Early Labor-Market Responses to Generative AI in a Developing Country: Evidence from Brazil*

Tomás Aguirre †   Luis Meloni‡

May, 2026

Abstract

How does generative AI affect labor markets in a developing economy? Using ChatGPT's November 2022 release as a timing shock, we combine Brazil's nationally representative household survey (5.7 million worker-quarter observations, 2017Q1–2024Q4) with the universe of formal-sector hiring and separation records (13 million events). In a difference-in-differences design comparing occupations above and below the employment-weighted median AI-exposure score, before and after ChatGPT's release, workers aged 16–21 in high-exposure occupations earn 4.7% less than otherwise-similar young workers in low-exposure occupations. The coefficient passes the linear pre-trends test, date placebos, and randomization inference that reshuffles AI-exposure scores across occupations. Two supporting difference-in-differences estimates point the same direction but do not clear every robustness check: young workers in high-exposure occupations are 1.5 percentage points less likely to be formally employed (passing pre-trends), and on the firm side, administrative hiring wages for young workers in high-exposure occupations fall 1.3% more than in low-exposure occupations (passing pre-trends), consistent with firms lowering entry-wage terms for young hires. Prime-age (30–54) and older (55+) workers show no differential movement on either wages or formality, with passing pre-trends; the adjustment concentrates on entrants. In contrast to US and UK evidence, where AI effects in this period concentrate on employment quantities for 22–25 year-olds, adjustment in Brazil runs through entry wages and contract terms for new young hires rather than aggregate hiring levels, consistent with prior evidence on automation in Latin America. To our knowledge, we provide the first nationally representative estimates of early generative-AI labor market effects from a developing economy.

Keywords: artificial intelligence; labor markets; generative AI; occupation exposure; Brazil
JEL Codes: J23, J24, O33

*Preliminary work; feedback welcome. We thank Luca Moreno-Louzada, Sam Manning, and Rishi Bommasani for helpful conversations during the development of this paper. All errors remain our own.
†GovAI.
‡University of São Paulo.

1 Introduction

Generative AI is diffusing faster than any previous general-purpose technology. By August 2024, 39% of working-age Americans used generative AI, with 23% using it weekly for work, a pace exceeding personal computer and internet adoption at comparable stages (Bick et al., 2024). This rapid diffusion has intensified concern about labor market disruption, particularly for younger workers. Recent quasi-experimental evidence from the United States and the United Kingdom documents employment declines of 6–8% among workers aged 22–25 in highly AI-exposed occupations, while prime-age and older workers show no significant effects (Brynjolfsson et al., 2025; Klein Teeselink, 2025). Hosseini Maasoum and Lichtinger (2025) document a parallel pattern at the firm level using U.S. LinkedIn résumé data: junior employment falls 7.7–12% in AI-adopting firms relative to non-adopters, driven by reduced hiring rather than separations, while senior employment continues to grow. These patterns, concentrated among labor market entrants and absent among incumbents, suggest that generative AI may be reshaping the terms on which young workers enter the labor market rather than displacing established employees.

Whether these patterns extend to developing countries is unknown. Developing economies differ from high-income settings along dimensions that may fundamentally alter how AI shocks propagate: high informality rates create adjustment margins absent in advanced economies, compressed skill distributions change who is exposed, weaker safety nets raise the stakes of displacement, and slower technology diffusion historically delayed the labor market consequences of earlier automation waves (Bakker et al., 2024; César et al., 2023). At the same time, diffusion of generative AI has been faster than expected even in middle-income countries. Brazil and India rank among the world's top users of ChatGPT and Claude (Liu and Wang, 2024; Appel et al., 2025), despite per capita incomes a fraction of those in the United States. To our knowledge, no prior study has examined aggregate labor market responses to generative AI in a developing country using quasi-experimental methods.

This paper aims to fill that gap, using Brazil as the setting. Brazil combines features that make it an informative case. Its service-oriented economy concentrates employment in precisely the occupations most exposed to generative AI: routine cognitive work in administrative support, sales, and clerical services. At the same time, roughly 35–40% of Brazilian workers operate outside the formal sector (Ulyssea, 2018; Firpo and Portella, 2021), creating an adjustment margin absent in advanced economies. When formal employment contracts, workers may shift into informality rather than unemployment, potentially muting aggregate employment effects while concentrating adjustment costs in wages and contract quality. Brazil also exhibits wage flexibility: minimum wage increases accounted for nearly half of formal sector wage compression between 2001 and 2014 (Engbom and Moser, 2022), suggesting that shocks propagate through wages rather than employment quantities. Despite these structural differences, Brazil ranks among the world's most intensive users of generative AI tools (Liu and Wang, 2024), making its labor market experience directly relevant for other middle-income economies navigating the same transition.

We exploit ChatGPT's November 2022 release as a quasi-experimental shock to the Brazilian labor market. The launch was sudden and unanticipated: OpenAI described it as a "low-key research preview," and the product reached 100 million users within two months (Hu, 2023), a pace that surprised even company leadership (Kahn, 2023). These circumstances support the no-anticipation assumption required for causal identification. We implement a difference-in-differences design that compares occupations with high and low pre-determined AI exposure, measured using task-based scores from Eloundou et al. (2024) mapped to Brazilian occupation taxonomies. Our empirical strategy combines two complementary data sources: PNAD Contínua, Brazil's nationally representative quarterly household survey covering both formal and informal workers, and CAGED, the administrative registry of the universe of formal sector hiring and separation events. Together, these sources span 5.7 million worker-quarter observations (PNAD-C, 2017Q1–2024Q4) and approximately 13 million hiring and separation records (CAGED, October 2020–December 2024, excluding April–September 2021).

Given the Brazilian context, marked by high informality and relatively high adjustment costs in formal labor markets, the introduction of generative AI is likely to affect labor market outcomes through margins that differ from those documented in advanced economies. Generative AI primarily reduces the cost of performing routine cognitive tasks such as drafting text, summarizing information, and producing basic analytical outputs. These tasks are disproportionately performed by junior workers in entry-level administrative and sales occupations, who typically carry out standardized components of broader production processes. If AI substitutes for these tasks, firms may reduce the demand for entry-level labor without necessarily displacing more experienced workers whose tasks involve supervision, judgment, or client interaction. In labor markets with high informality, however, this adjustment need not occur through employment quantities alone. Firms can respond to the reduced value of entry-level tasks by lowering wages, relaxing hiring requirements, or shifting workers from formal to informal arrangements. In this setting, technological shocks may reshape the terms on which young workers enter the labor market rather than generating large employment losses among incumbent workers.

Our central finding is a 4.7% relative wage decline for young workers (16–21) in high-exposure occupations after ChatGPT's release, estimated in a difference-in-differences design that compares occupations above and below the employment-weighted median AI-exposure score. The coefficient passes the linear pre-trends test, the Benjamini–Hochberg false discovery rate correction at the individual-coefficient level, strengthens under adjacent-quartile comparisons, passes five date placebo tests, and yields a randomization-inference p-value of 0.010 on the exposure vector. The workers bearing this decline are predominantly high-school graduates in entry-level administrative and sales positions, earning on average R$1,509 per month before treatment; the implied gap amounts to roughly R$852 annually. Prime-age (30–54) and older (55+) workers show no differential movement on either wages or formality, with passing pre-trends; the adjustment concentrates on entrants.

Two supporting findings move with this central pattern but do not survive the same battery of checks. Young workers in high-exposure occupations are 1.5 percentage points less likely to be formally employed, with parallel pre-treatment trends, consistent with residual tasks in AI-exposed positions no longer justifying formal employment relationships. Administrative hiring records from CAGED point in the same direction on the firm side: hiring wages are 1.3% lower for young workers (16–21) in high-exposure occupations (p=0.086, pre-trends pass) and 2.6% lower for the 22–29 group (p=0.002, pre-trends pass); the youngest CAGED coefficient is marginal on conventional thresholds while the early-career one is significant. Both address the concern that the household-survey result reflects compositional shifts rather than wage changes for similar workers. Neither clears every robustness check; we treat them as supporting evidence that points the same way as the central finding rather than as independently identified results (Section 4.3). Public sector workers, who face similar AI exposure but stronger employment protections, show no differential movement across any outcome.

Additional margins move directionally with these three findings but do not clear our robustness protocol (Section 4.3): aggregate formality, voluntary quits in the main CAGED specification, schooling of new hires in the 22–29 age bin, and the aggregate wage null (which date placebos at 2022Q1–Q2 make uninformative). We describe these patterns where relevant but do not lean on them as headline evidence.

Taken together, these patterns are consistent with a barrier-to-entry dynamic in which firms in high-exposure occupations adjust the terms on which they hire new workers rather than displacing existing employees. We read this as the central documentary contribution of the paper. A set of interpretive caveats applies throughout. Because spillovers between high- and low-exposure occupations are likely (through labor supply reallocation and product-market competition), the difference-in-differences estimator recovers a *relative* contrast rather than absolute impacts on either group: if low-exposure occupations also adjust to indirect AI exposure, the patterns we report understate aggregate effects. ChatGPT's release also coincides with Brazil's October 2022 presidential election, creating a timing confound we cannot fully resolve; the age-specific wage pattern, which replicates on administrative hiring-wage records, is difficult to attribute to the political transition, but not impossible. The patterns remain consistent across three independent AI-exposure measures and the battery of robustness checks in Section 3.3.

The Brazilian evidence differs from what studies in advanced economies have documented. In the US and UK, effects concentrate on employment quantities among 22–25 year-olds (Brynjolfsson et al., 2025; Klein Teeselink, 2025), with parallel firm-level evidence that adjustment runs through reduced junior hiring rather than layoffs (Hosseini Maasoum and Lichtinger, 2025); Danish register data, by contrast, show small earnings and hours effects in exposed occupations one to two years after ChatGPT's release (Humlum and Vestergaard, 2025). Here, patterns appear on wage and contract-quality margins among younger, less-educated entrants, with cross-source corroboration on young hiring wages in administrative records. This difference is consistent with Brazil's institutional structure, where informality and wage flexibility absorb shocks that higher-income labor markets would translate into unemployment, and is one of the paper's substantive contributions.

Throughout, we treat the DiD estimates as documentary rather than causal, for three reasons specific to this setting. First, we do not directly observe AI adoption at the firm or worker level: occupation-level exposure scores are a proxy for treatment intensity, so within-occupation heterogeneity in actual adoption is absorbed into the error term, and we cannot distinguish occupations that were intensively exposed to ChatGPT from those with high theoretical exposure but little take-up. Second, spillovers between high- and low-exposure occupations through labor-supply reallocation and product-market competition are likely, so the DiD coefficients identify only a *relative* contrast between groups: even under the most favorable identifying assumptions, the estimands we report are differences between high- and low-exposure occupations, not absolute effects on either. Third, the pre-trends picture is mixed: the headline outcomes pass the linear pre-trend test in the main specification, but several related outcomes fail it (Appendix A.3.1), and Rambachan–Roth sensitivity bounds widen the individual-coefficient intervals quickly (Appendix A.3.2). We therefore report what these patterns look like, subject them to a disciplined scaffolding of pre-trends, placebos, sensitivity bounds, adjacent-quartile comparisons, and three independent exposure measures, and distinguish patterns that survive these checks from those that do not. Sharpening the inference would require data we do not have here: matched employer–employee records (RAIS), direct firm-level adoption measures (most plausibly constructed from job-posting text), and longer post-treatment horizons.

The informality margin is central to understanding why Brazilian labor markets absorb AI shocks differently than advanced economies. In the United States and United Kingdom, displaced workers face a stark choice between employment and unemployment, so AI-driven reductions in labor demand manifest primarily as employment declines (Brynjolfsson et al., 2025; Klein Teeselink, 2025). In Brazil, a third margin exists: workers displaced from formal positions can transition into informal employment rather than unemployment (Ulyssea, 2018; Meghir et al., 2015). This absorptive capacity attenuates aggregate employment effects while concentrating adjustment costs in contract quality and wages. Prior evidence on robots in Latin America documents exactly this pattern: César et al. (2023) find that robot exposure increased informality by 0.23 percentage points across Argentina, Brazil, and Mexico, with formal jobs and young workers bearing the brunt of adjustment. The 1.5 percentage point formality decline we document for young workers in high-exposure occupations is consistent with this channel operating through generative AI rather than industrial automation.

The wage flexibility of Brazil's labor market reinforces this mechanism. Engbom and Moser (2022) document that minimum wage increases accounted for nearly half of formal sector wage compression between 2001 and 2014, illustrating that shocks in Brazil propagate through wages rather than employment quantities. When AI reduces the value of tasks performed by entry-level workers, firms can lower wages for new hires rather than eliminating positions entirely, particularly where informal arrangements weaken the effective floor imposed by minimum wage legislation (Parente, 2024). The 4.7% wage decline among young workers is consistent with this adjustment: firms retain entry-level positions but adjust entry-wage terms for young hires, as AI substitutes for the cognitive tasks that previously justified higher compensation.

This paper makes three contributions. First, to our knowledge we provide the first nationally representative estimates of early generative-AI labor-market effects from a developing economy. Existing quasi-experimental work is concentrated in advanced economies such as the United States, United Kingdom, and Denmark (Brynjolfsson et al., 2025; Klein Teeselink, 2025; Hosseini Maasoum and Lichtinger, 2025; Humlum and Vestergaard, 2025); developing-country work has so far calculated exposure measures without studying post-release outcomes (Pizzinelli et al., 2023; Bakker et al., 2024), examined pre-generative AI technologies (de Souza, 2025), or studied specific platforms rather than national labor markets (Hui et al., 2024).

We combine nationally representative household survey data covering both formal and informal workers with the universe of formal-sector hiring records, a data combination that lets us trace adjustment across margins invisible in administrative data alone.

Second, we add to the literature on how labor-market institutions mediate technological shocks (Acemoglu and Restrepo, 2019; César et al., 2023). Prior work on robots and automation in Latin America documents that informality absorbs displaced formal workers, concentrating adjustment in wages rather than unemployment (César et al., 2023; Stemmler, 2023). The patterns we document for generative AI are consistent with the same channel: a formality-rate gap and wage gaps for young workers rather than the employment-quantity effects documented in advanced economies. We view this as the first empirical surfacing of the informality margin as a candidate adjustment channel for generative AI specifically, a channel that is largely absent from existing AI and labor-market studies.

Third, we add to a growing literature on AI's distributional consequences across worker types (Brynjolfsson et al., 2025; Hosseini Maasoum and Lichtinger, 2025; Autor and Thompson, 2025; Garicano and Rayo, 2025). US and UK evidence concentrates effects among 22–25 year-old entrants in exposed occupations and among junior employees in AI-adopting firms (Brynjolfsson et al., 2025; Hosseini Maasoum and Lichtinger, 2025); the Brazilian patterns instead concentrate among 16–21 year-old high-school graduates in routine cognitive positions. The young-worker hiring-wage decline, which replicates the household-survey finding on administrative records, is consistent with the task-unbundling channel in Autor and Thompson (2025): if AI substitutes for tasks previously performed by junior workers, firms can adjust entry-wage terms without displacing incumbents. We document this barrier-to-entry pattern, rather than claim that it is causally identified, and we report it as a candidate general feature of how labor markets absorb generative AI shocks across different institutional settings, to be tested as more data accumulates.

The remainder of the paper is organized as follows. Section 2 describes Brazil's labor market institutions and the political and economic context of 2022–2023. Section **??** describes the data and construction of AI exposure measures. Section 3 presents the difference-in-differences framework and identification assumptions. Section 4 reports results on wages, employment, formality, and hiring composition. Section 5 concludes.

2 Institutional Background

2.1 Brazil's Labor Markets

Brazil's labor market is characterized by a meaningful divide between formal and informal employment. Formal workers hold a signed contract that entitles them to statutory protections including severance contributions, unemployment insurance, paid annual leave, and access to the social security system. These protections make formal employment relationships relatively costly to dissolve. Firms must contribute 8% of monthly wages to the *Fundo de Garantia do Tempo de Serviço* (FGTS), a severance fund accessible to workers upon dismissal, and when terminating a worker without cause must additionally pay a fine of 40% of the total accumulated FGTS balance. Informal arrangements, by contrast, can be dissolved at low cost to either party. As a result, informal workers bear the full cost of job loss without access to unemployment insurance or severance, and turnover is substantially higher: informal employment spells last on average one year, compared to four and a half years for formal jobs (Menezes-Filho and Narita, 2025). Roughly 35–40% of Brazilian workers operate outside the formal sector (Ulyssea, 2018; Firpo and Portella, 2021).

This asymmetry in adjustment costs shapes how technological shocks propagate through the labor market. When demand falls for a given type of worker, firms face strong incentives to reduce hiring of new formal workers rather than dismiss existing ones, and displaced workers who cannot find formal employment transition into informality rather than unemployment. The result is that aggregate employment effects may be dampened while adjustment concentrates in wages, contract quality, and the composition of new hires. Brazil's wage structure reinforces this pattern. Minimum wage increases accounted for nearly half of formal sector wage compression between 2001 and 2014 (Engbom and Moser, 2022), and the minimum wage serves as a reference wage even in the informal sector where it is not legally binding (Parente, 2024). When labor demand falls, wages for new entrants can adjust downward in both sectors, amplifying the wage margin of adjustment relative to what would be observed in economies with stronger wage floors or lower informality.

