omega_{\text{SOC2010} \rightarrow \text{ISCO08}} \times \omega_{\text{ISCO08} \rightarrow \text{COD}} \quad (9)$$

O escore final da ocupação na COD, para a dimensão, foi calculado como média ponderada dos escores provenientes de todos os caminhos válidos:

$$S_{dk}^{\text{COD}} = \frac{\sum_{p \in P(d)} \omega_p^{\text{COD}} T_{pk}}{\sum_{p \in P(d)} \omega_p^{\text{COD}}} \quad (10)$$

A construção da COD é documentada em dois estágios diretamente comparáveis aos da CBO. A COD harmonizada contém os 415 códigos alcançados pela correspondência ISCO-08 para COD. A COD completa contém os 437 códigos classificáveis da lista oficial. Os 22 códigos adicionais são preenchidos em etapa separada: oito códigos militares ou de proteção recebem, dimensão por dimensão, a mediana de doadores selecionados manualmente por proximidade de conteúdo e previamente documentados; 13 recebem a mediana dos códigos harmonizados que compartilham o mesmo prefixo de três dígitos; e um recebe a mediana no prefixo de dois dígitos. Apenas os 415 códigos harmonizados podem atuar como doadores, de modo que valores preenchidos nunca alimentam novas imputações. A base registra, para cada preenchimento, o procedimento empregado e os códigos usados como doadores. O código 0000, referente a ocupações mal definidas, não recebe escore.

A COD completa é o produto recomendado para as aplicações empíricas, pois evita que a disponibilidade de uma correspondência determine quais trabalhadores permanecem na amostra. Usar somente a COD harmonizada excluiria observações de microdados de forma potencialmente não aleatória e faria o denominador variar conforme a cobertura da correspondência. A COD harmonizada é preservada como estágio diagnóstico, permitindo mensurar o efeito dos 22 preenchimentos sem criar especificações concorrentes. O ganho da completude é de cobertura e comparabilidade, não de certeza; por isso, a identificação da origem de cada escore deve acompanhar a base e pode orientar análises de sensibilidade.

A comparação interna usa pares no mesmo estágio de completude: COD harmonizada com CBO harmonizada e COD completa com CBO completa. Em cada par, calculam-se as médias dos escores nos dez grandes grupos comuns. Para cada dimensão k , o coeficiente de Pearson r_k mede a associação linear entre a média da COD e a média da CBO. Seja G o conjunto dos grandes grupos, a média da dimensão k no grupo g da COD e S_{gk}^{COD} a média correspondente na CBO; então:

$$r_k = \frac{\sum_{g \in G} \left(\bar{s}_{gk}^{COD} - \bar{s}_k^{COD} \right) \left(\bar{s}_{gk}^{CBO} - \bar{s}_k^{CBO} \right)}{\sqrt{\sum_{g \in G} \left(\bar{s}_{gk}^{COD} - \bar{s}_k^{COD} \right)^2 \sum_{g \in G} \left(\bar{s}_{gk}^{CBO} - \bar{s}_k^{CBO} \right)^2}} \quad (11)$$

Pearson é utilizado na comparação agregada entre CBO e COD porque ambas as bases estão expressas na mesma escala e o interesse é verificar se seus escores médios variam conjuntamente. Na validação externa e nos testes de robustez, a medida principal é o coeficiente de Spearman. Em vez de comparar os valores absolutos, ele compara a posição ocupada por cada observação nos dois conjuntos. Isso o torna adequado para confrontar medidas construídas em escalas diferentes, como os escores deste estudo e os de Mihaylov e Tijdens (2019).

$$\rho_{s,k} = \frac{\sum_{i=1}^n \left(R_{ik} - \bar{R}_k \right) \left(Q_{ik} - \bar{Q}_k \right)}{\sqrt{\left[\sum_{i=1}^n \left(R_{ik} - \bar{R}_k \right)^2 \right] \cdot \left[\sum_{i=1}^n \left(Q_{ik} - \bar{Q}_k \right)^2 \right]}} \quad (12)$$

Na Equação 12, R_{ik} e Q_{ik} representam os postos da observação i nas duas medidas da dimensão k , e as barras indicam os postos médios. Quando há empate, atribui-se às observações empatadas a média das posições que ocupariam. Assim, o coeficiente mostra em que medida as duas fontes ordenam as ocupações de maneira semelhante, mesmo que os valores numéricos de seus escores não sejam diretamente comparáveis.

2.3 ESTRATÉGIAS DE VALIDAÇÃO E ROBUSTEZ

A avaliação empírica combina três procedimentos. Primeiro, a comparação entre CBO e COD por grande grupo examina a consistência dos perfis produzidos pelas duas rotas de harmonização. Segundo a validação convergente confronta os escores ISCO-08 deste estudo com a medida independente de Mihaylov e Tijdens (2019) no nível de quatro dígitos. Terceiro, os testes de robustez avaliam a imputação e a sensibilidade às regras de ponderação. Spearman é a estatística principal nas duas últimas etapas, e Pearson é apresentado como medida complementar de associação linear.

