

Inteligência artificial generativa e mercado de trabalho no Brasil: evidências iniciais sobre ocupação e renda

Daniel Duque

TECNOLOGIA

Estudo indica que a IA generativa impacta 30 milhões de trabalhadores no Brasil, com efeitos concentrados em jovens de 18 a 29 anos, resultando em queda na ocupação e renda. A análise aponta para a substituição de tarefas em funções de entrada.

Introdução

Com o avanço recente da inteligência artificial, em especial da IA generativa, o debate sobre uma questão clássica da economia do trabalho ressurgiu. Admite-se que as novas tecnologias elevam produtividade, mas também podem deslocar tarefas, alterar o perfil das ocupações e produzir efeitos distributivos relevantes. Em um recente estudo, [Paulo Peruchetti, Fernando de Holanda Barbosa Filho e Janaína Feijó](#) mapeiam a exposição potencial das ocupações à IA generativa com base em um índice da OIT, estruturado a partir de tarefas e alinhado ao padrão ISCO-08. Ainda, eles discutem que o problema não deve ser lido como uma dicotomia simples entre “substituição” e “complementaridade”, mas como uma reconfiguração do conteúdo do trabalho.

Esse mapeamento já mostra que o impacto potencial é relevante. No terceiro trimestre de 2025, quase 30 milhões de trabalhadores no Brasil estavam em ocupações com algum grau de exposição à IA generativa, o equivalente a 29,6% da população ocupada. Ainda, cerca de 5,2 milhões estavam no gradiente mais elevado de exposição. A exposição é maior entre mulheres, jovens, mais escolarizados, na região Sudeste e no setor de serviços, especialmente em informação e comunicação e serviços financeiros. O próprio estudo também recorre à evidência do FMI para mostrar que, no caso brasileiro, algo em torno de 20% dos ocupados combina alta exposição com baixa complementaridade, sendo esse o

grupo mais vulnerável a perdas de emprego, enquanto pouco mais de 20% combina alta exposição com elevada complementaridade, o que sugere potencial de ganhos de produtividade e salário.

Mas exposição potencial não é impacto realizado. Um índice ocupacional baseado em tarefas informa onde a IA pode incidir, não como ela já está afetando o mercado de trabalho. Este artigo procura justamente dar esse passo adicional. Com isso, estimo efeitos iniciais da difusão pública da IA generativa sobre ocupação e renda do trabalho no Brasil, inspirando-se conceitualmente no estudo de Peruchetti, Barbosa Filho e Feijó e, metodologicamente, numa estratégia semelhante à que utilizei em um estudo anterior sobre os efeitos dos aplicativos no mercado de trabalho: comparar, ao longo do tempo, grupos *ex ante* mais expostos com grupos menos expostos, usando uma medida de exposição predita construída antes do choque tecnológico se difundir.

Metodologia

A base empírica é a PNAD Contínua trimestral em cortes repetidos de 2022T1 a 2025T2, com dois desfechos principais: um indicador de ocupação (ocupado versus não ocupado) e o logaritmo da remuneração real do trabalho. O desenho econométrico segue uma lógica de diferenças-em-diferenças dinâmico com “dose” contínua de exposição. O problema central é evitar endogeneidade por realocação ocupacional: se a IA já estiver afetando o mercado, medir a exposição pela ocupação corrente mistura efeito causal com seleção pós-choque. Para contornar esse problema, a exposição é congelada no cenário pré-IA, em 2022T3, o último trimestre plenamente anterior ao lançamento público inicial das ferramentas que marcam o início da difusão mais ampla.

Primeiro estágio (exposição predita). Para evitar endogeneidade por realocação ocupacional após a difusão da IA, construímos um grau contínuo de exposição predita a partir de um *snapshot* pré-IA. No 2022T3 (último trimestre plenamente anterior ao lançamento público inicial), estimamos, em amostra de 14+ anos, a probabilidade de um indivíduo estar numa ocupação altamente exposta como função do perfil demográfico e geográfico (idade×sexo, escolaridade, raça/cor, unidade da federação, urbano/rural). Formalmente:

$$\Pr(H_i = 1 | X_i) = \alpha + X_i' \gamma + \delta_{UF} + \delta_{urb/rur} + u_i,$$

e definimos ($\widehat{E}_i \equiv \widehat{Pr}(H_i = 1 | X_i)$). (LPM com pesos; resultados do primeiro estágio apenas para construir ($\widehat{E}_i$).)