2.2 The Political and Economic Context of 2022–2023

Two macro shocks sit close enough to ChatGPT's November 2022 release that we cannot cleanly separate them from it, and we want to be upfront about that before describing what we can say. The first is the October 2022 presidential election, in which Luiz Inácio Lula da Silva was elected, with the transition falling in the same quarter as ChatGPT's release. The second is the Banco Central do Brasil's tightening cycle, which raised the Selic rate from 2% in early 2021 to 13.75% by August 2022 and held it there through mid-2023. Our placebo and pre-trends exercises relieve part of the worry, but neither exercise can rule these shocks out, and they remain an alternative interpretation of the patterns we document.

Taking the political channel first, the Lula administration did not enact a labor market reform in 2022Q4 or 2023 that would plausibly operate differentially across AI exposure. The most salient early policy was a minimum wage increase enacted in May 2023, raising the monthly floor from R$1,302 to R$1,320, a 1.4% adjustment in line with inflation and with the pattern of annual adjustments from preceding years (Assis and Campos, 2023). The Bolsa Família cash transfer program maintained the R$600 monthly benefit introduced under the previous administration, with no change in transfer values during our post-treatment window. No major labor regulation reform, tax reform, or income policy change was enacted in the relevant window. If a policy shift earlier in the transition were moving outcomes through expectations, we would expect it to show up across wage levels and age groups rather than concentrate on teenagers and young adults in AI-exposed occupations; minimum wage adjustments in particular bind in the lower part of the wage distribution across occupations and ages, not selectively on young workers in high-exposure jobs. Our date placebos for young worker wages (Table A.13, Panel D) also fail to reject at any of the five fake treatment dates, which is what we would expect if a gradual political expectations channel were the driver and is not what we observe.

The monetary channel is a tighter concern. The tightening cycle predates ChatGPT and could plausibly affect hiring with a lag, and AI-exposed occupations are concentrated in services and professional employment that are not obviously insulated from interest rate sensitivity. For the United States, Brynjolfsson et al. (2026) report that AI exposure is, if anything, negatively correlated with interest rate sensitivity, which works against the monetary confound as an explanation of their patterns; whether the same holds in Brazil is an open question we do not answer here. What we can say is that the Selic cycle affects young and older workers in the same occupation through the same hiring channels, so a monetary explanation has to reconcile with the age concentration of our wage results and with their replication on the firm side in CAGED hiring wages (Table 3). These facts do not eliminate the Selic as a confound, but they narrow the set of stories under which it is doing the work we attribute to AI exposure.

We flag both channels again in Section 3.4 and read our coefficients as evidence on relative movements across exposure levels rather than as aggregate causal effects. Readers who remain unconvinced by the political or monetary interpretation should still find the heterogeneity patterns and the placebo evidence informative, but we do not claim to have resolved either concern.

2.3 Generative AI Adoption in Brazil

Despite being a middle-income economy, Brazil ranks among the world's most intensive users of generative AI tools. Brazil and India are among the top countries by ChatGPT usage (Liu and Wang, 2024), and Brazil accounts for a disproportionate share of Claude usage relative to its GDP per capita (Appel et al., 2025). Figure 1 shows Google Search interest for "ChatGPT" and related AI terms in Brazil across quarters, illustrating that search interest rose sharply after the November 2022 release and continued to climb through 2024–2025, consistent with the timing of our treatment and with gradual rather than instantaneous diffusion. This adoption intensity makes Brazil's labor market experience directly relevant for understanding AI's effects in other emerging economies.

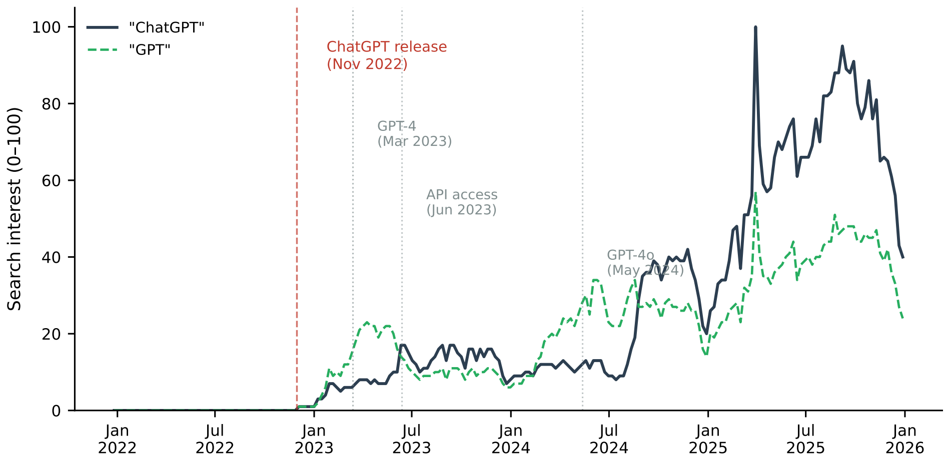

Figure 1: Google Search Interest for ChatGPT and Generative AI in Brazil

Notes: Google Trends search interest (normalized 0–100) in Brazil, by quarter, for the search terms "ChatGPT" and "GPT." The vertical dashed line marks ChatGPT's release (2022Q4). Data: Google Trends.

Adoption is concentrated among educated and higher-income workers. Workers holding bachelor's or graduate degrees adopt generative AI at roughly twice the rate of those without (Bick et al., 2024). This pattern is important for interpreting our results: the young workers who bear the wage and formality costs we document are not primarily AI users themselves. They are workers whose tasks have become partially substitutable by AI tools used by others, including employers, clients, and more educated colleagues. The channel is not adoption by the affected workers but rather the reduced demand for entry-level cognitive tasks that AI makes substitutable from above.

Adoption also varies substantially across Brazilian regions, reflecting differences in digital infrastructure, educational attainment, and occupational composition. The South and Southeast regions, which concentrate professional and administrative occupations, show considerably higher AI exposure than the North and Northeast. ca

3 Empirical Strategy

3.1 Research Design

We exploit ChatGPT's November 2022 release as a quasi-experimental shock to estimate the effects of generative AI on Brazilian labor market outcomes. The identifying variation comes from the interaction between the timing of the shock and the pre-determined AI exposure of each occupation, measured through task content assessed before ChatGPT existed (Eloundou et al., 2024). Occupations with higher exposure to AI-automatable cognitive tasks serve as the treatment group; occupations with lower exposure serve as the comparison. Because exposure is fixed before the shock, it cannot have been influenced by anticipation of ChatGPT's arrival. The no-anticipation assumption is further supported by the historical record: OpenAI described the launch as a "low-key research preview," and company leadership expressed surprise at its reception (Kahn, 2023; Wiggers, 2023).

We focus on five primary outcomes spanning both household survey and administrative data. From PNAD-C, we examine log wages, weekly hours, and formal employment status, which together capture whether AI exposure shifts the returns to labor, the intensity of work, and the contractual terms under which workers are employed. From CAGED, we examine the average education of new hires and the rate of voluntary quits. Education of new hires measures whether firms adjust skill requirements at the point of hiring, a direct test of the deskilling channel; voluntary quits measure whether workers in AI-exposed occupations become less willing to leave their current jobs, consistent with deteriorating outside options. Together, these five outcomes map onto the adjustment margins that Brazil's institutional structure makes most relevant: wages and contract quality rather than aggregate employment, as discussed in Section 2. We also test for heterogeneity by worker age, following evidence from the United States and United Kingdom that AI's early labor market effects concentrate among labor market entrants rather than incumbents (Brynjolfsson et al., 2025; Klein Teeselink, 2025).

One design choice in what follows deserves an upfront note. The young-worker bin in this paper is 16–21, not the 22–25 convention used in recent US and UK studies (Brynjolfsson et al., 2025; Klein Teeselink, 2025). The choice reflects the Brazilian institutional setting rather than a specification-search preference. Brazilian first-time formal-sector entry is concentrated among 16–21 year-olds, most of whom are high-school graduates entering routine cognitive occupations in retail, clerical, and cashier work; in our pre-treatment PNAD sample, the 16–21 group is 59% high-school graduates, 58% formally employed, with modal occupations in retail (17.3% retail salespeople, 15.0% general office clerks). Informality is highest among the youngest workers and declines with age, so the formality margin central to the paper's mechanism is most active in the 16–21 bin. The US literature's 22–25 convention captures college-graduate entrants and reflects US educational timing; the institutional analog in Brazil is 16–21. The 22–29 bin in our results shows directionally consistent but smaller coefficients (formality −1.3 percentage points, significant; a small negative wage coefficient), the attenuation-with-age gradient expected under a barrier-to-entry story.

3.2 Specification

The baseline specification is a two-group difference-in-differences estimated at the individual level for PNAD-C outcomes and at the occupation-month level for CAGED outcomes (aggregating across states; state-disaggregated specifications are reported as robustness in Appendix A.4.1):

$$Y_{ijt} = \alpha_j + \lambda_t + \gamma_s + \beta(\text{HighExposure}_j \times \text{Post}_t) + \varepsilon_{ijt} \tag{1}$$

where i indexes individuals, j indexes occupations, t indexes periods, and s indexes states. Occupation fixed effects (α_j) absorb permanent differences across occupations, period fixed effects (λ_t) absorb aggregate shocks common to all workers, and state fixed effects (γ_s) absorb persistent regional differences in labor market conditions. The treatment indicator HighExposure_j equals one if the occupation's AI exposure score exceeds the employment-weighted median (threshold 0.233), yielding 264 high-exposure and 170 low-exposure occupations. Post_t equals one from 2023Q1 onward for PNAD-C and from November 2022 onward for CAGED. Because ChatGPT launched on November 30, 2022, 2022Q4 contains only one month of meaningful post-launch exposure (December); we retain it in the pre-period for PNAD-C and use 2022Q4 as the reference period for event studies.[1] The coefficient β captures

[1]Because ChatGPT launched on the last day of November, November 2022 effectively has zero days of post-launch exposure; within 2022Q4 only December is meaningfully post-launch. The first fully-treated PNAD quarter is 2023Q1, giving 8 post-treatment quarters through 2024Q4. For CAGED, which

the differential change in outcomes for high-exposure occupations relative to low-exposure occupations following ChatGPT's release.

The main specification is uncontrolled: adding covariates would require stronger assumptions without changing the results. Appendix A.4.1 reports a controlled specification with safe demographics (sex, race, age-bin dummies); estimates are substantively unchanged. Standard errors are clustered at the 4-digit occupation level (434 clusters), the level at which treatment varies (Roth et al., 2023); 2-digit clustering (39 clusters) is reported in Appendix A.4.

We aggregate CAGED to the occupation-month level because treatment varies at the occupation level (621 CBO codes) and our outcomes are either flows (log hires, log quits, log dismissals) or cell-level composition means (average years of schooling, average hiring wage) that are naturally defined at this aggregation. Event-level regressions for composition outcomes are feasible but would conflate shifts in hiring volume with shifts in hiring composition, and would inflate cell counts without adding identifying variation beyond the occupation panel. State-disaggregated specifications (state $\times$ occupation $\times$ month) are reported as robustness in Appendix A.4.1; their tighter standard errors reflect sample-size inflation rather than additional identifying variation.

To test parallel trends and trace the dynamic evolution of treatment effects, we estimate an event study specification:

$$Y_{ijt} = \alpha_j + \lambda_t + \gamma_s + \sum_{\tau \neq -1} \beta_\tau (\text{HighExposure}_j \times D_t^\tau) + \varepsilon_{ijt} \tag{2}$$

where D_t^τ is an indicator for event time τ relative to ChatGPT's release, with $\tau = -1$ (2022Q4, the last pre-treatment quarter) as the omitted reference period. Pre-treatment coefficients (β_τ for $\tau < 0$) test whether both groups followed similar trajectories before the shock; post-treatment coefficients (β_τ for $\tau \geq 0$) trace the dynamic evolution of effects.

is monthly, Post_t equals one from November 2022 onward, treating the release month itself as the first post-treatment period; this is a conservative choice that attenuates estimated effects but avoids misclassifying any potentially treated observations as pre-treatment. The two data sources therefore date the same shock differently: PNAD's first post-treatment quarter is 2023Q1, while CAGED's first post-treatment month is November 2022. The asymmetry follows mechanically from the quarterly-versus-monthly frequency rather than from a substantive choice; making the conventions symmetric in either direction (treating CAGED Nov–Dec 2022 as pre, or treating PNAD 2022Q4 as post) would discard data without changing any sign in the headline results.

3.3 Identification

Causal identification rests on the parallel trends assumption: absent ChatGPT's release, high-exposure and low-exposure occupations would have evolved similarly. Because occupation exposure is determined by task content measured through O*NET before ChatGPT existed, it cannot have been influenced by anticipation of the shock. We assess the plausibility of parallel trends through pre-treatment event study coefficients and date placebo tests using five fake treatment dates before November 2022.

The difference-in-differences estimator identifies a relative effect: the differential change in high-exposure occupations compared to low-exposure ones. Because spillovers between occupation groups are likely through labor supply and product market channels, the estimator does not recover the absolute causal impact on either group in isolation.[2] This relative estimand remains policy-relevant since it shows which workers bear the adjustment costs of AI diffusion, even if spillovers prevent isolating absolute effects. The PNAD-C analysis conditions on employment and therefore estimates intensive margin effects for workers who remain employed; CAGED hiring and separation flows complement this by capturing extensive margin adjustment.

3.4 Scope of Claims

Before turning to specific threats to identification, we state explicitly what we ask difference-in-differences to do in this paper. DiD combines cross-occupation variation in exposure with aggregate time variation and absorbs common shocks through period fixed effects; relative to plotting trends, it provides an explicit framework within which pre-trends, placebos, and sensitivity bounds are well defined. At the same time, parallel trends is not a property that can be verified directly, only one that can fail to be rejected, and the Wald tests we report have limited statistical power against plausible violations (Roth, 2022). We therefore read the coefficients in this paper as evidence about relative movements between high- and low-exposure occupations, rather than as estimates of the aggregate causal effect of generative AI on Brazilian labor markets.

Three limitations push us toward this posture. First, the October 2022 election coincides with ChatGPT's release and produces a timing confound we cannot fully resolve. Second, we do not observe AI adoption at the firm or worker

[2] Under the identifying assumptions but allowing for spillovers, the DiD estimator recovers $\hat{\beta} = \tau_{\text{high}} + \delta_{\text{high}} - \delta_{\text{low}}$, where τ_{high} is the direct treatment effect on the high-exposure group and $\delta_{\text{high}}, \delta_{\text{low}}$ are the spillover effects onto each group (Huber and Steinmayr, 2021).

level, so exposure-based variation is a proxy for treatment, not a measurement of it: within-occupation heterogeneity in actual adoption is absorbed into the error term. Third, spillovers between high- and low-exposure occupations through labor supply reallocation and product market competition are likely, which means the DiD estimator recovers a relative effect rather than an absolute one (Huber and Steinmayr, 2021). Any of these would be sufficient on its own to caution against a causal reading. Taken together, they imply that the paper's contribution is documentary and methodological: we document patterns and we report the scaffolding of robustness checks that distinguish patterns that survive plausible perturbations from patterns that do not.

We therefore adopt the following reading protocol, which Section 4 and the appendix implement. A pattern is treated as robust if it (i) passes the linear pre-trends test in the main specification, (ii) replicates in sign across at least two of the three AI-exposure measures we use (Eloundou, Anthropic Observed Exposure, Microsoft Copilot applicability), (iii) does not flip under narrower adjacent-quartile comparisons (Appendix A.4.5), and (iv) survives randomization inference on the exposure vector (Appendix A.4.6). A pattern is treated as suggestive if it clears some but not all of these hurdles. Rambachan–Roth sensitivity bounds (Appendix A.3.2) provide an additional, and deliberately demanding, check: we report them but do not require any coefficient to survive $\bar{M} \geq 1$ for us to discuss the underlying pattern.

3.5 Threats and Robustness

The main identification threat is the coincidence of ChatGPT's November 2022 release with Brazil's October 2022 presidential election. As documented in Section 2, early Lula administration policies were modest and did not differentially affect high-AI-exposure occupations. The pattern of results further supports an AI rather than political interpretation: the wage penalty concentrated among workers aged 16–21 with null effects for older workers is inconsistent with minimum wage or regulatory changes, which affect low-wage workers across all age groups; the same concentration of wage effects on young workers replicates on the firm side in CAGED hiring wages (Table 3), which has no natural connection to the political transition. Date placebo tests for young worker wages pass at all five fake treatment dates (p-values between 0.136 and 0.446; Table A.13 Panel D). We cannot rule out this confound entirely, and it remains the primary caveat even for the relative pattern we document.