A robustez da imputação da CBO é examinada por dois exercícios. No teste leave-one-out, cada ocupação harmonizada é temporariamente ocultada e prevista a partir das demais ocupações elegíveis. No teste por níveis hierárquicos, retiram-se progressivamente a ocupação-alvo e os grupos ocupacionais mais próximos, obrigando

o procedimento a recorrer a doadores mais distantes. Em ambos os exercícios, os valores previstos são comparados aos harmonizados pela correlação de Spearman e pela concordância da dimensão predominante.

Em uma verificação distinta, avalia-se a sensibilidade dos resultados à ponderação dos caminhos da harmonização. A regra principal divide o peso a cada ramificação da cadeia de correspondências, enquanto a alternativa atribui pesos iguais aos caminhos completos que alcançam a CBO. Na COD, verifica-se se os 415 escores harmonizados permanecem inalterados após a inclusão dos 22 códigos faltantes e realiza-se um exercício exploratório de retirada individual para avaliar a aproximação hierárquica.

A concordância da dimensão predominante é avaliada pela proporção de classificações recuperadas e pelo kappa de Cohen, que desconta a parcela de concordância esperada a partir das distribuições marginais das categorias (Cohen, 1960):

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (12)$$

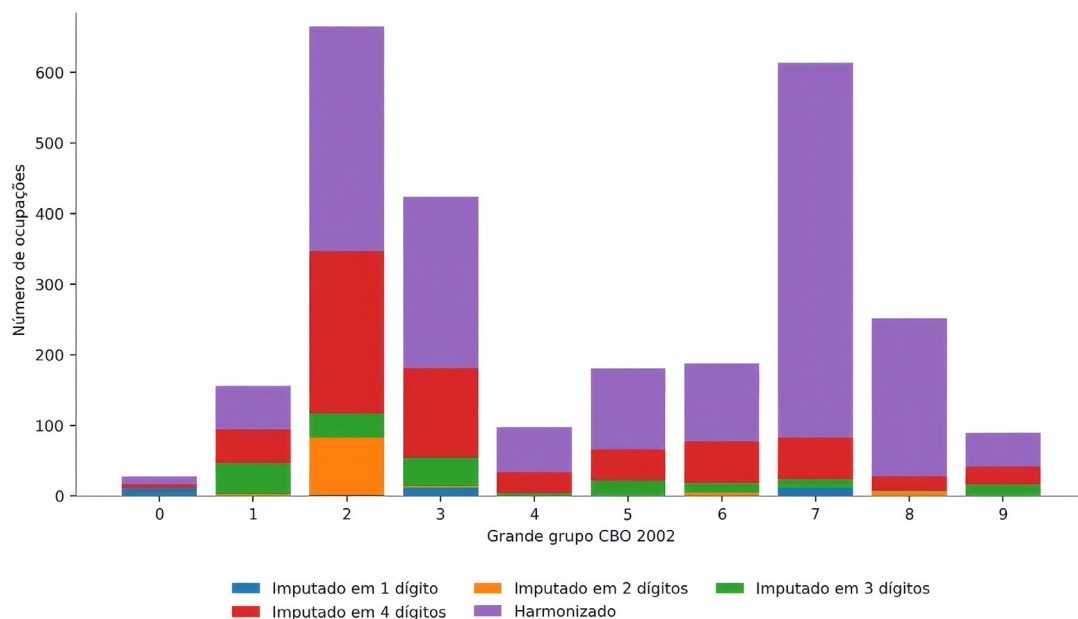
Na Equação 13, representa a proporção de concordância observada entre a dimensão original e a prevista, enquanto representa a concordância esperada a partir das frequências marginais das cinco dimensões. Quanto mais próximo de um, maior a concordância não explicada apenas pela distribuição das categorias.

3 VALIDAÇÃO TÉCNICA E ESTRUTURA FINAL

A construção da CBO 2002 foi avaliada em dois estágios. A CBO harmonizada reúne as 1.728 ocupações alcançadas diretamente pela cadeia de correspondências, equivalentes a 64,14% da lista mestra de 2.694 códigos. A CBO completa acrescenta 966 ocupações por imputação hierárquica: 645 em nível de família de quatro dígitos, 189 em três dígitos, 97 em dois dígitos e 35 em um dígito, sem casos não resolvidos. Os escores principais mantêm a referência O*NET; a padronização calculada após o preenchimento é mantida apenas como medida suplementar para análises internas da CBO completa.

A seguir, o GRAF. 1 apresenta a origem dos escores na CBO completa, segundo o grande grupo ocupacional. O gráfico distingue as ocupações da CBO harmonizada e aquelas acrescentadas por imputação hierárquica em 4, 3, 2 ou 1 dígito.