Estimo então, separadamente por coorte etária, um *event-study* que interage a dose contínua pré-determinada com indicadores de trimestre, tomando um período base (τ_0) pré-IA:

$$Y_{it} = \alpha + \sum_{\tau \neq \tau_0} \beta_{\tau} (E_i^* \times \mathbb{1}\{t = \tau\}) + \lambda' X_i + \mu_g + \theta_t + \xi_{c \times t} + \varepsilon_{it}.$$

Aqui, (Y_{it}) é (i) o indicador de ocupação; (X_i) inclui idade×sexo e escolaridade; (μ_g) são fixos de grupos demográficos; (θ_t) são efeitos fixos de tempo; e ($\xi_{c \times t}$) são efeitos fixos de UF×área (urbano/rural) por trimestre, absorvendo choques locais específicos no tempo. Os erros-padrão são agrupados em células demográficas (idade×sexo). O coeficiente (β) mede a variação diferencial (vs. (τ_0)) na probabilidade de estar empregado por unidade de exposição predita.

A mesma metodologia é aplicada aos salários reais. A interpretação de (β_{τ}) passa a ser o efeito percentual (aproximação log-linear) na remuneração por unidade de exposição predita, relativamente ao período base.

Adicionalmente, para cada coorte etária, estimo a tendência prévia da dose contínua sobre a variável nos trimestres pré-IA:

$$Y_{it} = \alpha + \theta_{pre} (E_i^* \times t) + \lambda' X_i + \mu_g + \theta_t + \xi_{c \times t} + \varepsilon_{it} \quad (\text{só pré-IA}).$$

Em seguida, ajustamos os (β_{τ}) do *event-study* pós-IA por um *de-trending* linear:

$$\widetilde{\beta}_{\tau} = \widehat{\beta}_{\tau} - (\tau - \tau_0), \widehat{\theta}_{pre},$$

A estratégia, portanto, é um estimador de diferenças em diferenças com exposição contínua fixa pré-IA (derivada de 2022T3) e choques