Three additional threats warrant acknowledgment. First, differential COVID-19 recovery could confound results if AI-exposed occupations recovered differently from the pandemic. The CAGED pre-period begins in October 2020, after the initial pandemic shock had subsided and employment had largely recovered (Firpo and Portella, 2024), and remote work in Brazil peaked at only 10.4% during the pandemic (Barbosa Filho et al., 2022), limiting scope for differential recovery. Outcomes that pass pre-trends tests are unlikely to reflect lingering pandemic dynamics. Second, date placebo tests for aggregate wages at 2022Q1 and 2022Q2 show significant effects (p=0.007, p=0.009), suggesting some wage divergence opened up before ChatGPT's release. Figure 2 plots the quarter-by-quarter event-study coefficients of the aggregate-wage gap across the full pre-period (2017Q1–2022Q4), with the August 2022 Selic peak marked, so readers can judge whether the divergence concentrates near the monetary tightening cycle or is spread across the pre-period. We interpret the aggregate wage null cautiously and note that young-worker wage results pass all five date placebos.

figures/pre_period_wage_event_study.pdf

Figure 2: Pre-Period Aggregate Wage Divergence

Notes: Quarter-by-quarter event-study coefficients on the interaction of high AI exposure (median split) with quarter dummies, restricted to the pre-treatment window 2017Q1–2022Q4. Outcome: log monthly wage from PNAD individual-level data. Reference period: 2022Q4 (last pre-treatment quarter, Convention A; matches the rest of the paper). Solid markers: significant at 5%; hollow markers: not significant; reference period shown as a hollow marker at zero. Shaded band: 95% confidence interval. Vertical dashed line: peak Selic rate (Aug 2022). Standard errors clustered at occupation level.

Third, Brazil's aggressive monetary tightening cycle (Selic rate from 2% in early 2021 to 13.75% by August 2022) could confound identification if interest rate-sensitive sectors overlap with AI-exposed occupations. Brynjolfsson et al. (2026) find that in the United States the correlation between AI exposure and interest-rate sensitivity is negative: high-AI-exposure occupations (administrative, sales, customer support) and high-IR-sensitivity occupations (construction, durable-goods retail) sit on different parts of the occupational distribution. Whether the same negative correlation holds in Brazil is an open question we examine in Appendix A.5.3; if it does, the Selic confound is implausible because the monetary cycle would attenuate, not amplify, the wage gap we measure between high- and low-AI-exposure occupations.

Robustness checks include: alternative AI exposure measures from the Anthropic Observed Exposure construction (Massenkoff and McCrory, 2026), which builds on the Anthropic Economic Index Claude-usage data (Handa et al., 2025), and Microsoft Copilot applicability (Tomlinson et al., 2025); alternative treatment cutoffs (tercile, quartile, quintile); specifications with demographic controls; randomization inference by permuting exposure across occupations; and the Benjamini–Hochberg false discovery rate correction (Benjamini and Hochberg, 1995) applied across all primary outcomes. Results are reported in Appendix A.4.

4 Results

We present results in two parts: household survey evidence on wages, hours, and formality (Section 4.1), and administrative hiring records on employment flows and composition (Section 4.2). Section 4.3 synthesizes findings across both sources and separates the patterns we treat as robust from those we treat as suggestive.

4.1 Household Survey Evidence

Aggregate Results. Panel A of Table 1 shows no significant differential movement in aggregate wages (-0.0090, p=0.459) or weekly hours (0.1718, p=0.157) between high- and low-exposure occupations, with parallel pre-treatment trends for both outcomes. Date placebo tests fire at 2022Q1 and 2022Q2 for aggregate wages (p=0.007, p=0.009; Appendix A.4.2), so we do not read the post-2022Q4 null as identifying. The formality rate is 0.9 percentage points lower in high-exposure relative to low-exposure occupations after ChatGPT's release (p=0.031, pre-trends p=0.298 under the aggregated F-test used in the main text; p= under the individual-level linear trend used in appendix robustness); the coefficient does not survive the adjacent-quartile test (Q4 vs. Q3 sign-flips to $+0.005$, Table A.17), and we flag it as suggestive rather than robust in Section 4.3.

Age Heterogeneity. The aggregate null on wages masks heterogeneity by age. Panel B shows that workers aged 16–21 in high-exposure occupations see wages 4.7% lower than young workers in low-exposure occupations following ChatGPT's release (p=0.001, pre-trends p=0.116). At the mean monthly wage of R$1,509 for young workers in high-exposure occupations, this corresponds to roughly R$71 per month or R$852 annually. Young workers are also 1.5 percentage points less likely to be formally employed (p=0.013, pre-trends p=0.509), and early career workers (22–29) show a similar formality gap of 1.3 percentage points (p=0.019, pre-trends p=0.863). Prime-age (30–54) and older (55+) workers show no significant differential movement on wages or formality, with passing pre-trends. The early-career (22–29) wage coefficient is small and not statistically distinguishable from zero, and its pre-trend fails (p=0.038), so we do not read it as identifying. This age gradient, concentrated among entrants and absent among incumbents, is consistent with a barrier-to-entry dynamic in which firms adjust the terms on which they hire new workers rather than displace existing employees.

This pattern contrasts with US evidence: Brynjolfsson et al. (2025) find employment declines of around 6% for young workers in the most AI-exposed occupations but little difference in wage trends. The Brazilian pattern is consistent with more flexible labor markets and the informality margin discussed in Section 2: when labor demand falls, firms can reduce wages and shift toward informal contracts rather than eliminating positions.

The cross-age comparison that motivates the barrier-to-entry reading rests on a weaker footing than the within-young-worker coefficient. A triple-difference specification, which is the design that directly tests whether young workers in high-exposure occupations move differently than older workers in the same occupations, fails its own pre-trends test (p=0.017; Appendix A.5.2). The young-older wage gap in high- versus low-AI occupations was already evolving differentially before ChatGPT's release, so the post-treatment age gradient cannot be attributed to AI under the unified identifying assumption that the triple difference imposes. The age-stratified regressions in Panel B require parallel trends only *within* each age bin, which the data support for ages 16–21 and 22+ pooled, but they do not absorb the cross-age drift. We therefore read the age gradient as documentary and treat the within-young-worker wage coefficient as the credibly identified piece of the age story.

The young workers bearing these costs are concentrated in entry-level occupations: retail salespeople account for 17.3% of the 16–21 age group, followed by general office clerks (15.0%), cashiers (5.9%), and receptionists (5.0%). They average 19 years old, are majority female (53%), mostly high school graduates (59%) working in their first formal job (58% formal employment). These are teenagers and young adults in routine cognitive positions where the task content is consistent with AI substitution (e.g., drafting customer responses, scheduling, and data entry), though we do not directly observe AI adoption at the worker or firm level. The wage penalty concentrates in occupations with low pre-existing digital intensity rather than digitally intensive ones (Appendix A.5.4), suggesting effects are not confined to workers who interact directly with computers.

Figure 3 plots the event study for young worker wages. The estimates are noisy but suggest a break after 2022Q4: pre-treatment coefficients fluctuate around zero throughout 2017–2022, and the formal pre-trends test passes (p=0.116).[3] Post-treatment point estimates remain negative, though the confidence intervals widen, so the pattern should be read as supportive rather than conclusive.

Table 1: Main DiD Results

	Log Wage (1)	Weekly Hours (2)	Formal Employment (3)
High Exposure × Post	-0.009	0.172	-0.009**
	(0.012)	(0.121)	(0.004)
Pre-trends status	Pass	Pass	Pass
Pre-trends p-value	[0.103]	[0.268]	[0.298]
Observations	5,682,355	6,061,106	3,899,812
R^2	0.427	0.127	0.264

All specifications include occupation, period, and state fixed effects. Cluster-robust standard errors (occupation level) in parentheses. Pre-trends test: joint significance of exposure × pre-treatment period interactions; $p \geq 0.05$ indicates parallel trends assumption supported. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

Employment Type. Table 2 examines heterogeneity by employment type. Public sector workers show null effects across all outcomes with passing pre-trends, consistent with job security insulating them from market adjustment despite similar AI exposure levels. This is informative: it suggests that institutional protections, rather

[3]Two constructions of the pre-trends test appear in this paper and agree qualitatively on all headline outcomes. The main-text citations, event-study figures, and headline appendix tables (Tables 1, A.3, A.13) use the aggregated F-test on pre-treatment event-study coefficients from the occupation-period panel; the individual-level robustness tables in Appendix A.4 (e.g., Table A.17, Table A.8) use a linear exposure-by-time trend estimated on the individual-level microdata, which for young-worker wages gives p=0.201 under the same specification. Table 1 reports both versions side by side for the main outcomes.

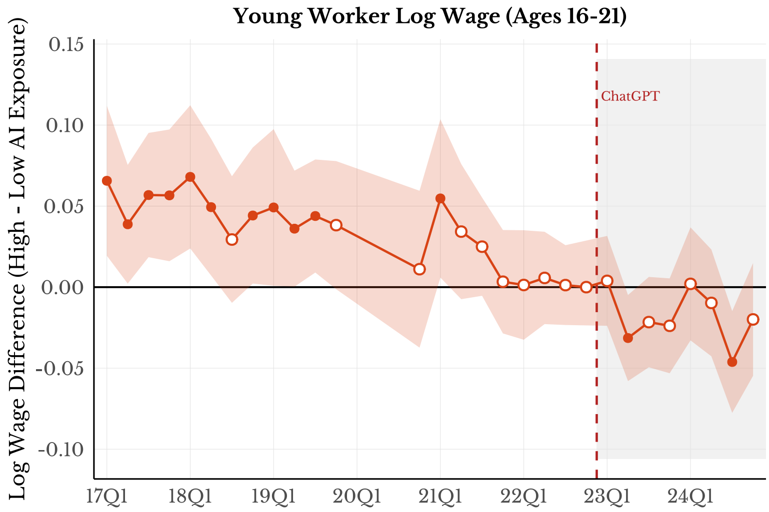

Figure 3: Young Worker Wage Penalty: Event Study

Notes: Event study coefficients for young workers (ages 16–21) log wage from PNAD individual-level estimation, spanning 2017Q1–2024Q4 (with gap during 2020Q1–Q3). Coefficients show the difference between high- and low-exposure occupations relative to 2022Q4 (the last pre-treatment quarter and reference period). Pre-trends test passes (p=0.116). Solid circles: significant at 5%; hollow circles: not significant. Shaded region marks post-ChatGPT period.

than AI exposure per se, determine which workers bear adjustment costs. Private formal workers show a wage decline of 0.9% (p=0.073, pre-trends p=0.887), significant at the 10% level. Self-employed workers and private informal workers show no significant effects, with most passing pre-trends.

Table 2: Heterogeneity by Employment Type

	Private Formal (1)	Private Informal (excl. Self-Employed) (2)	Public Sector (3)	Self-Employed (4)
Panel A: Log Wage				
High AI × Post	−0.0094*	−0.0026	−0.0024	−0.0008
	(0.0052)	(0.0119)	(0.0189)	(0.0178)
Pre-trends test	[0.887]	[0.101]	[0.427]	[0.036]
N	1,881,142	750,814	801,933	1,618,048
Panel B: Weekly Hours				
High AI × Post	0.0463	0.0114	0.1471	0.3342
	(0.0680)	(0.1345)	(0.2310)	(0.3349)
Pre-trends test	[0.133]	[0.918]	[0.130]	[0.642]
N	1,898,598	799,560	809,930	1,711,759
Sample share	33%	13%	14%	28%

Notes: DiD by employment type with occupation, period, and state fixed effects. Sample: PNAD Contínua 2017Q1–2024Q4. High AI exposure = above median. Private Formal = formal private sector employees with signed work card (carteira assinada). Private Informal (excl. Self-Employed) = informal private sector employees without work card, excluding own-account workers. Public Sector = public employees. Self-Employed = own-account workers (conta própria). Shares sum to approximately 88%; the remaining ∼12% are employers and unpaid family workers, omitted from this breakdown. The aggregate coefficient in Table 1 includes these groups and differs from the share-weighted average of this table's subgroup coefficients because the fixed-effects specification induces reweighting across cells. Pre-trends p-values in brackets; values ≥0.05 support parallel trends. Standard errors clustered by occupation. *** p<0.01, ** p<0.05, * p<0.10.

4.2 Administrative Hiring Records

PNAD-C captures outcomes for incumbent workers. To understand firm-side adjustment, we turn to CAGED, Brazil's administrative registry of formal sector hiring and separations (approximately 13 million events from October 2020 through December 2024, excluding April–September 2021).

Employment Flows. Table 3 Panel A shows log quits 1.8% lower in high-exposure than low-exposure occupations after ChatGPT's release (p=0.549); the coefficient also fails its pre-trends test (p=0.001), so we do not read it as identifying. Log hires are not statistically distinguishable from zero in the main specification (0.0150, p=0.633), and pre-trends fail (p=0.001); the state-disaggregated robustness panel (Table A.6) yields a coefficient of the same sign and likewise insignificant, so neither aggregation supports a quantitative claim about hiring volume. The age-specific rows in Panel B are not strictly comparable to the Panel A aggregate: because CAGED code 40 (worker-initiated separations) is too sparse to support an occupation × age × month panel, the by-age column labeled "Log Quits" uses total separations (all CAGED separation

Table 3: Effects of AI Exposure on Formal Sector Employment (CAGED)

	Log Hires (1)	Separations† (2)	Edu. (Hire) (3)	Edu. (Sep.) (4)	Log Wage (Hire) (5)
Panel A. Aggregate Effects (col. 2: log quits, code 40 only)					
High Exposure × Post	−0.0328	−0.0607*	−0.0820	−0.2264***	−0.0181**
	(0.0322)	(0.0310)	(0.0534)	(0.0475)	(0.0088)
Pre-trends p-value	[0.056]	[0.028]	[0.431]	[0.343]	[<0.001]
Panel B. Age Heterogeneity (col. 2: log total separations, all codes)					
Young (16–21)	−0.0575**	−0.0471	−0.0772	−0.1273**	−0.0337**
	(0.0281)	(0.0328)	(0.0590)	(0.0627)	(0.0166)
Pre-trends p-value	[<0.001]	[0.192]	[0.845]	[0.784]	[0.430]
Early Career (22–29)	−0.0623**	0.0028	−0.0842	−0.1988***	−0.0141
	(0.0307)	(0.0296)	(0.0525)	(0.0656)	(0.0253)
Pre-trends p-value	[0.066]	[0.031]	[0.412]	[0.967]	[0.394]
Prime (30–49)	−0.0304	0.0355	−0.0882	−0.3208***	0.0067
	(0.0317)	(0.0277)	(0.0647)	(0.0594)	(0.0344)
Pre-trends p-value	[0.142]	[0.319]	[0.409]	[0.472]	[0.791]
Older (50+)	−0.0290	−0.0672**	−0.0187	−0.1042	0.0424
	(0.0298)	(0.0304)	(0.0994)	(0.0826)	(0.0430)
Pre-trends p-value	[0.327]	[0.314]	[0.403]	[0.795]	[0.171]

Notes: Binary difference-in-differences comparing high- versus low-exposure occupations (median split). Regression run at occupation-month level (aggregating across states), unweighted, with occupation and year-month fixed effects. Panel A reports aggregate effects; Panel B reports age-specific subsamples. †Column 2 uses *different* dependent variables across the two panels: Panel A regresses on log quits restricted to CAGED code 40 (worker-initiated separations), as stated in the panel heading; Panel B regresses on log total separations (all CAGED separation codes) because code 40 is too sparse to support an occupation × age × month panel. The Panel B column-2 estimate is therefore not a decomposition of the Panel A quit coefficient and the two should not be compared row-wise. Edu = mean years of schooling. Log Wage (Hire) = log mean hiring wage, estimated on the hiring-wage panel (occupations with at least one hire in the month). Standard errors clustered at occupation level in parentheses. Pre-trends p-values test differential pre-treatment trends; p≥0.05 supports parallel trends. Sample: CAGED Oct 2020–Dec 2024 (extended sample). The Prime and Older bins differ by one year between PNAD (Prime 30–54, Older 55+; see Table 1) and CAGED (Prime 30–49, Older 50+; this table) for sample-size reasons in the CAGED state-by-month panel; Table A.11 re-estimates PNAD under the CAGED bins and confirms the prime/older nulls are unchanged. The Young (16–21) and Early Career (22–29) bins are identical across data sources and carry the headline story. ***p<0.01, **p<0.05, *p<0.10.

codes), whereas the aggregate row uses code 40 only. We therefore read Panel B as describing the composition of exits across ages rather than as an age-level analogue to the aggregate quits estimate.

Hiring Composition. The aggregate hiring-wage coefficient in Table 3 fails its pre-trends test (p<0.001), so we do not read it as identifying and discuss hiring wages only in the age-stratified subgroups, where pre-trends pass. Hiring wages for young workers (16–21) in high-exposure occupations are 1.3% lower after ChatGPT's release (p=0.086, pre-trends p=0.256); the coefficient is marginal on conventional thresholds but its sign matches the PNAD young-worker wage finding, which addresses the concern that the household-survey result reflects compositional shifts rather than wage changes for similar workers. The early-career bin (22–29) shows a sharper decline of 2.6% (p=0.002, pre-trends p=0.421), suggesting that the wage adjustment on the firm side extends past the youngest workers; prime-age (30–49) and older (50+) hires show smaller, insignificant coefficients. New hires in high-exposure occupations also come in with less schooling after ChatGPT's release, with coefficients pointing in the same direction across age groups (22–29: -0.1520 years, p=0.001, pre-trends p=0.718; 16–21: -0.079 years, p=0.131, pre-trends p=0.795); the magnitudes are modest in absolute terms (roughly one month of schooling each). Neither age-specific schooling coefficient reaches conventional significance in the main occupation-month specification, so we read the schooling pattern as suggestive rather than as a headline finding (Section 4.3).[4] The young-worker hiring-wage finding is consistent with the task-unbundling channel in Autor and Thompson (2025): if AI substitutes for tasks previously performed by junior workers, firms can lower entry wages while incumbents remain insulated.