GRÁFICO 1 – Origem do escore por grande grupo ocupacional



FONTE: Os autores (2026)

O GRAF.1 mostra que a maior parte dos escores da CBO foi obtida diretamente pela cadeia de harmonização. As imputações concentram-se nos grupos em que os crosswalks preservaram menor cobertura. Como o preenchimento foi realizado sempre no nível hierárquico mais desagregado disponível, a maior parte ocorreu em 4 dígitos, isto é, dentro da própria família ocupacional. Esse padrão reduz o grau de arbitrariedade e preserva a rastreabilidade da passagem da CBO harmonizada à CBO completa.

3.1 ESTRUTURA E COBERTURA DA BASE EM COD

A base em COD foi compatibilizada diretamente a partir da ISCO-08. A rota parte de 894 ocupações O*NET-SOC com escores, agregadas em 768 códigos SOC 2010 e, depois, em 423 códigos ISCO-08 com escores válidos. A correspondência reúne 436 unidades ISCO-08: 435 associadas a destinos COD em quatro dígitos e uma associada apenas em dois dígitos. Esses números medem a cobertura das unidades ISCO, não a cobertura do universo COD.

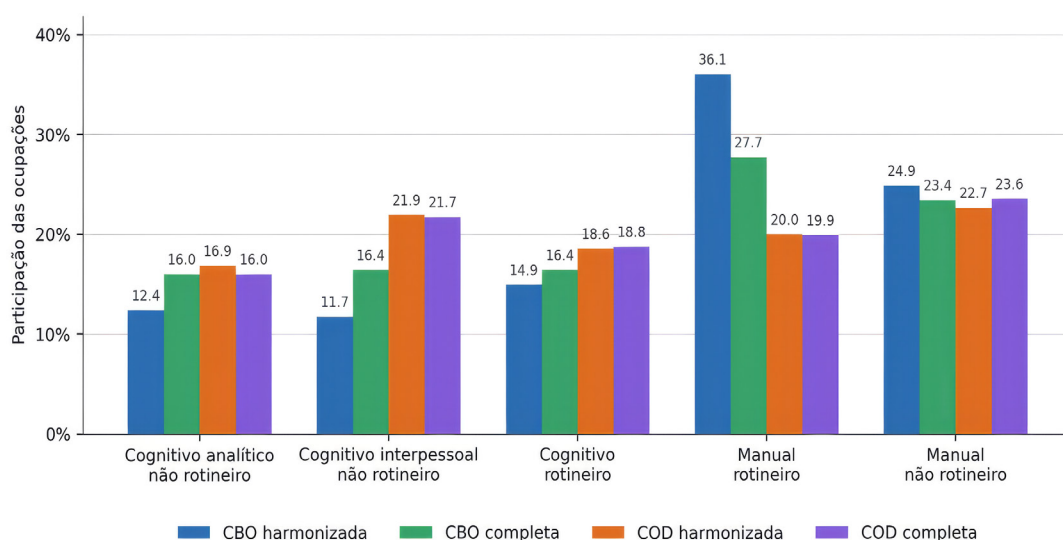
A lista oficial da COD contém 438 códigos: 437 classificáveis e o código residual 0000. A COD harmonizada alcança 415 dos códigos classificáveis, enquanto a COD completa acrescenta os 22 restantes: oito por doadores selecionados por proximidade de conteúdo, 13 por mediana hierárquica no prefixo de três dígitos e um no prefixo de dois dígitos. O código 0000 permanece sem escore por representar ocupações mal definidas, sem conteúdo ocupacional específico que sustente o preenchimento.

A COD harmonizada apresenta 392 vetores únicos entre 415 códigos, ou 5,54% de redundância. A COD completa apresenta 404 vetores únicos entre 437 códigos, ou 7,55%. O pequeno aumento decorre das medianas usadas nos 22 preenchimentos e permanece muito inferior ao observado na CBO. Os 415 escores harmonizados são preservados sem alteração, e a procedência dos demais é documentada para uso em filtros e análises de sensibilidade.

3.2 COMPARAÇÃO AGREGADA ENTRE AS DUAS BASES

O GRAF. 2 apresenta a distribuição da dimensão predominante nas quatro etapas: CBO harmonizada, CBO completa, COD harmonizada e COD completa. Para evitar que a escolha do universo de padronização determine o resultado, todas são expressas na mesma referência da O*NET.