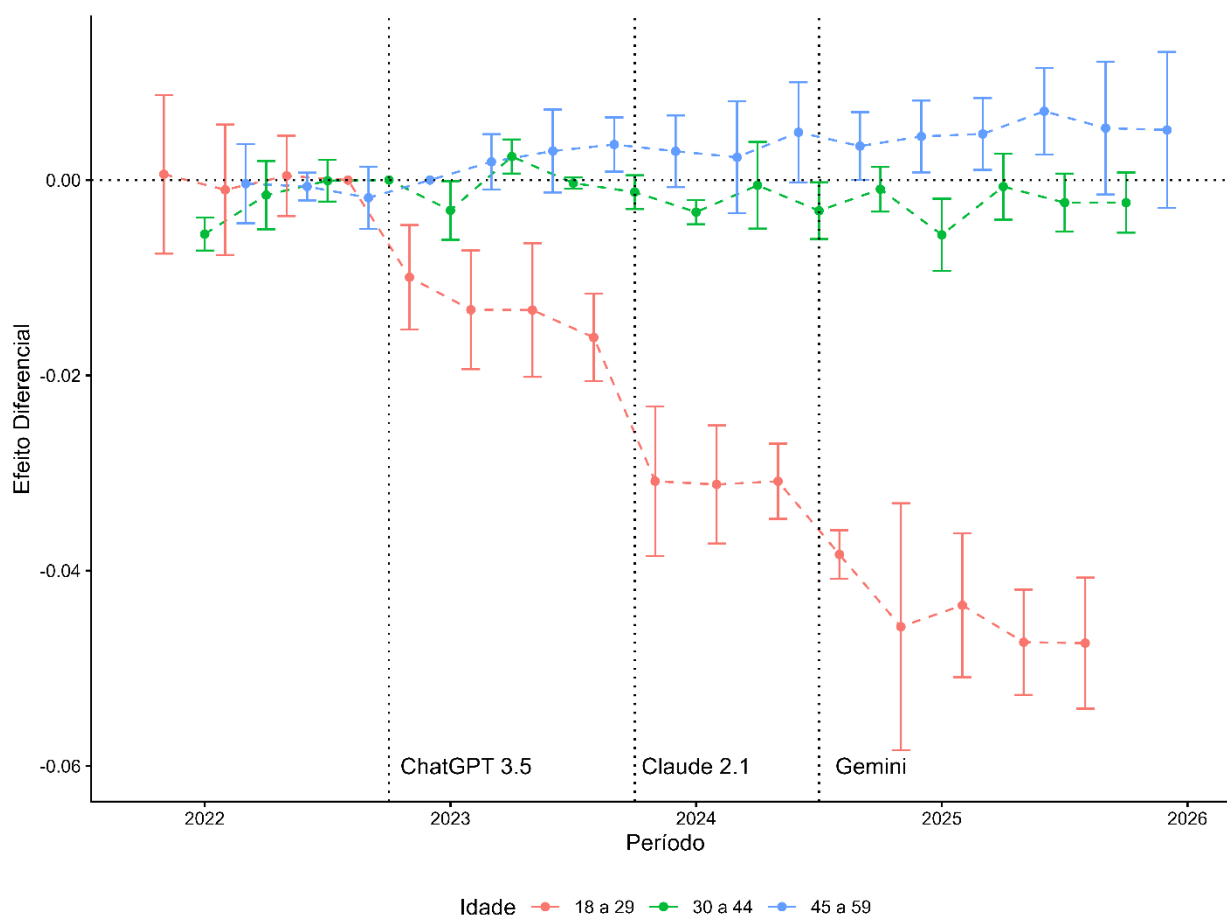
temporais comuns, contrastando grupos ex-ante mais expostos com menos expostos. As hipóteses-chave são: (i) tendências paralelas condicionais entre grupos de exposição dentro de células $UF \times \text{área}$ e perfis demográficos/educacionais; (ii) estabilidade da composição dos grupos (tratada ao fixar a dose com base no cenário pré-IA); e (iii) ausência de choques locais não observados que afetem diferencialmente grupos de alta exposição dentro da mesma célula quando já controlamos por $(\xi_{c \times t})$. Em termos substantivos, a exposição (do índice ocupacional) é ligada a variações no mercado de trabalho.

Relação com o índice da OIT e limitações. O uso de um índice baseado em tarefas/ocupações ISCO-08 oferece comparabilidade e fundamentação metodológica, e as classes de exposição derivam de médias e desvios dos escores de tarefas e foram refinadas para melhorar estabilidade e interpretação em relação ao exercício de 2023. Isso permite mapear ocupações brasileiras a um gradiente global, inclusive para exercícios de ligação com microdados nacionais. Ainda assim, o índice mede potencial e “o que a tecnologia pode fazer com tarefas”, não o que já foi feito. Vieses de mensuração da dose, heterogeneidade na velocidade de difusão setorial e limitações de infraestrutura/habilidades significam que são estimados efeitos diferenciais compatíveis com substituição na margem de entrada, não efeitos totais.

Resultados

Os resultados para ocupação apontam um padrão bastante claro. Entre os jovens de 18 a 29 anos, os coeficientes do *event-study* tornam-se crescentemente negativos no período pós-IA e se aprofundam ao longo de 2024 e 2025. Portanto, entre os mais jovens, perfis demográficos mais expostos à IA generativa passaram a apresentar uma queda diferencial na probabilidade de estar ocupado, relativamente ao período base. Já entre os grupos de 30 a 44 anos e 45 a 59 anos, os coeficientes oscilam ao redor de zero durante a maior parte da amostra, com efeitos médios muito menores e em geral compatíveis com ausência de impacto estatisticamente robusto. O padrão sugere que, ao menos neste estágio inicial, a IA generativa está incidindo sobretudo sobre tarefas de entrada, mais frequentes entre trabalhadores jovens.