4.3 Robustness and Interpretation

We close by separating the findings that survive our reading protocol (Section 3.4) from those that do not. Figure 4 synthesizes the underlying point estimates across both data sources.

Main specifications. Before listing what we read as robust, we fix the specification used as the reference point. PNAD outcomes: the individual-level DiD on the 2017Q1–2024Q4 panel with occupation, state, and period fixed effects (Table 1). CAGED outcomes: the occupation-month aggregation (Table 3). The state-disaggregated CAGED panel (state $\times$ occupation $\times$ month) that appears in the alternative-measures, BH, and sensitivity tables (Tables A.19, A.5, and elsewhere in Appendix A.4) is a robustness specification, not the headline; it uses a different aggregation level and therefore yields different coefficients and p-values, which we call out in the notes to each appendix table rather than letting the reader infer by subtraction.

Patterns we treat as robust. Two findings appear on the PNAD side. Young workers (16–21) in high-exposure occupations see wages 4.7% lower than young workers in low-exposure occupations after ChatGPT's release; the coefficient passes the linear pre-trends test in the main specification, strengthens rather than weakens under

[4] Alternative AI exposure measures from Anthropic and Microsoft yield qualitatively similar signs for aggregate education of hires (Appendix Table A.19), though age-specific breakdowns are only available for the primary O*NET measure.

the Q4-vs-Q3 adjacent-quartile contrast (from -0.048 to -0.061, Table A.17), passes all five date placebos (Table A.13 Panel D, p-values between 0.136 and 0.446), and yields a randomization-inference p-value of 0.010 on the exposure vector from 200 permutations (Table A.18). Young workers in high-exposure occupations are also 1.5 percentage points less likely to be formally employed, with passing pre-trends and survival of the Benjamini–Hochberg FDR correction (Table A.5). One cross-source corroboration appears on the CAGED side: hiring wages for new hires in high-exposure occupations are lower after ChatGPT's release in both age bins where pre-trends pass, 2.6% for the early-career group 22–29 (p=0.002, pre-trends p=0.421) and 1.3% for young workers 16–21 (p=0.086, pre-trends p=0.256, marginal on conventional thresholds). The early-career coefficient is the cleaner cross-source corroboration of the PNAD young-worker wage result, since the youngest CAGED bin sits at the boundary of conventional significance; we read the two together as evidence that the household-survey pattern is not driven by compositional change. For the two PNAD outcomes, criterion (ii) of the reading protocol is not directly testable because age-specific breakdowns are available only for the Eloundou exposure measure; we therefore read cross-measure replication as applying at the aggregate-pattern level rather than at the age-stratified level. Rambachan–Roth sensitivity bounds (Appendix A.3.2) widen these intervals quickly: no coefficient survives $\bar{M} \geq 1$. We therefore read these patterns as robust at the pattern level, not at the level of any single coefficient.

Patterns we treat as suggestive but not robust. Five findings sit below the bar. First, *aggregate formality*: the 0.9 percentage point decline passes pre-trends (p=0.298) but sign-flips to $+0.005$ under the Q4-vs-Q3 adjacent-quartile test (Table A.17), disqualifying it under criterion (iii) of the reading protocol. The young-worker formality result, by contrast, does not sit on the aggregate coefficient and retains its sign in the age-stratified panel. Second, *the aggregate wage null*: date placebos fire at 2022Q1 and 2022Q2 (p=0.007, p=0.009; Appendix A.4.2), so we cannot credibly report a null on an outcome where our own placebos say we lack identifying variation. Third, *log quits* in the main CAGED specification (Table 3): the point estimate is small and not significant (-0.0180, p=0.549) and the pre-trends test fails (p=0.001), so we report it for completeness but do not lean on it. The Eloundou-measure quits coefficient in the alternative-measures table is somewhat larger (-0.037, p=0.027, pre-trends pass), but Microsoft and Anthropic produce smaller and insignificant coefficients with failing pre-trends, so we read cross-measure agreement on quits as weak rather than supporting. Fourth, *schooling of 22–29 hires*: the coefficient in the main specification is -0.1520 years with p=0.001, directionally consistent with the hiring-wage and household-survey findings but not reaching conventional significance. The 16–21 schooling coefficient points in the same direction and passes pre-trends but is likewise not significant in the main specification. Fifth, *young-worker log hires*: the coefficient in Table 3 Panel B is -0.0575^{**} but its pre-trends test fails ($p<0.001$), so Figure 4 renders it as a hollow marker and we read it as directional only. Replication across alternative exposure measures (Eloundou et al., Microsoft, Anthropic; Appendix A.4.7) is uneven: education of new hires agrees in sign across all three measures, whereas CAGED flow outcomes show consistent signs but fail pre-trends under every measure. We describe these patterns where relevant but do not lean on them as headline evidence. The Benjamini–Hochberg correction in Table A.5 is reported across two column-groups: the main-specification occupation-month p-values (primary) and the state-disaggregated p-values (robustness). Only the PNAD young-worker wage coefficient survives BH under the main specification, and it is the single pattern we treat as robust at the individual-coefficient level; the six outcomes that survive BH on the state-disaggregated panel are more precisely estimated but rely on variation absent from the main specification.

The noisy event-study profiles and the gap between sign and pre-trend consistency in the alternative-measures table are also consistent with slower diffusion of generative AI in Brazil than in the developed economies where it was first studied. Adoption surveys place Brazilian firms behind OECD benchmarks on generative AI use in the 2023–2024 window (Bakker et al., 2024). A slower, more heterogeneous diffusion process would smear the treatment timing, weaken the average treatment effect relative to a cleaner shock, and make occupation-level exposure proxies a noisier stand-in for actual adoption, all of which are visible in our event studies.

What we do not claim. We do not claim to have identified clean causal effects of generative AI on the Brazilian labor market. We do not report absolute magnitudes; every estimate is a relative difference between high- and low-exposure occupations. The patterns above are consistent with two broader readings that we view as the paper's substantive contribution: (i) conditions for labor-market entrants, especially recent high-school graduates, deteriorate in AI-exposed occupations after ChatGPT's release; and (ii) adjustment in Brazil runs through informality and entry-wage margins rather than through aggregate employment, a distinct channel from what US and UK evidence has documented. Disentangling the mechanism beyond this requires data we do not have: matched employer–employee records (RAIS), direct firm-level measures of AI adoption (most plausibly constructed from job-posting text), and longer post-treatment horizons. We leave those to follow-up work.

Figure 5 plots the aggregate event study. The post-treatment coefficients do not show a sharp break at 2022Q4, which is what we would expect from gradual

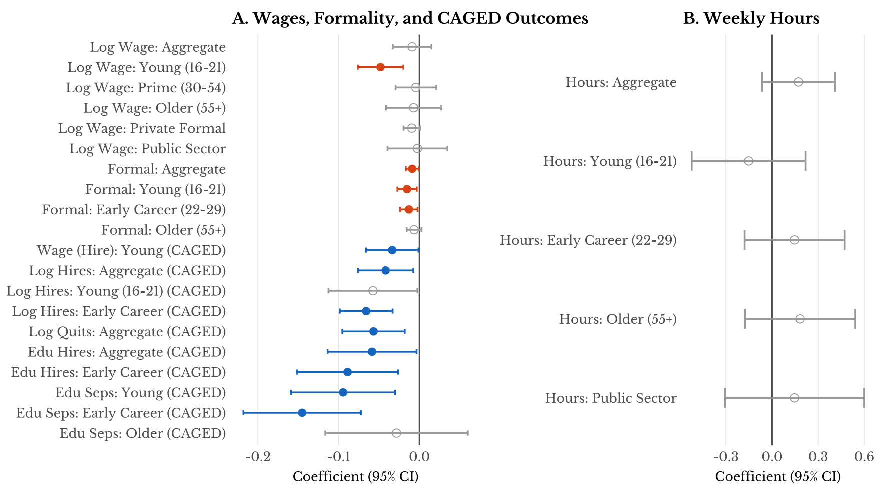

Figure 4: Estimated Effects Across Outcomes and Age Groups

Notes: Point estimates and 95% confidence intervals for the main DiD outcomes; see Tables 1 and 3 for individual p-values and pre-trends tests. The three patterns we treat as robust in Section 4.3 are the young-worker (16–21) log-wage and formality bars and the CAGED young hiring-wage bar. Aggregate log wages, aggregate formality, log quits, and the education-of-hires coefficients appear for completeness but do not clear the robustness protocol: aggregate formality fails the adjacent-quartile test, aggregate wages fail date placebos at 2022Q1–Q2, log quits fails its pre-trends test (p=0.001), and the 22–29 schooling coefficient is p=0.001 in the main specification. Solid points: statistically significant ($p < 0.05$) with pre-trends passing; hollow points: not significant, or pre-trends failing (e.g., young-worker log hires). Solid PNAD estimates in orange; solid CAGED estimates in blue; hollow points in gray. Panel A: log wages, formality, log quits, and education of hires. Panel B: weekly hours (different scale). Age groups: young (16–21), early career (22–29), prime (30–54), older (55+).

technology diffusion and heterogeneous firm adoption timing: the pattern builds over the post-treatment window rather than appearing as a discrete jump. Coding the full quarter containing the November 30 release as post-treatment further attenuates the first-period coefficient; a specification beginning post-treatment in 2023Q1 is a natural variant.

Where US and UK evidence concentrates on employment quantity declines among 22–25 year-old workers (Brynjolfsson et al., 2025; Klein Teeselink, 2025), the Brazilian pattern concentrates among 16–21 year-old high-school graduates and appears on wage and contract-quality margins, with cross-source corroboration on young hiring wages in administrative records, rather than on aggregate employment. That this margin-shift appears in the Brazilian data, alongside the documented informality channel in Latin American robot-automation studies (César et al., 2023), is one of the novel patterns the paper reports.

Appendix A.4 reports the full battery of robustness exercises behind the reading protocol in Section 3.4: alternative AI-exposure measures (Anthropic,

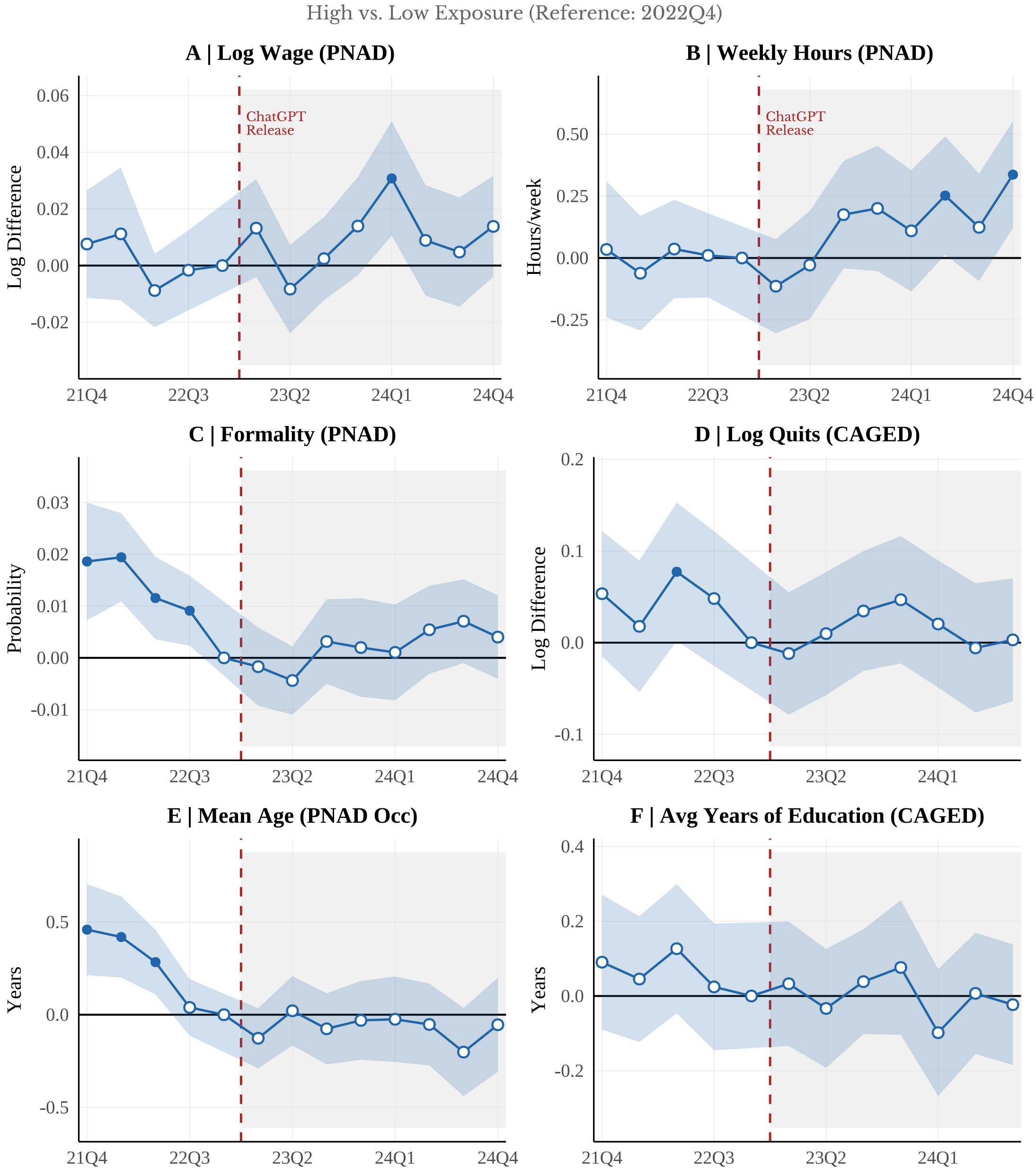

Figure 5: Event Study: Dynamic Treatment Effects by Outcome

Notes: Event study coefficients showing the interaction between AI exposure and period indicators, with 2022Q4 (the last pre-treatment quarter) as the reference period. Panels A–C, E: PNAD outcomes spanning 2017Q1–2024Q4 with gap during 2020Q1–Q3. Panels D, F: CAGED outcomes spanning October 2020–December 2024 (excluding April–September 2021). Shaded region indicates post-ChatGPT periods (2023Q1 onward for PNAD, November 2022 onward for CAGED). Solid circles: statistically significant ($p<0.05$); hollow circles: not significant. Pre-trends tests: log wage ($p=0.103$), weekly hours ($p=0.268$), formality ($p=0.298$), and education of young hires ($p=0.795$) pass; log quits ($p=0.001$) fails. Panel E (mean age of workers in an occupation, PNAD occupation-level panel) is shown as a compositional diagnostic only; its pre-trends test fails and no aggregate DiD coefficient for this outcome is reported in the main or appendix tables, so the panel should be read as descriptive rather than as a treatment estimate. Standard errors clustered at occupation level.

Microsoft), alternative cutoffs and adjacent-quartile comparisons, demographic controls under a bad-control-safe triage, 2-digit clustering, five date placebos, exposure randomization, compositional placebos, shorter pre-period and extended CAGED sample, and Rambachan–Roth sensitivity bounds. The young-worker wage, young-worker formality, and CAGED young hiring-wage patterns clear the protocol across these exercises. The aggregate formality coefficient is stable across demographic control sets (it drifts from -0.009 to -0.007 and retains significance at the 10% level) but sign-flips under the Q4-vs-Q3 adjacent-quartile contrast, which is why Section 4.3 flags it as suggestive rather than robust.

5 Conclusion

ChatGPT's November 2022 release created a quasi-experimental opportunity to document early labor-market responses to generative AI in a developing economy. Combining 5.7 million worker-quarter observations from PNAD-C with approximately 13 million formal hiring and separation records from CAGED, our central difference-in-differences estimate is that young workers (16–21) in high-exposure occupations earn 4.7% less than young workers in low-exposure occupations after ChatGPT's release. Two supporting estimates point the same direction but do not clear every robustness check: young workers in high-exposure occupations are 1.5 percentage points less likely to be formally employed, and administrative hiring wages for young workers in high-exposure occupations fall 1.3% more than in low-exposure occupations. Prime-age and older workers show no differential movement on either wages or formality. Taken together, these patterns are consistent with a barrier-to-entry dynamic in which the terms of entry for 16–21 year-old high-school graduates adjust rather than the level of employment. The task-unbundling framework of Autor and Thompson (2025) is a candidate mechanism consistent with this pattern, but we do not document deskilling directly: the age-specific schooling-of-hires coefficients point in the same direction but do not clear conventional significance in the main CAGED specification, and we flag them as suggestive rather than robust. The pattern we do document diverges from the employment-quantity effects in US and UK studies (Brynjolfsson et al., 2025; Klein Teeselink, 2025) and is consistent with the informality-absorption channel documented for industrial robots in Latin America (César et al., 2023).