GRÁFICO 2 – Distribuição da dimensão predominante por base de dados



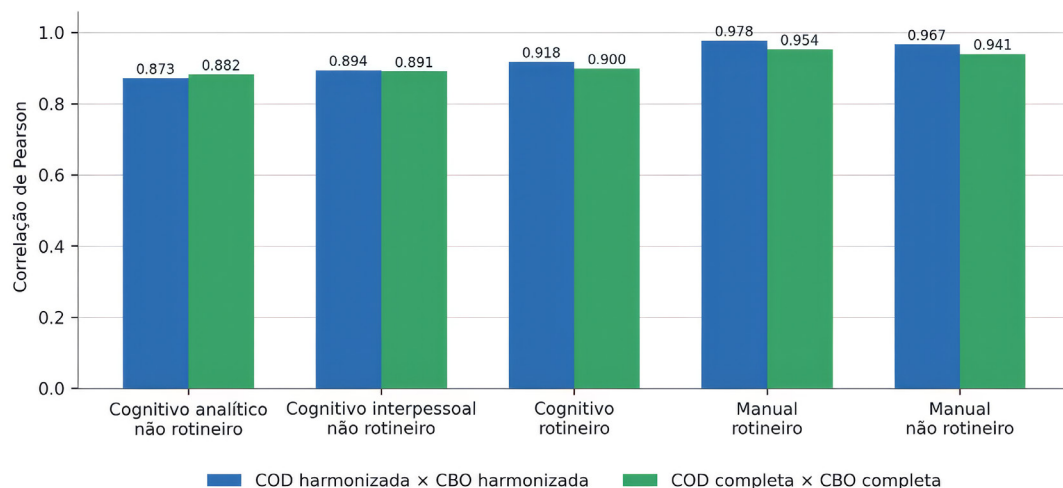
FONTE: Os autores (2026)

O GRAF. 2 mostra que a CBO harmonizada concentra maior parcela de ocupações nas dimensões manuais, sobretudo na manual rotineira. O preenchimento redistribui parte dessa concentração para as dimensões cognitivas, indicando que as ocupações não alcançadas pelas correspondências não se distribuem uniformemente entre os perfis de tarefas. Na COD, a composição já é mais equilibrada e muda pouco com a inclusão dos códigos faltantes, resultado coerente com sua cobertura inicial mais ampla.

O GRAF. 3 compara os escores médios da CBO e da COD nos dez grandes grupos ocupacionais, separando as versões harmonizadas e completas. O objetivo é verificar

se as duas classificações reproduzem perfis semelhantes nas cinco dimensões. Como a comparação é agregada e parte de uma fonte comum, ela é interpretada como evidência de consistência entre as duas rotas; a validação com uma fonte independente é apresentada no GRAF. 5.

GRÁFICO 3 – Compatibilidade dos escores médios por grande grupo e dimensão

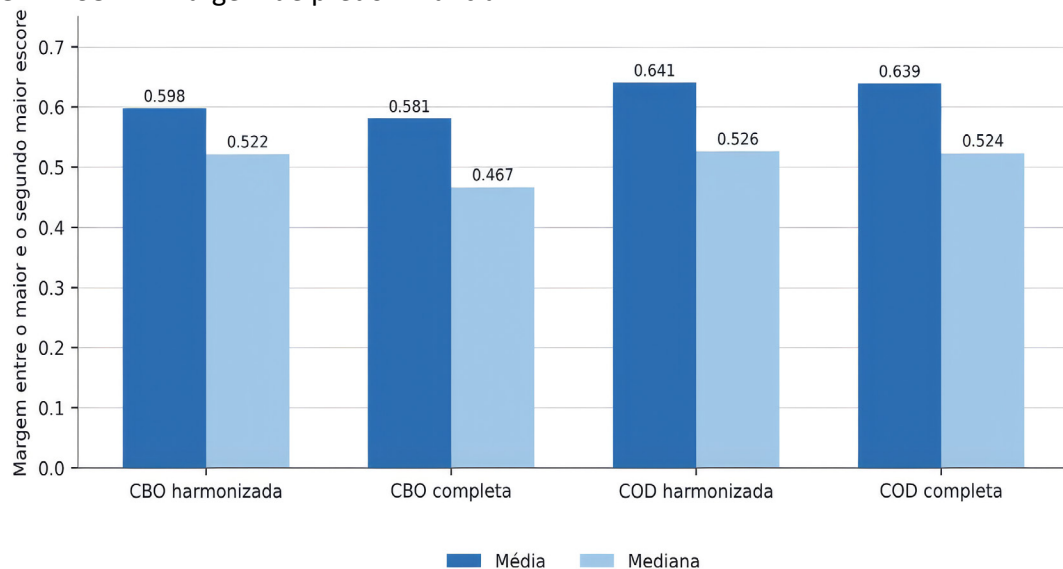


FONTE: Os autores (2026)

Os coeficientes são positivos e elevados nas cinco dimensões, indicando que CBO e COD reproduzem perfis médios semelhantes nos grandes grupos ocupacionais. A concordância é mais forte nas dimensões rotineiras e manuais, enquanto as dimensões cognitivas apresentam alguma variação adicional na ordenação dos grupos. A passagem das bases harmonizadas para as completas preserva o padrão geral, embora a ordenação da dimensão manual não rotineira seja mais sensível ao preenchimento.