Event-study dos efeitos diferenciais de exposição à IA sobre ocupação por grupo etário

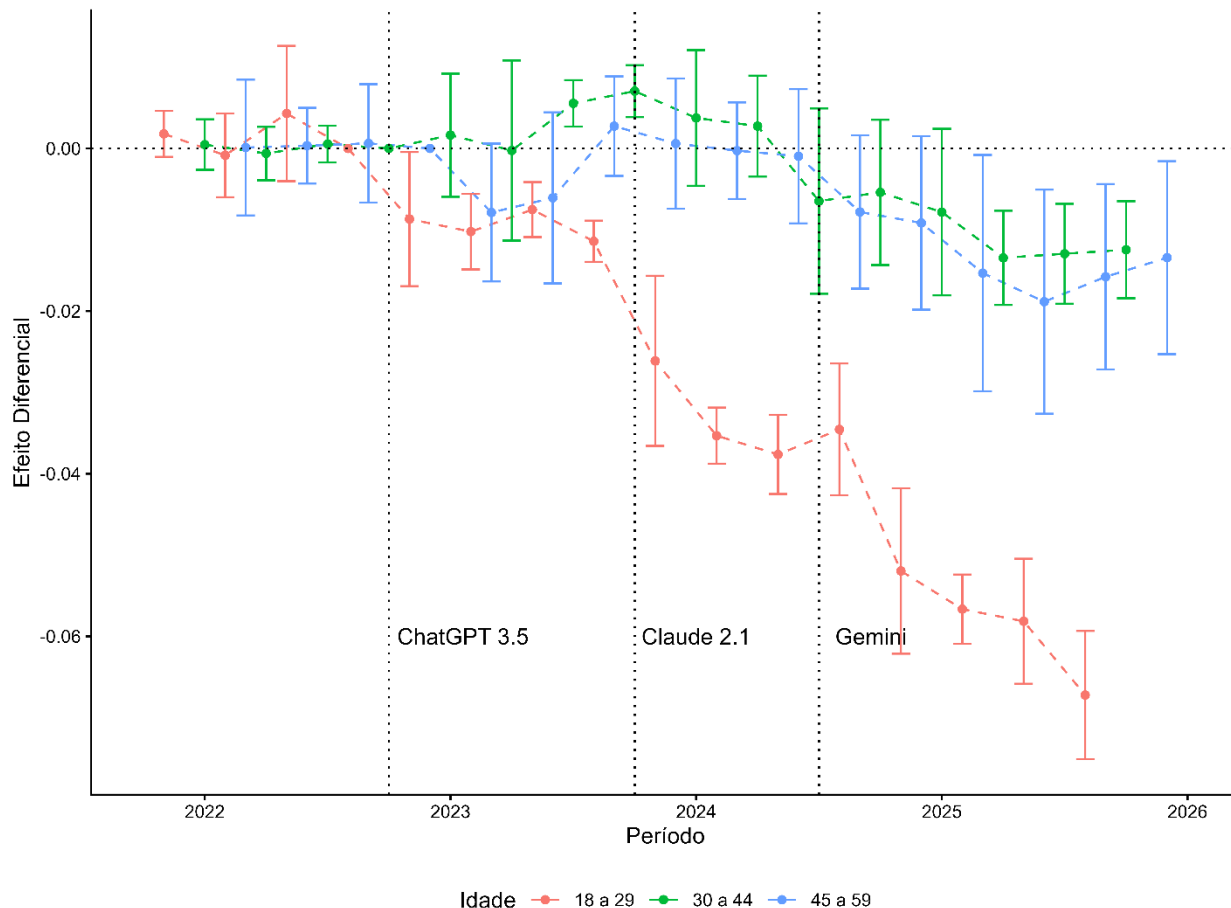


Fonte: Elaboração própria com dados da PNADC (IBGE)

Os resultados para salário real reforçam ainda mais essa leitura. O gráfico de renda mostra que, novamente, a deterioração se concentra entre os 18-29 anos: os coeficientes tornam-se mais negativos à medida que o tempo avança, enquanto as trajetórias das coortes mais velhas permanecem próximas de zero ou mesmo levemente positivas em alguns momentos. Em termos econômicos, isso é compatível com um processo no qual a IA não apenas afeta a margem extensiva, como a entrada e permanência no emprego, mas também pressiona a remuneração relativa de perfis mais jovens em ocupações mais expostas. Uma interpretação possível é que a tecnologia esteja reduzindo o valor relativo de tarefas mais padronizadas e rotinizáveis, justamente aquelas que costumam compor a porta de entrada de muitas carreiras administrativas, de apoio e de serviços qualificados. Vale notar que, em períodos mais

recentes, uma tendência negativa também se observa sobre trabalhadores de faixas etárias mais avançadas.