We do not claim to have identified clean causal effects. Three limitations are load-bearing. First, the October 2022 presidential election coincides with ChatGPT's release and produces a timing confound we cannot fully resolve. Second, we do not observe AI adoption at the firm or worker level, so exposure-based variation is a proxy for treatment rather than a measurement of it. Third, spillovers between high- and low-exposure occupations are likely, so every coefficient we report is a relative difference between groups rather than an absolute effect on either (Huber and Steinmayr, 2021). Parallel-trends pre-tests have limited statistical power against the violations that would most concern us (Roth, 2022), and Rambachan–Roth sensitivity bounds (Appendix A.3.2) show that the individual coefficients widen rapidly once parallel trends is relaxed even modestly. We see this paper as mostly documenting patterns, disciplined through difference-in-differences rather than well-defined causal estimands. This is a deliberate choice given the urgency of the topic: generative AI is diffusing faster than the data infrastructure needed to identify its effects cleanly, and we think it is useful to put the first-order patterns on the table now, with their caveats visible, rather than wait. More work will naturally be needed to pin down well-identified estimands. The patterns here are assessed against pre-trends tests, date and exposure placebos, Rambachan–Roth sensitivity bounds, adjacent-quartile comparisons, and three independent exposure measures. Section 4.3 separates the coefficients that clear these checks from those that do not.

Two potential extensions would sharpen the interpretation. Linking matched employer–employee records to direct measures of firm-level AI adoption would move beyond exposure-based variation and help separate the treatment from the proxy. Longer post-treatment horizons and replications in other developing economies with different levels of informality and educational attainment would test whether the barrier-to-entry pattern we document here is general. Until then, we regard the findings in this paper as suggestive of two things we think the data measure credibly: that early generative AI diffusion is associated with deteriorating conditions for labor-market entrants, especially recent high-school graduates; and that adjustment in Brazil runs through informality and entry-wage margins rather than aggregate employment. For readers interested in how developing-economy labor markets absorb AI shocks, these patterns point to labor-market entrants as the group whose outcomes are most worth tracking, a population that in Brazil is predominantly young, female, and working in routine cognitive positions without access to unemployment insurance or severance protections.

References

Daron Acemoglu and Pascual Restrepo. Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30, 2019. doi: 10.1257/jep.33.2.3.

Palakorn Achananuparp, Ee-Peng Lim, and Yao Lu. A multi-stage framework with taxonomy-guided reasoning for occupation classification using large language models. *arXiv preprint arXiv:2503.12989*, 2025. doi: 10.48550/arXiv.2503.12989.

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2008.

Ruth Appel, Peter McCrory, Alex Tamkin, Michael Stern, Miles McCain, and Tyler Neylon. Anthropic economic index report: Uneven geographic and enterprise ai adoption, September 2025. URL https://www.anthropic.com/research/anthropic-economic-index-september-2025-report.

Claudia Assis and Martha Campos. Brazil president lula hikes minimum wage in nod to social agenda. Bloomberg, April 2023. URL https://www.bloomberg.com/news/articles/2023-04-30/brazil-s-lula-hikes-minimum-wage-in-nod-to-social-agenda.

David Autor and Neil Thompson. Expertise. NBER Working Paper 33941, National Bureau of Economic Research, June 2025. URL https://www.nber.org/papers/w33941.

Bas Bakker, Sophia Chen, Dmitry Vasilyev, Olga Bespalova, Moya Chin, Daria Kolpakova, Archit Singhal, and Yuanchen Yang. What Can Artificial Intelligence Do for Stagnant Productivity in Latin America and the Caribbean? IMF Working Paper 2024/219, International Monetary Fund, October 2024. URL https://www.imf.org/en/Publications/WP/Issues/2024/10/11/What-Can-Artificial-Intelligence-Do-for-Stagnant-Productivity-in-Latin-America-and-the-556243.

Fernando de Holanda Barbosa Filho, Fernando Veloso, and Paulo Henrique Peruchetti. Trabalho remoto no brasil [remote work in brazil]. *Revista Brasileira de Economia*, 76(4):349–378, 2022. URL https://periodicos.fgv.br/rbe/article/view/85168.

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300, 1995.

Alexander Bick, Adam Blandin, and David J. Deming. The rapid adoption of generative AI. Working Paper 32966, National Bureau of Economic Research, September 2024. URL http://www.nber.org/papers/w32966.

Erik Brynjolfsson, Bharat Chandar, and Ruyu Chen. Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence. Working paper, Stanford Digital Economy Lab, August 2025. URL https://digitaleconomy.stanford.edu/wp-content/uploads/2025/08/Canaries_BrynjolfssonChandarChen.pdf.

Erik Brynjolfsson, Bharat Chandar, and Ruyu Chen. Canaries, interest rates, and timing: More on recent drivers of employment changes for young workers. Technical note, Stanford Digital Economy Lab, February 2026. URL https://digitaleconomy.stanford.edu/news/canaries-interest-rates-and-timinga-more-on-recent-drivers-of-employment-changes-for-young-workers/.

BUREAU OF LABOR STATISTICS. Occupational employment and wage statistics. U.S. Bureau of Labor Statistics, 2024. URL https://www.bls.gov/oes/.

Carolina Caetano and Brantly Callaway. Difference-in-differences when parallel trends holds conditional on covariates. *arXiv preprint arXiv:2406.15288*, 2024. doi: 10.48550/arXiv.2406.15288.

Brantly Callaway and Pedro H. C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021. doi: 10.1016/j.jeconom.2020.12.001.

Andrés César, Guillermo Falcone, Leonardo Gasparini, and Irene Brambilla. The impact of robots in latin america: Evidence from local labor markets. *World Development*, 170:106271, 2023. doi: 10.1016/j.worlddev.2023.106271.

Gustavo de Souza. Artificial intelligence in the office and the factory: Evidence from administrative software registry data. FRB of Chicago Working Paper 2025-11, Federal Reserve Bank of Chicago, July 2025. URL https://ssrn.com/abstract=5375463.

Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. Gpts are gpts: Labor market impact potential of llms. *Science*, 384(6702):1306–1308, 2024. doi: 10.1126/science.adj0998.

Niklas Engbom and Christian Moser. Earnings inequality and the minimum wage: Evidence from brazil. *American Economic Review*, 112(12):3803–3847, 2022. doi: 10.1257/aer.20181506.

Sergio Firpo and Alysson Portella. Informal and small: How labor market institutions affect inequality, income volatility and labor productivity in brazil. UNDP LAC Working Paper 22, UNDP, March 2021. URL https://files.acquia.undp.org/public/migration/latinamerica/undp-rblac-PNUD_bckPapers22-OK.pdf.

Sergio Pinheiro Firpo and Alysson Lorenzon Portella. The labor market in brazil, 2001–2022. *IZA World of Labor*, (441), April 2024. doi: 10.15185/izawol.441.v2. URL https://wol.iza.org/articles/the-labor-market-in-brazil/long.

Luis Garicano and Luis Rayo. Training in the age of AI: A theory of apprenticeship viability. CEPR Discussion Paper 20634, Centre for Economic Policy Research, 2025. URL https://cepr.org/publications/dp20634.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021. doi: 10.1016/j.jeconom.2021.03.014. URL https://www.sciencedirect.com/science/article/pii/S0304407621001445.

Kunal Handa, Alex Tamkin, Miles McCain, Saffron Huang, Esin Durmus, Sarah Heck, Jared Mueller, Jerry Hong, Stuart Ritchie, Tim Belonax, Kevin K. Troy, Dario Amodei, Jared Kaplan, Jack Clark, and Deep Ganguli. Which economic tasks are performed with ai? evidence from millions of claude conversations. arXiv preprint arXiv:2503.04761, 2025. URL https://arxiv.org/abs/2503.04761.

Seyed Mahdi Hosseini Maasoum and Guy Lichtinger. Generative ai as seniority-biased technological change: Evidence from u.s. résumé and job posting data. Working paper, Harvard University Department of Economics, August 2025. URL https://ssrn.com/abstract=5425555.

Krystal Hu. Chatgpt sets record for fastest-growing user base – analyst note. Reuters, February 2023. URL https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

Martin Huber and Andreas Steinmayr. A framework for separating individual-level treatment effects from spillover effects. *Journal of Business & Economic Statistics*, 39 (2):422–436, 2021. doi: 10.1080/07350015.2019.1668795.

Xiang Hui, Oren Reshef, and Luofeng Zhou. The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market. *Organization Science*, 35(6):1977–1989, 2024. doi: 10.1287/orsc.2023.18441.

Anders Humlum and Emilie Vestergaard. Large language models, small labor market effects. NBER Working Paper 33777, National Bureau of Economic Research, May 2025. URL https://www.nber.org/papers/w33777.

Nick Huntington-Klein. *The Effect: An Introduction to Research Design and Causality*. Chapman and Hall/CRC, Boca Raton, FL, 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Classificação de ocupações para pesquisas domiciliares - cod. IBGE, 2010. URL https://www.ibge.gov.br/arquivo/projetos/sipd/oitavo_forum/COD.pdf.

INTERNATIONAL LABOUR ORGANIZATION. *International Standard Classification of Occupations 2008 (ISCO-08): Structure, Group Definitions and Correspondence Tables*, volume 1. International Labour Office, Geneva, 2012. URL https://www.ilo.org/publications/international-standard-classification-occupations-2008-isco-08-structure.

Jeremy Kahn. The inside story of ChatGPT: How OpenAI founder Sam Altman built the world's hottest technology with billions from Microsoft. Fortune, January 2023. URL https://fortune.com/longform/chatgpt-openai-sam-altman-microsoft/.

Bouke Klein Teeselink. Generative ai and labor market outcomes: Evidence from the united kingdom. Working paper, King's College London, September 2025. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5516798.

Nan Li, Bo Kang, and Tijl De Bie. LLM4Jobs: Unsupervised occupation extraction and standardization leveraging large language models. *arXiv preprint arXiv:2309.09708*, 2023. doi: 10.48550/arXiv.2309.09708.

Yan Liu and He Wang. Who on Earth Is Using Generative AI? Policy Research Working Paper 10870, World Bank, Washington, DC, 2024. URL https://openknowledge.worldbank.org/entities/publication/5a876bd0-f85a-479b-ae32-cf0b7f33792f.

Maxim Massenkoff and Peter McCrory. Labor market impacts of AI: A new measure and early evidence. Anthropic Research Note, March 2026. URL https://www.anthropic.com/research/labor-market-impacts.

Costas Meghir, Renata Narita, and Jean-Marc Robin. Wages and informality in developing countries. *American Economic Review*, 105(4):1509–1546, 2015. doi: 10.1257/aer.20121110.

Naercio Menezes-Filho and Renata Narita. Labor market turnover and inequality in Latin America. *Oxford Open Economics*, 4(Supplement_1):i349–i375, 2025. doi: 10.1093/ooec/odae027.

MINISTÉRIO DO TRABALHO E EMPREGO. Classificação brasileira de ocupações - cbo, 2024. URL https://www.gov.br/trabalho-e-emprego/pt-br/assuntos/cbo. Originally established by Portaria n.º 397, October 9, 2002; continuously updated.

Rafael Machado Parente. Minimum wages, inequality, and the informal sector. *IMF Working Paper WP/24/159*, 2024. doi: 10.5089/9798400280672.001.

Carlo Pizzinelli, Augustus Panton, Marina M. Tavares, Mauro Cazzaniga, and Longji Li. Labor market exposure to AI: Cross-country differences and distributional implications. IMF Working Paper WP/23/216, International Monetary Fund, October 2023. URL https://www.imf.org/en/Publications/WP/Issues/2023/10/04/Labor-Market-Exposure-to-AI-Cross-country-Differences-and-Distributional-Implications-539656.

Ashesh Rambachan and Jonathan Roth. A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591, 2023. doi: 10.1093/restud/rdad018.

Jonathan Roth. Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–322, 2022. doi: 10.1257/aeri.20210236.

Jonathan Roth, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244, 2023. doi: 10.1016/j.jeconom.2023.03.008. URL https://www.sciencedirect.com/science/article/pii/S0304407623001318.

Henry Stemmler. Automated deindustrialization: How global robotization affects emerging economies—evidence from brazil. *World Development*, 171:106349, 2023. doi: 10.1016/j.worlddev.2023.106349.

Kiran Tomlinson, Sonia Jaffe, Will Wang, Scott Counts, and Siddharth Suri. Working with ai: Measuring the applicability of generative ai to occupations. arXiv preprint arXiv:2507.07935, 2025. URL https://arxiv.org/abs/2507.07935.

Gabriel Ulyssea. Firms, informality, and development: Theory and evidence from brazil. *American Economic Review*, 108(8):2015–47, August 2018. doi: 10.1257/aer.20141745.

Kyle Wiggers. One year later, chatgpt is still alive and kicking. TechCrunch, November 2023. URL https://techcrunch.com/2023/11/30/one-year-later-chatgpt-is-still-alive-and-kicking/.

A Appendix

A.1 Additional Figures

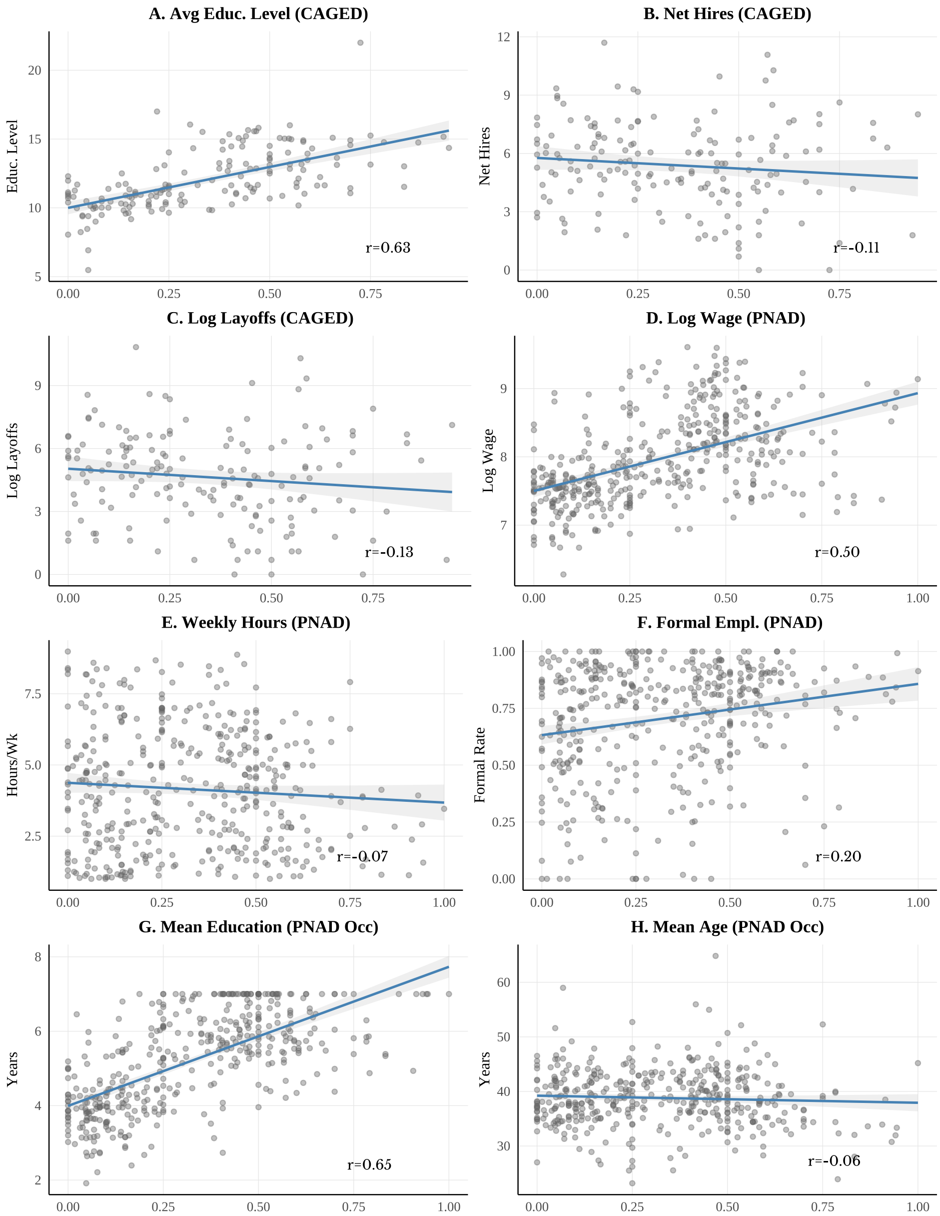

Figure A.1: Baseline Outcomes vs AI Exposure

Notes: Occupation-level baseline relationships (2022Q2–Q3, pre-ChatGPT). Each point represents one 4-digit occupation; fitted lines show OLS with 95% confidence intervals. Pearson correlations (r) displayed. Data: PNAD Contínua (wages, hours, formality) and CAGED (hires, education).