O GRAF. 4 apresenta a margem entre o maior e o segundo maior escore de cada ocupação. Valores mais altos indicam maior separação entre as duas dimensões mais intensas. Novamente, todas as bases são expressas na referência comum da O*NET.

GRÁFICO 4 – Margem de predominância



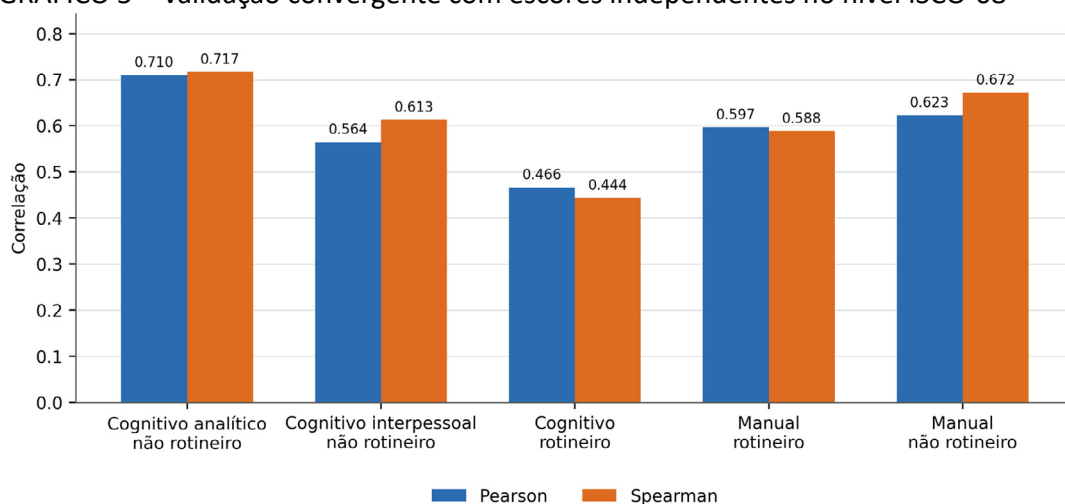
FONTE: Os autores (2026)

A margem média é 0,598 na CBO harmonizada, 0,581 na CBO completa, 0,641 na COD harmonizada e 0,639 na COD completa; as medianas são 0,522, 0,467, 0,526 e 0,524, respectivamente. Na referência comum, o preenchimento reduz ligeiramente a separação entre a primeira e a segunda dimensão nas duas classificações, mas o efeito é muito menor na COD. Uma conclusão oposta aparece apenas quando cada base completa é repadronizada internamente, pois essa operação altera sua média e sua dispersão; por isso, essa escala não é usada na comparação entre bases. A maior margem média da COD é descritiva e não implica, isoladamente, maior ou menor validade.

3.3 VALIDAÇÃO CONVERGENTE NO NÍVEL OCUPACIONAL

O GRAF. 5 apresenta a validação externa que utiliza os escores de Mihaylov e Tijdens (2019), construídos independentemente da O*NET a partir da classificação manual de 3.264 descrições de tarefas da ISCO-08. A comparação é realizada diretamente no nível de quatro dígitos, evitando a agregação em grandes grupos. Foram pareadas 414 ocupações entre os dois conjuntos, número próximo ao universo de 427 ocupações reportado pelos autores.

GRÁFICO 5 – Validação convergente com escores independentes no nível ISCO-08



FONTE: Adaptado de Mihaylov e Tijdens (2019)

O GRAF. 5 mostra associações positivas em todas as dimensões: ocupações que aparecem em posições mais altas nos escores deste estudo tendem também a ocupar posições mais altas na medida independente de Mihaylov e Tijdens. A convergência é mais nítida nas dimensões cognitiva analítica e manual não rotineira e mais fraca na cognitiva rotineira, sugerindo que repetição, precisão e conteúdo procedimental são mais sensíveis à forma como cada fonte descreve as tarefas. Em complemento, a dimensão predominante é compatível em aproximadamente dois terços das 414 ocupações comparadas.

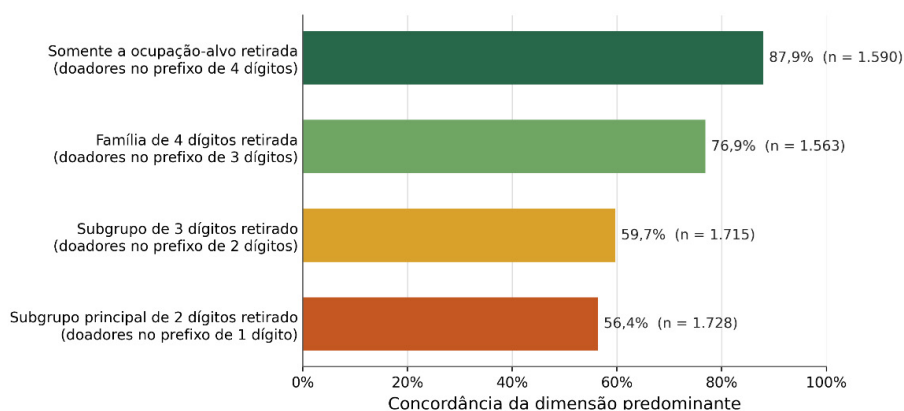
Como as duas medidas usam fontes e escalas distintas, somas padronizadas de itens do O*NET neste estudo e proporções de descrições de tarefas em Mihaylov e Tijdens (2019), não se espera igualdade numérica. As associações moderadas a fortes e a recuperação dos padrões ocupacionais conhecidos oferecem evidência convergente de validade de construto, sem eliminar incerteza específica de cada item ou crosswalk.