Event-study dos efeitos diferenciais de exposição à IA sobre rendimentos do trabalho por grupo etário



Fonte: Elaboração própria com dados da PNADC (IBGE)

Esses achados dialogam de maneira interessante com a literatura recente. O estudo de referência enfatiza que exposição não equivale automaticamente a substituição, e que parte importante dos efeitos da IA pode vir por reorganização de tarefas e complementaridades. A evidência encontrada é compatível com essa visão, mas acrescenta uma nuance importante: mesmo que o efeito agregado não seja de substituição ocupacional em massa, pode haver efeitos distributivos concentrados em grupos específicos (neste caso, trabalhadores jovens). Isso também conversa com a literatura micro que encontra ganhos de produtividade em tarefas bem definidas, frequentemente maiores para trabalhadores

menos experientes. A nível agregado, porém, esses ganhos podem coexistir com menor demanda por posições de entrada, se parte do trabalho júnior for absorvida por ferramentas de IA ou por trabalhadores mais experientes munidos dessas ferramentas.

Conclusão

O principal resultado deste exercício é que a difusão recente da IA generativa no Brasil parece já estar associada a efeitos diferenciais negativos sobre ocupação e renda entre jovens de 18 a 29 anos, enquanto os efeitos médios sobre trabalhadores de 30 a 44 e de 45 a 59 anos permanecem pequenos e, na maior parte do tempo, compatíveis com zero. Não se trata de afirmar que a IA esteja produzindo um colapso generalizado do emprego. O ponto é mais específico e, justamente por isso, mais relevante: a transição parece estar incidindo de forma desigual sobre o mercado de trabalho, penalizando mais intensamente os perfis que ocupam as posições de entrada e de menor senioridade.

Esses resultados devem ser lidos com cautela. O índice da OIT mede potencial de exposição, não adoção observada em tempo real, e a janela disponível ainda é curta. Ainda assim, os gráficos sugerem que os impactos iniciais da IA não são neutros do ponto de vista etário e distributivo. Se essa leitura estiver correta, a agenda de política pública torna-se bastante concreta: reforçar políticas ativas de mercado de trabalho, requalificação modular e contínua, certificação de competências, proteção durante transições, além de criar mecanismos de governança do uso da IA no ambiente de trabalho, com transparência, auditabilidade e possibilidade de contestação de decisões automatizadas. O próprio estudo de referência enfatiza que os efeitos da IA dependem menos de um determinismo tecnológico e mais do arranjo institucional que regula e distribui seus ganhos.

$$y_{it} = \alpha + X_i' \lambda + \sum_{\tau \neq \tau_0} \beta_{\tau} [E_i \times \mathbb{1}\{t = \tau\}] + \varepsilon_{it}$$

Índice de exposição ocupacional à IA. Partimos de uma classificação de exposição a IA generativa em nível ocupacional, alinhada ao padrão ISCO-08. A referência conceitual é o índice global refinado da OIT, que organiza ocupações em gradientes de exposição com base nas médias e dispersões de escores de tarefas, deixando claro tratar-se de exposição potencial (não impacto realizado) e privilegiando comparabilidade internacional com pesquisas de força de trabalho harmonizadas.

Utilizamos PNADC em cortes repetidos trimestrais de 2022T1–2025T2 (período efetivo usado nos gráficos). Os principais outcomes são: (i) indicador de ocupação (empregado vs. não), e (ii) logaritmo da remuneração do trabalho (deflacionada/real).