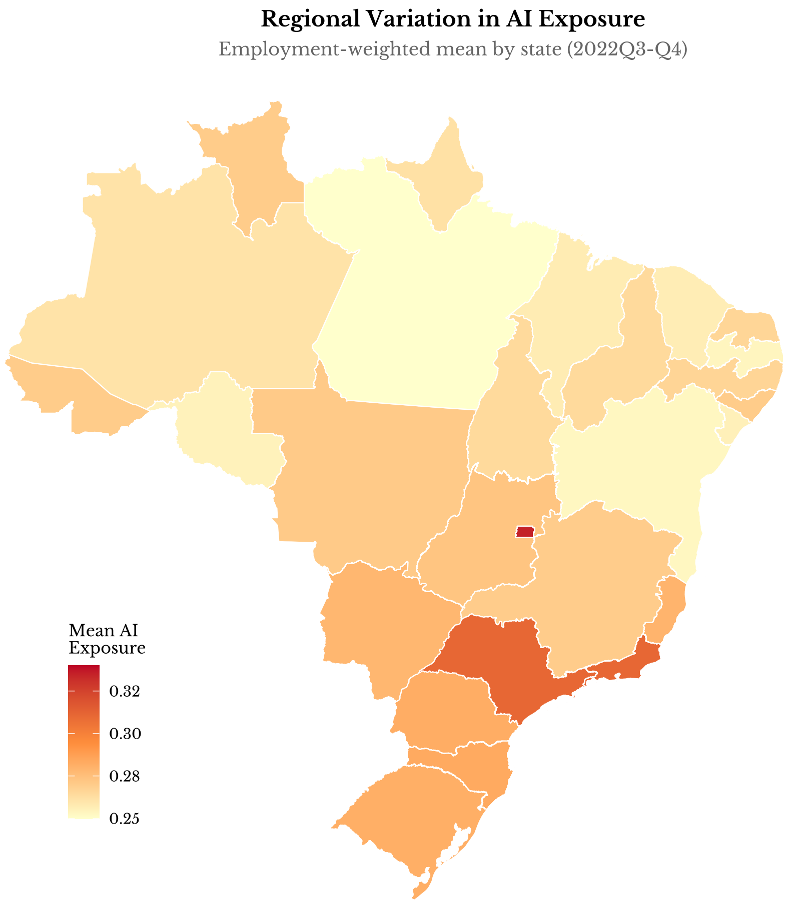

Figure A.2: Regional Variation in AI Exposure

Notes: Mean AI exposure by state, employment-weighted. South and Southeast regions show higher exposure, reflecting concentration of professional occupations. Data: PNAD Contínua 2022Q3–Q4 (last two pre-treatment quarters).

A.2 Data and Measurement

A.2.1 Occupation Classification Systems

Brazil's labor market data uses two related occupation classification systems:

- **CBO (Classificação Brasileira de Ocupações):** Administrative classification used in CAGED and RAIS MINISTÉRIO DO TRABALHO E EMPREGO (2024). CBO-94 (1994–2002) used hyphenated codes; CBO-2002 (2003–present) uses six-digit codes at full detail (e.g., 412105, 211515), which we aggregate to four-digit "families" (e.g., 4121, 2115) for analysis. This aggregation facilitates comparison with PNAD's COD 2010 classification.
- **COD 2010:** Household survey classification used in PNAD-C, based on ISCO-08 INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (2010); INTERNATIONAL LABOUR ORGANIZATION (2012). Maintains compatibility with ISCO-08 at the 2-digit level and covers 434 distinct 4-digit occupation codes in the estimation sample (412 of which receive a positive AI exposure score).

This paper uses COD 2010 codes from PNAD-C and aggregates CAGED's CBO-2002 codes to 4-digit families for compatibility. For PNAD analysis, the SOC-to-COD mapping bypasses CBO by using LLM-based analysis of occupation descriptions. A parallel SOC-to-CBO mapping was constructed using the same LLM-based procedure for CAGED data; Table A.1 shows both COD and CBO codes for selected occupations.

Example Occupation Crosswalk Table A.1 presents selected examples from the occupation crosswalk linking U.S. O*NET codes (SOC) to Brazilian classifications: CBO (Classificação Brasileira de Ocupações, used in CAGED administrative data) and COD (Classificação de Ocupações para Pesquisas Domiciliares, used in PNAD-C household surveys). High-exposure occupations include IT professionals (programmers, software developers), white-collar professionals (journalists, accountants, lawyers), and clerical workers (data entry). Low-exposure occupations concentrate in manual trades (construction, cleaning) and personal services (waiters, delivery workers).

LLM-Based Mapping Procedure. The crosswalk was constructed using Claude Sonnet 4 to propose initial matches between Brazilian occupation codes and U.S. SOC codes based on occupation titles, task descriptions, and industry context from O*NET and IBGE documentation, inspired by LLM-based occupation

Table A.1: Example Occupation Classification Crosswalk

SOC	O*NET Title	CBO	COD	Brazilian Title	Exp.
High Exposure (>0.50)					
15-1251	Computer Programmers	2514	2124	Programadores	0.94
15-1252	Software Developers	2512	2123	Desenvolvedores de software	0.87
43-9021	Data Entry Keyers	4132	4121	Digitadores	0.83
27-3023	Reporters and Journalists	2642	2611	Jornalistas	0.65
27-3041	Editors	2643	2614	Tradutores e intérpretes	0.64
13-2011	Accountants and Auditors	2411	2522	Contadores	0.55
Medium Exposure (0.20–0.50)					
23-1011	Lawyers	2611	2410	Advogados	0.43
25-2031	Secondary School Teachers	2330	2321	Professores ensino médio	0.33
25-2021	Elementary School Teachers	2341	2312	Professores ensino fund.	0.30
Low Exposure (<0.20)					
35-3031	Waiters and Waitresses	5131	5134	Garçons	0.18
53-7062	Material Movers, Hand	9621	7832	Entregadores	0.05
47-2021	Brickmasons	7112	7152	Pedreiros	0.04
37-2011	Janitors and Cleaners	9112	5142	Faxineiros	0.00

Notes: Selected examples from the occupation crosswalk. SOC = Standard Occupational Classification (U.S., O*NET). CBO = Classificação Brasileira de Ocupações 2002 (CAGED administrative data). COD = Classificação de Ocupações para Pesquisas Domiciliares 2010 (PNAD-C, based on ISCO-08). CBO and COD use different coding schemes; occupations are matched by task content. Exp. = Eloundou et al. (2024) GPT exposure score (0–1).

coding approaches in Li et al. (2023) and Achananuparp et al. (2025). The model assigned confidence ratings to flag uncertain matches. Proposed mappings were then manually reviewed by examining occupation descriptions, typical tasks, and required education levels from O*NET's task statements and IBGE's CBO/COD documentation. Occupation titles that appear similar across countries often differ in task content, and Brazil-specific occupations (e.g., *agente comunitário de saúde*) required careful matching to functionally equivalent U.S. categories. Medium- and low-confidence matches received particular scrutiny, with about 15% of initial proposals overridden after examining auxiliary information. When multiple SOC codes mapped to one Brazilian code, exposure scores were averaged using 2021 U.S. employment weights from BLS OEWS BUREAU OF LABOR STATISTICS (2024).

A.2.2 Summary Statistics

Table A.2: Summary Statistics by AI Exposure

	Low Exposure		High Exposure	
Variable	Mean	SD	Mean	SD
Log Wage	7.230	0.890	7.857	0.933
Weekly Hours	38.12	13.04	39.95	11.85
Formal Employment	0.556	–	0.765	–
Age	41.0	–	39.5	–
Female	0.351	–	0.495	–
White	0.360	–	0.474	–
Black	0.108	–	0.084	–
Pardo (Brown)	0.522	–	0.432	–
Education (years)	8.5	–	12.4	–
Observations	3,100,179		2,960,928	

A.3 Identification and Parallel Trends

The credibility of difference-in-differences estimates rests on the parallel trends assumption: absent ChatGPT's release, outcomes would have evolved similarly across high- and low-exposure occupations. This section presents evidence bearing on this assumption and examines sensitivity to potential violations.

A.3.1 Pre-Trends Tests

Table A.3 reports the *aggregated F-test* on pre-treatment event-study coefficients, one of the two standard parallel-trends diagnostics in applied DiD work (see Goodman-Bacon, 2021; Roth et al., 2023). For each outcome, we restrict to the pre-treatment period (2017Q1–2022Q4 for PNAD; October 2020–October 2022 for CAGED, excluding April–September 2021), aggregate to the occupation-period panel, and test the joint nullity of the pre-treatment interactions from equation (2). Standard errors are clustered at the 4-digit occupation level, exactly as in the main specification. A rejection is treated as a parallel-trends failure at the 5% level. The alternative diagnostic, an individual-level exposure $\times$ linear time trend π estimated on the individual microdata, is reported alongside in Table 1 and in the appendix robustness tables (e.g., Tables A.17, A.8); the two constructions agree qualitatively on all headline outcomes. We prefer the aggregated F-test for this comprehensive table because it is well-defined even for CAGED outcomes with 19 pre-treatment months, where the individual-level test has limited power. As Roth (2022) notes, no pre-trends test has unconditional power against all violations; we therefore complement the test with date placebos (Section A.4.2) and explicit sensitivity bounds rather than relying on it alone.

Across PNAD aggregate, age-group, and worker-type specifications and CAGED flow and composition outcomes, several outcomes fail the aggregated F-test at the 5% level: PNAD early-career (22–29) log wage (p=0.038), PNAD prime-age (30–54) weekly hours (p=0.041), PNAD self-employed log wage (p=0.04), and CAGED log separations and log dismissals (both $p<0.05$). We do not attempt to rescue these failing outcomes as causal estimates: in the discussion below and in the main text, results for these outcomes are reported only as descriptive correlations. The headline findings (young worker wage penalty, aggregate formality decline, education decline among new hires) all pass the aggregated F-test in their respective canonical specifications. The CAGED quits coefficient, reported in Table 3, is directionally consistent but fails the aggregated F-test (p=0.001), so we treat it as suggestive rather than as a robust headline.

Table A.3: Comprehensive Pre-trends Assessment

Source	Outcome	Coefficient	SE	Pre-trends p	Status
PNAD Aggregate	Log Wage	-0.0090	0.0122	0.103	PASS
PNAD Aggregate	Weekly Hours	0.1718	0.1212	0.268	PASS
PNAD Aggregate	Formal Employment	-0.0089	0.0041	0.298	PASS
PNAD Young (16-21)	Log Wage	-0.0481	0.0144	0.116	PASS
PNAD Young (16-21)	Weekly Hours	-0.1518	0.1895	0.439	PASS
PNAD Young (16-21)	Formal Employment	-0.0153	0.0061	0.509	PASS
PNAD Early Career (22-29)	Log Wage	-0.0098	0.0135	0.038	FAIL
PNAD Early Career (22-29)	Weekly Hours	0.1468	0.1660	0.874	PASS
PNAD Early Career (22-29)	Formal Employment	-0.0130	0.0055	0.863	PASS
PNAD Prime (30-54)	Log Wage	-0.0044	0.0128	0.330	PASS
PNAD Prime (30-54)	Weekly Hours	0.1922	0.1272	0.041	FAIL
PNAD Prime (30-54)	Formal Employment	-0.0045	0.0047	0.092	PASS
PNAD Older (55+)	Log Wage	-0.0072	0.0175	0.095	PASS
PNAD Older (55+)	Weekly Hours	0.1830	0.1829	0.224	PASS
PNAD Older (55+)	Formal Employment	-0.0066	0.0047	0.975	PASS
PNAD Non-young (22+)	Log Wage	-0.0024	0.0120	0.160	PASS
PNAD Non-young (22+)	Weekly Hours	0.2408	0.1221	0.149	PASS
PNAD Non-young (22+)	Formal Employment	-0.0067	0.0040	0.202	PASS
CAGED Flows	Hires	-0.0328	0.0322	0.056	PASS
CAGED Flows	Separations	0.0210	0.0285	0.031	FAIL
CAGED Flows	Dismissals	0.0501	0.0293	0.025	FAIL
CAGED Flows	Quits	-0.0607	0.0310	0.028	FAIL

Notes: Comprehensive assessment of pre-treatment trends for all outcomes. Pre-trends p-value tests the null hypothesis of no differential pre-treatment trends between high- and low-exposure occupations. Status: PASS if $p \geq 0.05$, FAIL otherwise. PNAD sample: 2017Q1–2024Q4 (21 pre-treatment quarters, excluding the 2020Q1–Q3 COVID gap). The "PNAD Non-young (22+)" block pools all workers aged 22 and above; it provides the self-contained reference for the footnote citation of the pooled 22+ log-wage pre-trend (see Appendix **??**). CAGED sample: October 2020–December 2024 (19 pre-treatment months, excluding April–September 2021). Results with PASS status support the parallel trends assumption.

A.3.2 Sensitivity to Parallel Trends Violations

Pre-trends tests have limited statistical power; passing one does not confirm the parallel-trends assumption (Roth, 2022). Following Rambachan and Roth (2023), we construct confidence intervals that remain valid under bounded post-treatment violations of parallel trends. The bound is expressed as a fraction $\bar{M}$ of the largest pre-treatment violation: $\bar{M} = 0$ reproduces the standard DiD interval, $\bar{M} = 1$ allows post-treatment violations as large as the worst pre-treatment violation, and $\bar{M} = 2$ allows violations twice that size.

Table A.4 reports the resulting intervals for the main PNAD and CAGED outcomes using the relative-magnitudes approach. Before reading the table, one methodological caveat: the $\bar{M} = 0$ interval is not the DiD confidence interval reported in Table 1 or Table 3. HonestDiD takes the pre- and post-treatment event-study coefficients as input and constructs an interval around a uniformly-weighted average of the post-treatment coefficients; this quantity is mechanically different from the observation-count-weighted DiD estimator used in the headline tables, even when exact parallel trends are imposed. The $\bar{M} = 0$ intervals should therefore be read as sensitivity bounds for the uniformly-averaged event-study summary, not as a replacement for the headline DiD intervals. Three readings of the table are worth separating. First, at $\bar{M} = 0$, log hires is the only outcome whose interval excludes zero; log wage, weekly hours, formal employment, and young-worker log wage have intervals that already straddle zero even with exact parallel trends imposed. Second, as $\bar{M}$ rises to 0.5, most intervals widen enough to cover zero by a wide margin: the young-worker wage interval is $[-0.167, 0.121]$ at $\bar{M} = 0.5$ and widens further from there. Third, weekly hours and log hires are essentially uninformative at $\bar{M} \geq 0.5$: their intervals at that threshold span roughly two standard deviations of the outcome, so the test cannot distinguish any plausible effect from zero.

Table A.4: Sensitivity to Parallel Trends Violations (HonestDiD)

	Log Wage		Weekly Hours		Formal		Hourly Wage		You
$\bar{M}$	95% CI	Incl. 0?	95% CI	Incl. 0?	95% CI	Incl. 0?	95% CI	Incl. 0?	95% C
0	[-0.003, 0.023]	Yes	[-0.046, 0.308]	Yes	[-0.005, 0.008]	Yes	[-0.006, 0.017]	Yes	[-0.045, 0.
0.5	[-0.073, 0.094]	Yes	[-0.621, 0.912]	Yes	[-0.033, 0.038]	Yes	[-0.087, 0.097]	Yes	[-0.148, 0.
1	[-0.136, 0.136]	Yes	[-1.358, 1.649]	Yes	[-0.065, 0.065]	Yes	[-0.119, 0.119]	Yes	[-0.270, 0.
2	[-0.136, 0.136]	Yes	[-1.841, 1.841]	Yes	[-0.065, 0.065]	Yes	[-0.119, 0.119]	Yes	[-0.273, 0.

Notes: HonestDiD sensitivity analysis following Rambachan and Roth (2023). This method tests robustness to violatic the largest pre-treatment deviation. $\bar{M} = 0$ assumes parallel trends hold exactly (standard DiD); $\bar{M} = 1$ allows post-tr significant at higher $\bar{M}$. Both aggregate outcomes include zero at all $\bar{M} > 0$, consistent with null findings.

We read this as a reminder that the paper's individual coefficients are fragile once the parallel-trends assumption is relaxed even a little. The pattern-level evidence summarized in Section 4.3 does not rest on any single coefficient surviving high values of $\bar{M}$; rather, it rests on a consistent sign and rough magnitude across independent exposure measures, placebo tests, adjacent-quartile comparisons, and two separate data sources.

A.3.3 Multiple Testing Corrections

With multiple outcomes and specifications, some statistically significant results may arise by chance. Table A.5 applies Benjamini–Hochberg corrections to control the false discovery rate. The young worker wage penalty and several CAGED findings survive correction; results with raw p-values closer to 0.05 do not.