3.4 ROBUSTEZ DA IMPUTAÇÃO E DAS REGRAS DE HARMONIZAÇÃO

No teste leave-one-out, cada uma das 1.728 ocupações harmonizadas é temporariamente retirada e reconstruída pela mediana dos demais doadores disponíveis. A dimensão predominante é recuperada em 86,2% dos casos, e o kappa de Cohen é 0,817, indicando que a elevada concordância não decorre apenas da distribuição das categorias. Como referência adicional, uma regra ingênua que atribísse sempre a dimensão mais frequente acertaria 36,1% dos casos. As correlações de Spearman entre os escores harmonizados e previstos variam de 0,819 a 0,939, mostrando que a ordenação das ocupações também permanece semelhante nas cinco dimensões.

Para avaliar como a distância entre a ocupação-alvo e seus doadores afeta a imputação, realizou-se um segundo exercício de retirada. No primeiro cenário, oculta-se somente a ocupação-alvo e preservam-se as demais ocupações de sua família de quatro dígitos. Nos cenários seguintes, retiram-se sucessivamente toda a família, o subgrupo de três dígitos e o subgrupo principal de dois dígitos, de modo que a previsão passe a depender de ocupações progressivamente mais distantes. O GRAF. 6 apresenta a proporção de casos em que a dimensão predominante original é recuperada em cada cenário, onde n corresponde ao número de ocupações harmonizadas para as quais permaneceu ao menos um doador elegível após a retirada definida em cada cenário.

GRÁFICO 6 – Concordância da imputação sob níveis crescentes de afastamento dos doadores



FONTE: Os autores (2026)

A concordância diminui à medida que os doadores mais próximos são retirados. Quando permanecem ocupações da mesma família ou do mesmo subgrupo, o procedimento recupera com maior frequência a dimensão predominante original. O desempenho se reduz quando a previsão depende apenas de unidades mais amplas da hierarquia, mostrando que a proximidade ocupacional contém informação relevante para a imputação. O resultado recomenda maior cautela no uso de escores preenchidos a partir de doadores disponíveis somente nos níveis de um ou dois dígitos, sobretudo em análises muito desagregadas.

Os demais testes reforçam a estabilidade geral dos resultados. A regra alternativa de pesos iguais por caminho terminal preserva a dimensão predominante em 90,7% da CBO harmonizada, com correlações de Spearman de 0,862 a 0,981. Na COD, a passagem do estágio harmonizado ao completo preserva os 415 códigos originais e acrescenta 22 casos com procedência documentada. No exercício exploratório de retirada individual, a regra hierárquica permite prever 414 dos 415 códigos harmonizados; a concordância da dimensão predominante é 65,7% e as correlações de Spearman variam de 0,624 a

0,841. Esse exercício avalia apenas a aproximação hierárquica e não valida a escolha manual dos oito conjuntos de doadores por proximidade de conteúdo. Entre esses oito casos, sete mantêm a mesma dimensão predominante nas variações examinadas dos doadores documentados, enquanto um alterna entre as duas dimensões manuais. Em conjunto, os testes delimitam onde os resultados são mais estáveis e onde análises de sensibilidade são recomendáveis.

CONSIDERAÇÕES FINAIS

Este artigo teve como objetivo construir e documentar bases de escores ocupacionais baseados em tarefas para a CBO 2002 e a COD, aproximando a literatura internacional de tarefas e tecnologia das classificações ocupacionais utilizadas no Brasil⁵. Partindo de informações da O*NET e da operacionalização de Acemoglu e Autor (2011), o estudo desenvolveu uma infraestrutura metodológica que combina padronização de itens, agregação em dimensões de tarefa, harmonização entre taxonomias ocupacionais e procedimentos explícitos de validação.

Os resultados indicam que as duas rotas geram produtos distintos, mas substantivamente coerentes. Na CBO 2002, a etapa harmonizada reúne 1.728 ocupações e a etapa completa acrescenta 966 imputações. Na COD, a etapa harmonizada reúne 415 códigos e a completa acrescenta 22 códigos com procedência documentada, alcançando os 437 classificáveis. A apresentação dos dois estágios permite medir o efeito do preenchimento; para aplicações em microdados, as versões completas evitam seleção decorrente da cobertura das correspondências, enquanto as versões harmonizadas permanecem úteis para auditoria e sensibilidade.