Primeiro estágio (exposição predita). Para evitar endogeneidade por realocação ocupacional após a difusão da IA, construímos um grau contínuo de exposição predita a partir de um *snapshot* pré-IA. No 2022T3 (último trimestre plenamente anterior ao lançamento público inicial), estimamos, em amostra de 14+ anos, a probabilidade de um indivíduo estar numa ocupação altamente exposta como função do perfil demográfico e geográfico (idade×sexo, escolaridade, raça/cor, unidade da federação, urbano/rural). Formalmente:

$$\Pr(H_i = 1 | X_i) = \alpha + X_i' \gamma + \delta_{UF} + \delta_{urb/rur} + u_i,$$

e definimos $(\hat{E}_i \equiv \widehat{Pr}(H_i = 1 | X_i))$. (LPM com pesos; resultados do primeiro estágio apenas para construir $(\hat{E}_i)$.)

Evento-estudo com dose contínua. Estimamos, separadamente por coorte etária, um evento-estudo que interage a dose contínua pré-determinada com indicadores de trimestre, tomando um período base (τ_0) pré-IA:

$$Y_{it} = \alpha + \sum_{\tau \neq \tau_0} \beta_{\tau} \cdot \text{big}(E_i^* \times 1_{t = \tau} \cdot \text{big}) + \lambda' X_i + \mu_g + \theta_t + \xi_{c \times t} + \varepsilon_{it}.$$

Aqui, (Y_{it}) é (i) o indicador de ocupação; (X_i) inclui idade×sexo e escolaridade; (μ_g) são fixos de grupos demográficos; (θ_t) são efeitos

fixos de tempo; e $(\xi_{c \times t})$ são efeitos fixos de UF×área (urbano/rural) por trimestre, absorvendo choques locais específicos no tempo. Os erros-padrão são agrupados em células demográficas (idade×sexo). O coeficiente (β) mede a variação diferencial (vs. (τ_0)) na probabilidade de estar empregado por unidade de exposição predita.

Repetimos a especificação substituindo (Y_{it}) por $(\ln(\text{salário real}_{it}))$. A interpretação de $(\beta\tau)$ passa a ser o efeito percentual (aproximação log-linear) na remuneração por unidade de exposição predita, relativamente ao período base.

Correção linear de pré-tendência. Para cada coorte etária, estimamos a tendência prévia da dose contínua sobre o desfecho nos trimestres pré-IA:

$$Y_{it} = \alpha + \theta_{\text{pre}}(E_i^* \times t) + \lambda' X_i + \mu_g + \theta_t + \xi_{c \times t} + \varepsilon_{it} \quad (\text{só pré-IA}).$$

Em seguida, ajustamos os (β) do evento-estudo pós-IA por um *de-trending* linear:

$$\widetilde{\beta}_\tau = \widehat{\beta}_\tau - (\tau - \tau_0), \widehat{\theta}_{\text{pre}},$$