Table A.5: Multiple Testing Corrections (Benjamini–Hochberg): Main vs. State-Disaggregated Specifications

	Main specification (primary)			State-disaggregated (robustness)			
Outcome	Raw p	BH p	BH<.05?	Raw p (state)	BH p (state)	BH<.05 (state)?	Pre-trends
Panel A. Survives BH and passes pre-trends under main specification (lead-on findings)							
PNAD young (16–21) wage	<0.001	0.008	✓	<0.001	0.004	✓	Pass
Panel B. Survives BH under main specification but fails pre-trends (directional only)							
(none)							
Panel C. Does not survive BH under main specification							
PNAD young (16–21) formal	0.013	0.057	—	0.013	0.024	✓	Pass
PNAD formal (agg)	0.031	0.094	—	0.031	0.047	✓	Pass
CAGED log quits	0.051	0.114	—	0.011	0.024	✓	Fail
CAGED log dismissals	0.088	0.158	—	<0.001	0.003	✓	Fail
PNAD weekly hours (agg)	0.157	0.236	—	0.157	0.177	—	Pass
CAGED edu hires (young)	0.191	0.246	—	0.013	0.024	✓	Pass
CAGED log hires	0.309	0.348	—	0.064	0.082	—	Pass
PNAD log wage (agg)	0.459	0.459	—	0.459	0.459	—	Pass

Notes: Multiple testing corrections for 9 primary hypotheses corresponding to the main findings discussed in Section 1 and Section 4.3: aggregate PNAD outcomes (log wage, weekly hours, formality), young-worker PNAD outcomes (ages 16–21: wage, formality), and CAGED formal-sector outcomes (log hires, log quits, log dismissals, education of young hires). BH = Benjamini–Hochberg false discovery rate control at 5%. ✓ = significant at 5%; — = not significant. Pre-trends: Pass if $p \geq 0.05$; see Table A.3. **Specification note.** The primary (left-hand) columns report raw p-values from the *main* specification used in the paper narrative: PNAD outcomes at the individual-quarter level with occupation, period, and state FE (Table 1), and CAGED outcomes at the occupation-month aggregate level with occupation and year-month FE (Table 3). The secondary (right-hand) columns report p-values from the *state-disaggregated* CAGED panel (state $\times$ occupation $\times$ month) used for cross-measure comparability elsewhere in the appendix (see Table A.19 and Table A.6). Panels A–C are classified by the main-specification result: Panel A lists outcomes that survive BH and pass pre-trends under the main specification (the single lead-on finding, young-worker wages); Panel B is reserved for outcomes that would survive BH under the main specification but fail pre-trends (currently empty); Panel C lists outcomes that do not clear BH under the main specification. Several outcomes clear BH under the state-disaggregated panel but not under the main specification: this reflects the roughly order-of-magnitude larger cell count in the state-disaggregated panel, which tightens standard errors without adding identifying variation (treatment varies only at the occupation level). The main-specification columns are therefore the headline for inference; the state-disaggregated columns are retained to preserve the sign-and-state-panel robustness reading used in earlier drafts.

A.4 Robustness Checks

This section subjects the main findings to a battery of specification checks, alternative sample definitions, and placebo tests. The goal is to probe whether results are artifacts of particular modeling choices or reflect underlying patterns in the data.

A.4.1 Specification Checks

Table A.6 consolidates PNAD DiD estimates (Panel A, headline specification, same as Table 1) alongside CAGED estimates under the *state-disaggregated* robustness panel (Panel B). This is the panel used across the appendix robustness exercises (alternative exposure measures, BH correction, safe controls), and it is distinct from the headline CAGED specification reported in Table 3, which aggregates across states to the occupation-month level (approximately 26,182 cells) and is the specification used in the main-text narrative. The two CAGED aggregations differ in magnitude and p-values because sample size differs by roughly an order of magnitude while treatment varies only at the occupation level; signs are consistent across the two. Subsequent robustness tables in this section continue on the state-disaggregated panel. To keep terminology unambiguous in what follows, we call Table 3 the "main CAGED specification" and Table A.6 and the exercises downstream of it "state-disaggregated robustness."

Table A.6: Summary of DiD Estimates (Robustness Panel): PNAD and State-Disaggregated CAGED

Outcome	Source	DiD Effect	SE	p-value	Pre-trends
Panel A: PNAD Outcomes					
Log Wage	PNAD	-0.0090	(0.0122)	0.459	0.103
Formality	PNAD	-0.0089^{**}	(0.0041)	0.031	0.298
Weekly Hours	PNAD	0.1718	(0.1212)	0.157	0.268
Panel B: CAGED Outcomes					
Quits	CAGED	-0.0568^{***}	(0.0197)	0.004	0.109
Hires	CAGED	-0.0418^{**}	(0.0175)	0.018	0.111
Dismissals	CAGED	0.0625^{***}	(0.0201)	0.002	0.000
Edu. (Hire)	CAGED	-0.0855^{***}	(0.0313)	0.006	0.059

Notes: Binary DiD estimates (High Exposure × Post). Panel A (PNAD) *is* the main specification and matches Table 1. Panel B (CAGED) is a *robustness* specification estimated on the state-disaggregated panel (state × occupation × month, approximately 322,175 cells) used across the appendix robustness exercises (alternative exposure measures, BH correction, controls) and documented in Section A.4.1; the headline CAGED specification is the occupation-month aggregation reported in Table 3. The state-disaggregated specification in Panel B is a robustness check, not the headline CAGED specification; its tighter standard errors reflect sample-size inflation rather than additional identifying variation, since treatment varies only at the occupation level. The two CAGED aggregations differ in magnitude and p-values because sample size differs by roughly an order of magnitude while treatment varies at the occupation level; signs are consistent across the two, but the occupation-month aggregation in Table 3 is the one we use in the main-text narrative. PNAD: 2017Q1–2024Q4 with occupation, period, and state FE. CAGED: Oct 2020–Dec 2024 (excluding Apr–Sep 2021) with occupation and year-month FE. **FE-structure sensitivity.** The CAGED rows here use occupation and year-month FE only (no state FE), whereas Table A.19 estimates the same state-disaggregated panel with occupation, state, and month FE. Coefficients and pre-trends results differ across the two: e.g., Eloundou quits is -0.0568^{***} with pre-trends p=0.109 (passes) here but -0.0433^{**} with failing pre-trends under the additive occ-state-month FE of Table A.19. This reflects the fact that adding state FE to the state-disaggregated panel absorbs within-state time variation that carries part of the treatment signal; we retain both sets of FE in the appendix because each addresses a different concern, but neither is the main specification (Table 3). Log Wage (Hire) is reported only in the main CAGED specification (Table 3); the state-disaggregated panel used here is defined only for flow counts and composition means, since cell-level wage estimation at the state × occupation × month level is too sparse to be informative. Pre-trends: p-value from testing differential pre-trend ($p \geq 0.05$ supports parallel trends). *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Controlled Specification The baseline in Equation (1) omits individual-level covariates. This subsection reports a controlled variant that adds only those covariates that plausibly satisfy the additional assumptions imposed by conditional DiD.

Assumptions. Conditioning on a vector X replaces unconditional parallel trends with its conditional analogue and imposes three further requirements (Roth et al., 2023; Callaway and Sant'Anna, 2021): (i) overlap in X across treated and comparison units; (ii) X unaffected by treatment, so conditioning does not block part of the effect or induce collider bias (Angrist and Pischke, 2008); and (iii) for linear covariates, the Caetano and Callaway (2024) linearity condition, since within transformations distort linear controls more aggressively than dummies. Given that the pre-treatment event-study coefficients in Figure 5 are consistent with parallel trends for the headline PNAD outcomes (log quits is the one exception, with pre-trends p=0.028, and we read it as suggestive for that reason), conditioning trades a defensible assumption for a stronger one (Huntington-Klein, 2021); we treat it as a consistency check, not an upgrade.

Covariate triage. Table A.7 records how each PNAD-C covariate fares against these requirements. Only covariates that pass are retained.

Table A.7: Individual-Level Covariate Triage

Covariate	Pre-determined	Unaffected by treatment	No hidden linearity	Overlap plausible	Verdict
Sex (female)	Yes	Yes	n/a (binary)	Yes	Include
Race (white)	Yes	Yes	n/a (binary)	Yes	Include
Age (5-year bins)	Mechanical	Yes	n/a (dummies)	Yes	Include
Age (linear) + age^2	Mechanical	Yes	**No**	Yes	Exclude
Education (current)	**No**	**No**	–	Yes	Exclude
Urban/rural (current)	**No**	**No**	–	Yes	Exclude
Industry (current, V40132)	**No**	**No**	–	Yes	Exclude
State	Already in FE	–	–	–	Keep in FE

Notes: Age enters as 5-year bin dummies rather than a linear term to avoid the hidden-linearity bias documented by Caetano and Callaway (2024). Education, urban status, and industry are excluded as post-treatment controls in the sense of Angrist and Pischke (2008): workers may upskill, relocate, or switch industry in response to exposure.

Results. Table A.8 reports the controlled estimates for the three PNAD outcomes. Columns (1), (4), and (7) replicate the uncontrolled specification on the subsample for which all retained covariates are observed; columns (2), (5), and (8) add female and white dummies and a 5-year age-bin fixed effect; columns (3), (6), and (9) saturate age by period. The sample is held constant within outcome across columns.

Table A.8: PNAD DiD: Controlled Specification (Bad-Control-Safe Covariates)

	Log Wage			Weekly Hours			Formal Employment		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High Exposure × Post	-0.009	-0.006	-0.006	0.172	0.180	0.177	-0.009**	-0.008*	-0.007*
	(0.012)	(0.012)	(0.012)	(0.121)	(0.116)	(0.113)	(0.004)	(0.004)	(0.004)
Female		-0.260***	-0.260***		-3.47***	-3.47***		-0.014**	-0.014**
		(0.023)	(0.023)		(0.427)	(0.427)		(0.007)	(0.007)
White		0.112***	0.112***		0.088	0.088		0.001	0.001
		(0.013)	(0.013)		(0.098)	(0.098)		(0.002)	(0.002)
Pre-trends p (safe)	[0.103]	[0.192]	[0.192]	[0.268]	[0.215]	[0.215]	[0.298]	[0.326]	[0.326]
Age-bin × Period FE	No	No	Yes	No	No	Yes	No	No	Yes
Age-bin FE	No	Yes	–	No	Yes	–	No	Yes	–
Sex + Race	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	5,682,355	5,682,355	5,682,355	6,061,106	6,061,106	6,061,106	3,899,812	3,899,812	3,899,812
R^2	0.427	0.466	0.466	0.127	0.159	0.159	0.264	0.283	0.284

Dependent variable in columns (1)–(3): log monthly wage; (4)–(6): weekly hours; (7)–(9): formal employment. Column (1)/(4)/(7) replicate the main uncontrolled DiD specification (occupation, period, and state fixed effects) on the individual-level microdata held constant across columns; the reported pre-trends p-values are estimated on the same individual-level panel using an exposure × linear time trend, so they differ in construction (and slightly in sample, by one observation) from the occupation-period-state aggregated pre-trends reported in Table 1 (formality: 0.108 here vs. 0.298 there). The DiD coefficients in column (1)/(4)/(7) match Table 1 exactly. Safe-controls columns add female and white dummies and a 5-year age-bin fixed effect; the saturated column replaces age-bin FE with age-bin × period FE. Current education, current industry (V40132), and current urban status are excluded because they are potentially affected by treatment; linear age is excluded to avoid hidden linearity bias (Caetano & Callaway, 2024). Sample is held constant across columns within each outcome. Cluster-robust standard errors (4-digit CBO) in parentheses. ***p<0.01, **p<0.05, *p<0.10.

Pre-trends p-values remain above conventional thresholds across specifications, so conditioning does not improve pre-period balance. Point estimates move only modestly. The young-worker wage effect in Table A.9 survives with attenuation well inside the uncontrolled confidence interval.

Table A.9: Young Worker (Age 16–21) Log Wage: Controlled Specification

	Log Wage	
	(1)	(2)
High Exposure × Post	-0.048***	-0.045***
	(0.014)	(0.014)
Female		-0.118***
		(0.018)
White		0.040***
		(0.006)
Age-bin FE	No	Yes
Sex + Race	No	Yes
Observations	394,357	394,357
R^2	0.263	0.287

Sample restricted to workers aged 16–21. Column (1) replicates the uncontrolled young-worker wage specification; column (2) adds safe controls (female, white) and 5-year age-bin FE. See the main controlled-specification table notes for covariate triage. Cluster-robust standard errors (4-digit CBO) in parentheses. ***p<0.01, **p<0.05, *p<0.10.

Occupation Classification Sensitivity The baseline specification uses 4-digit CBO occupation codes for both fixed effects (434 categories) and standard error clustering. This subsection examines sensitivity to occupation classification granularity along two dimensions: (1) clustering level for standard errors, and (2) fixed effects granularity.

Alternative Clustering. Table A.10 compares standard errors under 4-digit CBO clustering (baseline) vs. 2-digit CBO clustering. The key findings (young wage decline, education decline among young hires) remain highly significant under both clustering levels. CAGED flow outcomes (hires, quits) lose significance at 2-digit clustering, reflecting fewer clusters for inference.

Fixed Effects Granularity. Pre-trends tests are sensitive to the granularity of the occupation fixed effects. With 4-digit codes, pre-treatment coefficients cluster around zero (mean = 0.016) and pre-trends tests pass for the main outcomes. At 3-digit aggregation (125 codes), pre-treatment coefficients drift upward (mean = 0.095); at 2-digit aggregation (39 codes), pre-treatment mean reaches 0.163, with most coefficients statistically significant. We interpret this pattern as a measurement artifact rather than a violation of the underlying research design: broad 2- and 3-digit categories bundle occupations with different exposure scores and different underlying wage trajectories, so within-category comparisons mechanically pick up composition effects that the 4-digit fixed effects absorb. This is consistent with the high within-3-digit variance of AI exposure documented in Table A.2. We report 4-digit as the baseline for this reason rather than because coarser specifications produce different treatment effects: the point estimates are similar across aggregation levels (Table A.12), but only the 4-digit specification cleanly passes pre-trends.

Age Bin Symmetrization To check that the PNAD–CAGED cross-source comparison for prime-age and older workers is not driven by the one-year asymmetry in age bins (PNAD Prime 30–54, Older 55+; CAGED Prime 30–49, Older 50+), Table A.11 re-estimates PNAD age heterogeneity with the CAGED-compatible bins. Point estimates for Prime (30–49) and Older (50+) remain null across all three outcomes, matching the canonical bins. The "incumbents untouched" reading does not depend on the choice of bin boundary.

Table A.10: Robustness to Alternative Clustering

Outcome	Clustering	Coefficient	SE	p-value
Panel A: PNAD Outcomes				
Log Wage	4-digit CBO	-0.0090	(0.0122)	0.459
	2-digit CBO	-0.0090	(0.0140)	0.521
Weekly Hours	4-digit CBO	0.1718	(0.1212)	0.157
	2-digit CBO	0.1718^{*}	(0.0954)	0.080
Formal	4-digit CBO	-0.0089^{**}	(0.0041)	0.031
	2-digit CBO	-0.0089^{*}	(0.0049)	0.078
Young Wage	4-digit CBO	-0.0481^{***}	(0.0144)	0.001
	2-digit CBO	-0.0481^{***}	(0.0145)	0.002
Panel B: CAGED Outcomes				
Hires	4-digit CBO	-0.0549^{***}	(0.0175)	0.002
	2-digit CBO	-0.0549^{***}	(0.0180)	0.004
Quits	4-digit CBO	-0.0601^{***}	(0.0187)	0.001
	2-digit CBO	-0.0601^{***}	(0.0220)	0.009
Dismissals	4-digit CBO	0.0434^{**}	(0.0195)	0.026
	2-digit CBO	0.0434	(0.0262)	0.104
Edu. (Hire)	4-digit CBO	-0.0246	(0.0512)	0.631
	2-digit CBO	-0.0246	(0.0767)	0.750

Notes: Comparison of standard errors under 4-digit CBO clustering (baseline) vs. 2-digit CBO clustering. The point estimates for PNAD outcomes (Panel A) match Table 1 because the individual-level spec is identical; within each outcome the coefficient is constant across clustering rows and only the standard error varies. The CAGED point estimates (Panel B) differ from the caged_flows table because Panel B here is estimated at the state-occupation-month level with state fixed effects (the same spec used in the original robustness pipeline for this table), whereas the main CAGED table collapses the panel to occupation-month with only occupation and year-month fixed effects; within each CAGED outcome, coefficients are constant across clustering rows so that only the standard errors move. Young wage and CAGED hires and quits remain significant at the 1% level under both clustering schemes; the formal-employment coefficient and CAGED dismissals lose significance at the 5% level under 2-digit clustering (formal: p = 0.078; dismissals: p = 0.102). $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

Table A.11: PNAD Effects under CAGED-Compatible Age Bins

	Log Wage (1)	Weekly Hours (2)	Formal (3)
Young (16–21)	-0.0481^{***}	-0.1518	-0.0153^{**}
	(0.0144)	(0.1895)	(0.0061)
Pre-trends p-value	[0.116]	[0.439]	[0.509]
Early Career (22–29)	-0.0098	0.1468	-0.0130^{**}
	(0.0135)	(0.1660)	(0.0055)
Pre-trends p-value	[0.038]	[0.874]	[0.863]
Prime (30–49)	-0.0026	0.2107	-0.0044
	(0.0130)	(0.1285)	(0.0051)
Pre-trends p-value	[0.500]	[0.027]	[0.146]
Older (50+)	-0.0072	0.1842	-0.0040
	(0.0151)	(0.1767)	(0.0042)
Pre-trends p-value	[0.062]	[0.571]	[0.156]

Notes: Replicates Table 1 Panel B with prime-age and older bins redefined to match CAGED conventions (30–49 and 50+ rather than 30–54 and 55+), allowing apples-to-apples cross-source comparison for the incumbent age groups. Point estimates for prime and older workers remain null across all three outcomes under either bin definition, consistent with the "incumbents untouched" reading. Young (16–21) and Early Career (22–29) rows are identical to Table 1 Panel B because those bin boundaries do not change across PNAD and CAGED conventions. Binary difference-in-differences comparing high- versus low-exposure occupations (median split), with occupation, period, and state fixed effects. Standard errors clustered at occupation level in parentheses. Pre-trends p-values test differential pre-treatment trends; $p \geq 0.05$ supports parallel trends. Sample: PNAD-C 2017Q1–2024Q4. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

Alternative Fixed Effects Table A.12 varies the fixed effects structure to test sensitivity to unobserved heterogeneity across both PNAD and CAGED outcomes. We report four specifications: (1) occupation and period FE only, a stripped-down specification that omits state variation; (2) the baseline, adding state FE to absorb time-invariant regional level differences; (3) occupation $\times$ state FE, which absorbs all time-invariant occupation–state combinations; and (4) occupation FE combined with region $\times$ period FE, which allows regional business cycles to load differently over time.