A comparação por grande grupo mostra que CBO e COD recuperam perfis médios semelhantes, e a validação com Mihaylov e Tijdens (2019) encontra associações positivas nas cinco dimensões apesar das diferenças de fonte e escala. Os testes de retirada individual e por níveis hierárquicos mostram, por sua vez, que a imputação é mais estável quando os doadores estão próximos. A referência comum da O*NET também permite estabelecer que o preenchimento reduz ligeiramente a margem entre a primeira e a segunda dimensão, em vez de ampliá-la.

As limitações delimitam o uso das bases. A correspondência ISCO-08–COD foi construída pelos autores e sua estrutura original de uma linha por unidade ISCO não

⁵ As bases finais, os crosswalks, os scripts e os materiais de validação produzidos no âmbito desta pesquisa estão disponíveis no Zenodo, versão 1.3.1: <https://doi.org/10.5281/zenodo.21974455>.

representa todos os desdobramentos brasileiros. O fechamento da COD pressupõe semelhança entre ocupações associadas por conteúdo ou por prefixo, razão pela qual a procedência dos 22 preenchimentos deve ser considerada em análises mais desagregadas. Na CBO, a perda de precisão nos níveis hierárquicos mais amplos recomenda cautela com os casos preenchidos a partir de doadores distantes. Por fim, a O*NET descreve ocupações dos Estados Unidos e não capta diretamente variações da organização do trabalho no Brasil. Essas limitações não anulam a comparabilidade construída, mas indicam extensões relevantes com tarefas autorreportadas, salários e composição do emprego brasileiros.

REFERÊNCIAS

- ACEMOGLU, Daron; AUTOR, David. Skills, tasks and technologies: implications for employment and earnings. In: ASHENFELTER, Orley; CARD, David (Ed.). **Handbook of Labor Economics**. Amsterdam: Elsevier, 2011. p. 1043-1171.
- ADAMCZYK, Willian; EHRL, Philipp; MONASTERIO, Leonardo. Skills and employment transitions in Brazil. **International Labour Review**, [s.l.], v. 163, n. 2, p. 199–224, 2024. DOI: 10.1111/ilr.12412.
- AUTOR, David H.; DORN, David. The growth of low-skill service jobs and the polarization of the US labor market. **American Economic Review**, Pittsburgh, v. 103, n. 5, p. 1553-1597, 2013. DOI: 10.1257/aer.103.5.1553.
- AUTOR, David H.; LEVY, Frank; MURNANE, Richard J. The skill content of recent technological change: an empirical exploration. **Quarterly Journal of Economics**, Oxford, v. 118, n. 4, p. 1279-1333, 2003.
- BRASIL. Ministério do Trabalho e Emprego. **Classificação Brasileira de Ocupações**: Estrutura, tábua de conversão e índice de títulos. 3. Ed. Brasília: Ministério do Trabalho e Emprego, 2010. Disponível em: https://www.gov.br/trabalho-e-emprego/pt-br/assuntos/cbo/servicos/downloads/cbo2002_liv3.pdf. Acesso em: 27 abr. 2026.
- BRASIL. Ministério do Trabalho e Emprego. Downloads: Classificação Brasileira de Ocupações. **Ministério do Trabalho e Emprego**, Brasília, [s.d.]. Disponível em: <https://www.gov.br/trabalho-e-emprego/pt-br/assuntos/cbo/servicos/downloads>. Acesso em: 27 abr. 2026.
- BUREAU OF LABOR STATISTICS (UNITED STATES). **Standard Occupational Classification Manual 2018**. Washington: U.S. Department of Labor, 2018. Disponível em: https://www.bls.gov/soc/2018/soc_2018_manual.pdf. Acesso em: 27 abr. 2026.
- BUREAU OF LABOR STATISTICS (UNITED STATES). **2010 to 2018 SOC crosswalk**. Washington: U.S. Department of Labor, 2017. Disponível em: https://www.bls.gov/soc/2018/soc_2010_to_2018_crosswalk.xlsx. Acesso em: 27 abr. 2026.
- BUREAU OF LABOR STATISTICS (UNITED STATES). **ISCO-08 x SOC 2010 crosswalk**. Washington: U.S. Department of Labor, 2012. Atualizado em jun. 2015. Disponível em: https://www.bls.gov/soc/isco_soc_crosswalk.xls. Acesso em: 27 abr. 2026.
- COHEN, Jacob. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, [s.l.], v. 20, n. 1, p. 37–46, 1960. DOI: <https://doi.org/10.1177/001316446002000104>
- FIRPO, Sergio; PORTELLA, Alysson; RIVA, Flavio; ÚBIDA, Giovanna. Brazil: employment and inequality trends. In: GRADÍN, Carlos; LEWANDOWSKI, Piotr; SCHOTTE, Simone; SEN, Kunal (Ed.). **Tasks, skills, and institutions: the changing nature of work and inequality**. Oxford: Oxford University Press, 2023. p. 229–252. DOI: 10.1093/oso/9780192872241.003.0013.
- GOOS, Maarten; MANNING, Alan; SALOMONS, Anna. Explaining job polarization: routine-biased technological change and offshoring. **American Economic Review**, Pittsburgh, v. 104, n. 8, p. 2509–2526, 2014. DOI: 10.1257/aer.104.8.2509.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Comissão Nacional de Classificação. **Ata da 25ª Reunião Ordinária, de 18 de agosto de 2010**. Rio de Janeiro: IBGE/Concla, 2010. Disponível em: https://concla.ibge.gov.br/images/concla/resolucoes_e_atas/ataconcla25_180810.pdf. Acesso em: 27 abr. 2026.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Classificação de Ocupações para Pesquisas Domiciliares – COD**. Rio de Janeiro: IBGE, 2010. Disponível em: https://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/metodologia/anexos/anexo_7_ocupacao_cod.pdf. Acesso em: 30 jul. 2026.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Classificação de Ocupações para Pesquisas Domiciliares – COD**: principais diferenças entre a COD e a CBO-Domiciliar. In: 8º FÓRUM SIPD, 2010, Rio de Janeiro. Rio de Janeiro: IBGE, 09 jul. 2010. Disponível em: https://www.ibge.gov.br/arquivo/projetos/sipd/oitavo_forum/COD.pdf. Acesso em: 27 abr. 2026.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **IBGE x CBO 94**. Rio de Janeiro: IBGE/Concla, [s.d.]. [Arquivo XLS]. Disponível em: <https://concla.ibge.gov.br/images/concla/documentacao/ibgexcbo94.xls>. Acesso em: 27 abr. 2026.