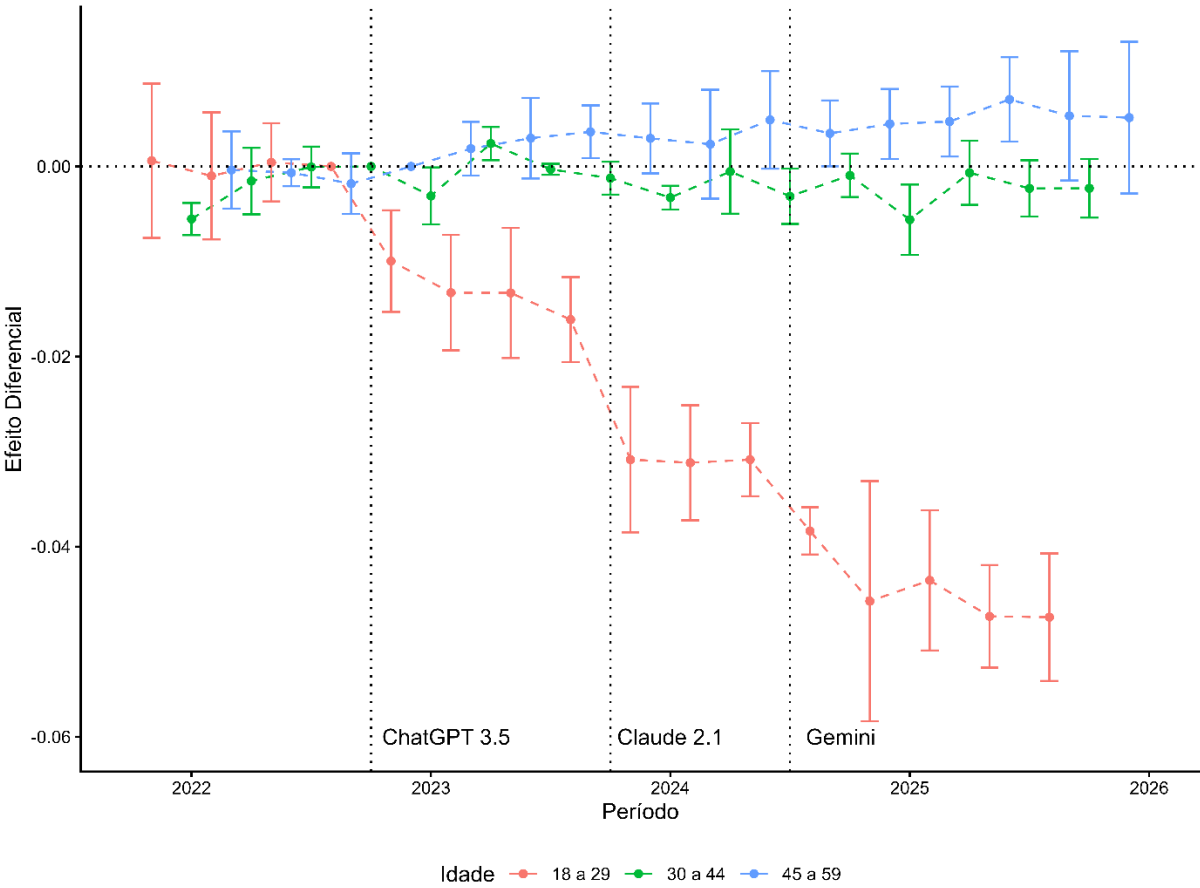
Identificação e leitura dos coeficientes. A estratégia é um DiD dinâmico com dose fixa pré-IA (derivada de 2022T3) e choques temporais comuns, contrastando grupos ex-ante mais expostos com menos expostos. As hipóteses-chave são: (i) tendências paralelas condicionais entre grupos de exposição dentro de células UF×área e perfis demográficos/educacionais; (ii) estabilidade da composição dos grupos (tratada ao fixar a dose com base no *snapshot* pré-IA); e (iii) ausência de choques locais não observados que afetem diferencialmente grupos de alta exposição dentro da mesma célula quando já controlamos por $(\xi_{c \times t})$. Em termos substantivos, ligamos exposição (do índice ocupacional) a ajustes no mercado de trabalho; o índice internacional de exposição é construído sobre tarefas e não é, por si, adoção tecnológica efetiva, conforme enfatiza a OIT.

Relação com o índice da OIT e limitações. O uso de um índice baseado em tarefas/ocupações ISCO-08 oferece comparabilidade e