INTERNATIONAL LABOUR ORGANIZATION (ILO). **International Standard Classification of Occupations**: Structure, group definitions and correspondence tables. Geneva: ILO, 2012. Disponível em: <https://www.ilo.org/publications/international-standard-classification-occupations-2008-isco-08-structure>. Acesso em: 27 abr. 2026.

INTERNATIONAL LABOUR ORGANIZATION (ILO). **Correspondence ISCO-08 to ISCO-88**. Geneva: ILO, [s.d.]. Arquivo XLSX. Disponível em: https://webapps.ilo.org/ilostat-files/Documents/Correspondence_EN_ISCO_08_to_ISCO_88.xlsx. Acesso em: 27 abr. 2026.

MACIENTE, Aguinaldo Nogueira. Competências e habilidades ocupacionais no Brasil. In: OLIVEIRA, Marina Pereira Pires de *et al.* (Orgs.). **Rede de pesquisa**: formação e mercado de trabalho : coletânea de artigos : Volume 2 : Qualificação profissional : demandas e estratégias das firmas. Brasília: Ipea: ABDI, 2014. v. 2, p. 20-44. Disponível em: <https://repositorio.ipea.gov.br/handle/11058/20550>. Acesso em: 30 jul. 2026.

MAHUTGA, Matthew C.; CURRAN, Michaela; ROBERTS, Anthony. Job tasks and the comparative structure of income and employment: routine task intensity and offshorability for the LIS. **International Journal of Comparative Sociology**, [s.l.], v. 59, n. 2, p. 81–109, 2018. DOI: 10.1177/0020715218765218

MIHAYLOV, Emil; TIJDENS, Kea. **Measuring the routine and non-routine task content of 427 four-digit ISCO-08 occupations**. Amsterdam: Tinbergen Institute, 2019. Disponível em: <https://papers.tinbergen.nl/19035.pdf>. Acesso em: 30 jul. 2026.

NATIONAL CENTER FOR O*NET DEVELOPMENT. O*NET-SOC 2019 occupation listings: crosswalk O*NET-SOC 2019 to 2018 SOC. **O*NET Resource Center**, [s.l.; s.d.]. Arquivo XLSX. Disponível em: https://www.onetcenter.org/taxonomy/2019/soc/2019_to_SOC_Crosswalk.xlsx?fmt=xlsx. Acesso em: 27 abr. 2026.

NATIONAL CENTER FOR O*NET DEVELOPMENT. O*NET database releases archive. **O*NET Resource Center**, [s.l.; s.d.]. Disponível em: https://www.onetcenter.org/db_releases.html. Acesso em: 27 abr. 2026.

NATIONAL CENTER FOR O*NET DEVELOPMENT. **O*NET database 30.1**. Sponsored by U.S. Department of Labor, Employment and Training Administration. [S. l.]: O*NET Resource Center, 2025. Disponível em: <https://www.onetcenter.org/dictionary/30.1/excel/>. Acesso em: 27 abr. 2026.