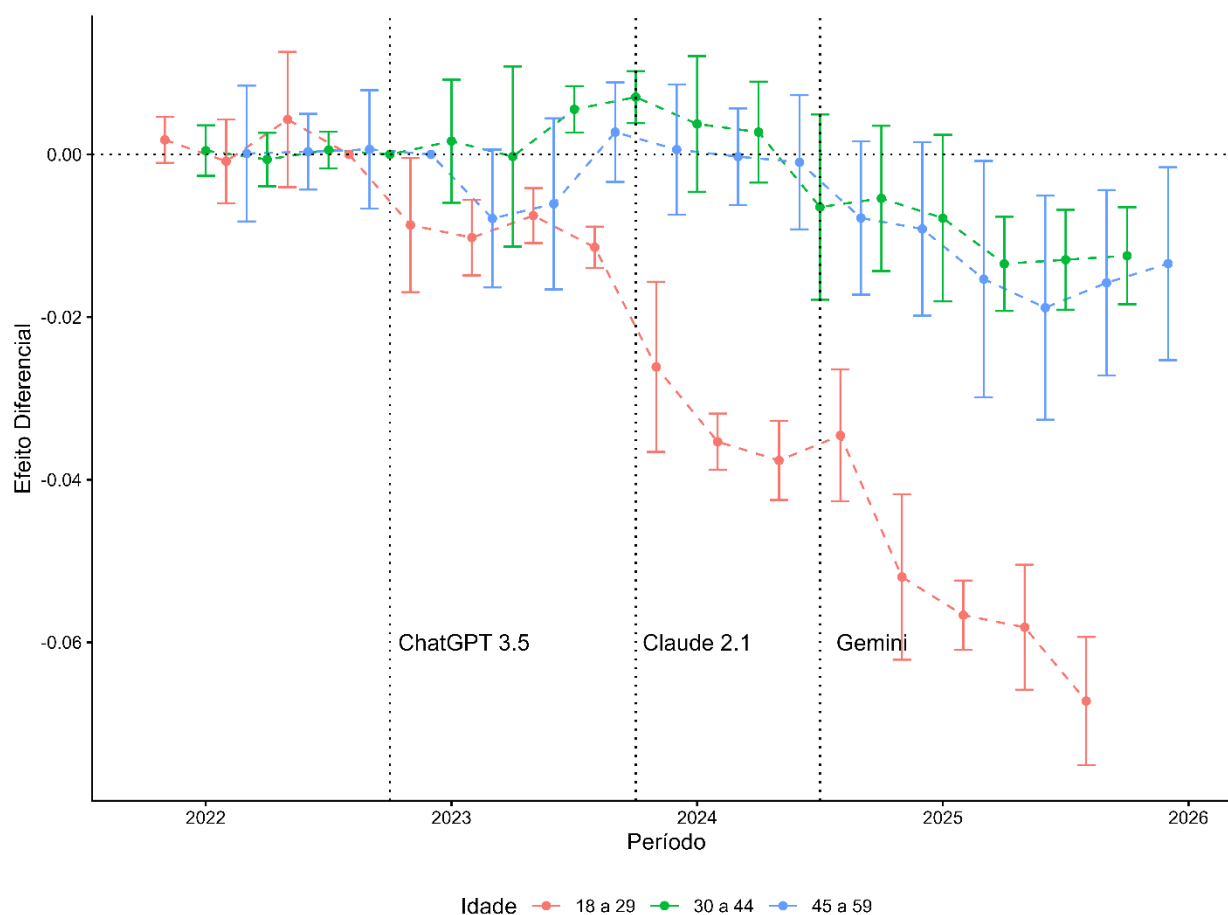
fundamentação metodológica – as classes de exposição derivam de médias e desvios dos escores de tarefas e foram refinadas para melhorar estabilidade e interpretação em relação ao exercício de 2023. Isso permite mapear ocupações brasileiras a um gradiente global, inclusive para exercícios de ligação com microdados nacionais, como fazemos aqui. Ainda assim, o índice mede potencial e “o que a tecnologia pode fazer com tarefas”, não o que já foi feito; vieses de mensuração da dose, heterogeneidade na velocidade de difusão setorial e limitações de infraestrutura/skills significam que estimamos efeitos diferenciais compatíveis com substituição na margem de entrada, não efeitos totais.

Observações finais de implementação. (a) As figuras marcam visualmente três marcos da difusão pública de LLMs (ChatGPT 3.5; Claude 2.1; Gemini); são apenas marcadores de tempo—a identificação não depende desses cortes. (b) Todos os modelos usam pesos da pesquisa, efeitos fixos ricos (incl. UF×área×trimestre), e erros agrupados por células demográficas; (c) estimamos separadamente por coorte etária (18–29, 30–44, 45–59). O achado central é um efeito negativo pronunciado entre jovens (18–29) em ocupação e salários nos períodos pós-IA, coerente com substituição de tarefas de entrada; entre 30–44 e 45–59 os efeitos são pequenos e estatisticamente compatíveis com zero na média pós. Esses resultados são consistentes com o arcabouço de exposição ocupacional baseado em tarefas da OIT.

Ocupação



Salário



As opiniões expressas neste artigo são de responsabilidade exclusiva do autor, não refletindo necessariamente a opinião institucional da FGV.