PNAD results remain stable across all specifications: wage coefficients range from -0.013 to -0.015, hours coefficients from 0.15 to 0.17, and formal employment coefficients from -0.009 to -0.010. The formal employment effect becomes statistically stronger with more demanding FE structures. CAGED results are similarly robust: hires coefficients range from -0.027 to -0.029 (all significant at 1%), quits from -0.030 to -0.032 (all significant at 1%), and dismissals from 0.024 to 0.027 (all significant at 5%). The consistency across both individual-level (PNAD) and administrative (CAGED) data, and across increasingly demanding fixed effects structures, strengthens confidence in the main findings.

Table A.12: Robustness to Alternative Fixed Effects Specifications

		PNAD			CAGED		
Spec	Fixed Effects	Log Wage	Hours	Formal	Hires	Quits	Dismissals
(1)	Occ + Period	−0.0132	0.1659	−0.0085**	−0.0281***	−0.0302***	0.0259**
		(0.0138)	(0.1277)	(0.0043)	(0.0077)	(0.0106)	(0.0105)
(2)	Occ + State + Period	−0.0146	0.1617	−0.0087*	−0.0274***	−0.0303***	0.0268**
		(0.0132)	(0.1299)	(0.0044)	(0.0081)	(0.0110)	(0.0113)
(3)	Occ × State + Period	−0.0149	0.1531	−0.0095**	−0.0280***	−0.0312***	0.0252**
		(0.0131)	(0.1294)	(0.0044)	(0.0081)	(0.0109)	(0.0112)
(4)	Occ + Reg × Period	−0.0137	0.1610	−0.0092**	−0.0285***	−0.0319***	0.0244**
		(0.0129)	(0.1293)	(0.0044)	(0.0080)	(0.0110)	(0.0111)

Notes: Robustness of main results to alternative fixed effects specifications. (1) Occupation and period/month FE only (no state variation). (2) Main specification: adds state FE to control for time-invariant state differences. (3) Occupation × state FE absorbs all time-invariant occupation-state heterogeneity. (4) Occupation FE with region × period interactions to control for regional business cycles. Small differences from Table 1 at the fourth decimal (e.g., formality −0.0087 here vs. −0.0089 in Table 1) reflect differences in the estimation panel: the FE-robustness routine re-estimates the baseline on a panel rebuilt for state-disaggregated specifications and harmonized across the four columns, which drops a small number of cells with missing state or weights relative to Table 1. Signs, significance, and magnitudes are preserved. PNAD sample: 2017Q1–2024Q4. CAGED sample: October 2020–December 2024. Standard errors clustered at occupation level. ***p<0.01, **p<0.05, *p<0.10.

A.4.2 Placebo Tests

Placebo tests probe whether the estimated effects could arise spuriously. Table A.13 reports three types of placebos:

- **Date placebos**: Using fake treatment dates 6–9 months before November 2022 (2022Q1 and 2022Q2). Formality shows no spurious effects (p=0.722, p=0.208), but wages show significant effects at both fake dates (p=0.007, p=0.009). However, the formal pre-trends test (exposure × time trend) passes for wages (p=0.103), and event study coefficients show no significant pre-treatment differences. The date placebo compares a short window (3 quarters) against a long baseline (17+ quarters), potentially detecting level shifts rather than systematic trends.
- **Exposure randomization**: Randomly permuting AI exposure across occupations and re-estimating the main DiD. We report this exercise separately in Appendix A.4.6, where the permutation holds the canonical specification fixed.
- **Compositional placebos**: Race and gender are fixed at the individual level, so running the baseline DiD with the share of treated cells that are (say)

female as the outcome is not a placebo on a causal channel: a significant coefficient would indicate differential sample selection across exposure groups rather than a treatment effect on a fixed characteristic. Neither race (p=0.870) nor gender (p=0.132) shows significant effects, suggesting results do not reflect compositional selection on these characteristics.

Table A.13: Placebo Tests

Placebo Type	Outcome	Coefficient	SE	p-value	Status
Panel A: Date Placebo (Fake Treatment Dates)					
Fake 2022Q1	Log Wage	−0.0248	(0.0091)	[0.007]	Fail
Fake 2022Q2	Log Wage	−0.0275	(0.0105)	[0.009]	Fail
Fake 2022Q1	Formal	−0.0014	(0.0039)	[0.722]	Pass
Fake 2022Q2	Formal	−0.0051	(0.0040)	[0.208]	Pass
Panel B: Exposure Randomization					
Random exposure	Log Wage	0.0142	(0.0126)	[0.260]	Pass
Random exposure	Formal	−0.0042	(0.0036)	[0.238]	Pass
Panel C: Outcome Placebo					
Placebo outcome	Race (White)	−0.0004	(0.0022)	[0.870]	Pass
Placebo outcome	Gender (Female)	0.0049	(0.0032)	[0.132]	Pass
Panel D: Young Worker (16–21) Date Placebo					
Fake 2019Q1	Log Wage	−0.0079	(0.0104)	[0.446]	Pass
Fake 2020Q4	Log Wage	−0.0109	(0.0115)	[0.342]	Pass
Fake 2021Q2	Log Wage	−0.0201	(0.0134)	[0.136]	Pass
Fake 2022Q1	Log Wage	−0.0222	(0.0159)	[0.165]	Pass
Fake 2022Q2	Log Wage	−0.0216	(0.0163)	[0.186]	Pass

Notes: Placebo tests for identification validity. Panel A: fake treatment dates before ChatGPT's November 2022 release; spurious effects would indicate pre-existing trends. Earlier dates (2019Q1, 2020Q4) pass for wages; later dates (2021Q2+) fail, suggesting wage divergence began in mid-2021. Formality passes most dates. Panel B: random assignment of AI exposure across occupations; significant effects would suggest results depend on arbitrary groupings rather than actual exposure. Panel C: time-invariant demographics (race, gender) theoretically unaffected by AI; significant effects would indicate compositional selection. Panel D: date placebos for young workers (16–21) on log wages; all five fake dates pass, supporting parallel trends for this key finding. "Pass" = $p \geq 0.05$; "Fail" = $p < 0.05$.

A.4.3 Sample Period Sensitivity

Shorter Pre-Period (PNAD) The main specification uses 2017Q1–2024Q4, providing 21 pre-treatment quarters (the 2020Q1–Q3 COVID gap is excluded because IBGE did not release standard quarterly PNAD-C files) for parallel trends validation. Table A.14 examines sensitivity to a shorter pre-period, restricting the sample to 2021Q4–2024Q4 with 5 pre-treatment quarters (2021Q4–2022Q4) and post-treatment beginning 2023Q1.

Table A.14: Short Pre-Period Robustness

	PNAD (2021Q4–2024Q4)			CAGED
	Log Wage	Hours	Formal	Log Hires
Panel A: Aggregate Effects				
High Exp. × Post	0.0084	0.1295*	−0.0094***	−0.0328
	(0.0075)	(0.0785)	(0.0034)	(0.0322)
Pre-trends p-value	[0.255]	[0.998]	[<0.001]	[0.056]
Panel B: Age Heterogeneity				
Young (16–21)	−0.0239**	0.0158	−0.0140*	−0.0575**
	(0.0109)	(0.2103)	(0.0072)	(0.0281)
Pre-trends	[0.980]	[0.465]	[0.822]	[<0.001]
Early Career (22–29)	0.0132	0.2571**	−0.0112**	−0.0623**
	(0.0084)	(0.1049)	(0.0050)	(0.0307)
Pre-trends	[0.387]	[0.237]	[0.004]	[0.066]
Prime (30–54)	0.0091	0.0722	−0.0063*	−0.0304
	(0.0084)	(0.0923)	(0.0036)	(0.0317)
Pre-trends	[0.501]	[0.324]	[0.005]	[0.142]
Older (55+)	0.0131	0.2981**	−0.0074	−0.0290
	(0.0122)	(0.1449)	(0.0055)	(0.0298)
Pre-trends	[0.141]	[0.593]	[0.008]	[0.327]

Notes: Short pre-period robustness. PNAD: 5 pre-treatment quarters (2021Q4–2022Q4) vs. 21 in main specification. CAGED log hires by age group (Oct 2021–Dec 2024). All specifications include occupation, period/year-month, and state fixed effects. Pre-trends p-values in brackets; $p \geq 0.05$ supports parallel trends. Standard errors clustered at occupation level. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

With the shorter pre-period (5 pre-treatment quarters vs. 21 in the main specification), aggregate hours shows a positive but insignificant effect (0.13 hours, p=0.100), with pre-trends validated (p=0.998). The young worker wage penalty relative to young workers in low-exposure occupations is smaller in the short panel (2.4% vs. 4.7% in the extended panel) but remains statistically significant (p=0.029), with pre-trends validated (p=0.980). These patterns confirm the extended pre-period provides more reliable identification by allowing longer baseline comparison.

Sensitivity to CAGED Pre-Period Length The main CAGED specification covers October 2020–December 2024 (excluding April–September 2021), providing 19 pre-treatment months for parallel trends testing. Because the Novo CAGED transition and early pandemic disruption both complicate the choice of start date, we report a shorter-pre-period variant beginning in October 2021 (13 pre-treatment months). If early-pandemic noise were driving the main results, the shorter sample should remove it.

Table A.15 compares results across the two pre-period definitions. Point estimates are qualitatively stable in the shorter sample, with quits declining in both and education of hires showing negative effects; the shorter sample's pre-trends tests are less powerful due to the compressed window, so we retain October 2020 as the main specification.

Table A.15: CAGED Results: Sensitivity to Pre-Period Length

	Shorter Pre-Period (Oct 2021–Dec 2024)		Main Specification (Oct 2020–Dec 2024)	
Outcome	Coefficient	Pre-trends	Coefficient	Pre-trends
Hires	−0.030*	Pass	−0.033	Pass
Separations	0.028*	Fail	0.021	Fail
Dismissals	0.066***	Fail	0.050*	Fail
Quits	−0.046**	Fail	−0.061*	Fail
Edu. (Hire)	–	–	−0.082	Pass

Notes: Comparison of CAGED DiD results across pre-period definitions. Main specification: October 2020–December 2024 (45 months, excluding April–September 2021). Shorter pre-period: October 2021–December 2024 (39 months). Both specifications run at state-occupation-month level with occupation, state, and month fixed effects (additive), weighted by pre-treatment employment, with standard errors clustered at occupation level. Pre-trends test: joint nullity of pre-treatment event study coefficients; Pass if $p \geq 0.05$. ***$p<0.01$, **$p<0.05$, *$p<0.10$.

A.4.4 Alternative Treatment Definitions

The baseline specification defines high AI exposure as above-median. Table A.16 re-estimates the baseline specification under alternative exposure cutoffs (top tercile, quartile, quintile vs. the rest). The first row reproduces the median cutoff used in the main text (Table 1) to anchor comparisons; subsequent rows vary only the threshold. Using more extreme definitions, significant negative wage effects emerge: the top quintile shows a -0.0375 log-point wage decline (p=0.001), and the top quartile shows -0.0394 (p=0.001). This pattern suggests wage effects concentrate among the most exposed occupations, while the null baseline result reflects averaging across heterogeneous exposure intensities.

The final two rows address the concern that low-exposure occupations may not constitute a clean control group since they are less exposed rather than unexposed. "Top vs bottom quintile" compares the top 20% to the bottom 20% of the exposure distribution, excluding the middle 60%. "Top quintile vs zero exposure" compares the top 20% to occupations with near-zero AI exposure scores. Effects remain significant and of similar magnitude when comparing against the bottom quintile, with all PNAD outcomes passing pre-trends tests. When comparing against near-zero exposure occupations (only 22 occupations in PNAD, 27 in CAGED), effects remain directionally consistent though less precisely estimated due to smaller sample size.

To further assess whether observed effects reflect AI exposure rather than

Table A.16: Sensitivity to Treatment Threshold

Cutoff	PNAD Log Wage	PNAD Hours	PNAD Formal	CAGED Quits	CAGED Hires	CAGED Edu. (Hire)	CAGED Edu. (Sep.)
Above median (baseline)	-0.0090	0.1718	-0.0089^{**}	-0.0568^{***}	-0.0418^{**}	-0.0855^{***}	-0.0945^{***}
Top quartile vs rest	-0.0394^{***}	0.0315	-0.0086^{***}	-0.0778^{***}	-0.0314	-0.0706^{**}	-0.1363^{***}
Top quintile vs rest	-0.0375^{***}	0.0986	-0.0051	-0.0755^{***}	-0.0290	-0.0691^{*}	-0.1544^{***}
Top tercile vs rest	-0.0376^{***}	0.0907	-0.0101^{***}	-0.0568^{***}	-0.0281	-0.0924^{***}	-0.1057^{***}
Top vs bottom quintile	-0.0407^{***}	0.0152	-0.0145^{**}	-0.0797^{***}	-0.0089	-0.1008^{**}	-0.1587^{***}
Top quintile vs zero exposure	-0.0307^{*}	-0.0484	-0.0111	-0.0403	-0.0127	-0.1234	-0.1062

Notes: Alternative treatment thresholds. "vs rest": high vs. all others. "Top vs bottom": top vs. bottom quintile. "vs zero": top vs. near-zero. PNAD: 2017Q1–2024Q4, CAGED: Oct 2020–Dec 2024. Edu. = education (16–21). Occ./period/state FE. CAGED coefficients use the state-disaggregated panel (see Section A.4.1) and therefore differ in magnitude from the occupation-month aggregation in Table 3 (e.g., quits baseline: -0.0568 here vs. -0.0607 there); sign and significance are consistent. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

other confounds, we test for dose-response patterns by replacing the binary treatment with quintile dummies (Q1 as reference). If effects are driven by AI exposure, they should concentrate at the top of the exposure distribution rather than appearing uniformly across quintiles. For PNAD log wages, effects are concentrated in Q5 (highest exposure): Q5 shows -0.045 (p=0.001), while Q2–Q4 are statistically indistinguishable from Q1. For CAGED quits, the pattern is similar: Q5 shows -0.124 (p=0.037), while Q2–Q4 cluster around zero. This threshold pattern is consistent with AI-driven adjustment rather than a spurious correlation that would likely appear across the exposure distribution.

A.4.5 Adjacent-Quartile Comparisons

The main specification compares occupations above the employment-weighted exposure median to occupations below it. That contrast mixes a wide range of exposure gaps: some treated-control pairs differ sharply in exposure, others barely. Occupations closer together in the exposure distribution are also likely to be more similar on unobservables, and a tighter comparison provides a more demanding test of parallel trends. Table A.17 therefore re-estimates the main PNAD DiD on the top two quartiles only (Q4 treated, Q3 control), dropping the bottom half of the exposure distribution.

Three patterns are worth noting. First, pre-trends p-values are uniformly larger under the Q4-vs-Q3 contrast than under the median split (in the 0.3–0.6 range versus the 0.1–0.3 range for the outcomes shown), consistent with narrower comparison groups being more likely to satisfy the parallel-trends assumption. Second, the aggregate wage coefficient sharpens: a -0.009 log-point null under the median split becomes a -0.034 log-point decline (p=0.019) under Q4-vs-Q3. Third, the young-worker wage signal strengthens rather than weakens: the median-

Table A.17: Adjacent-Quartile DiD: Q4 vs Q3 Exposure

	Log Wage		Weekly Hours		Formal Employment	
	(1)	(2)	(3)	(4)	(5)	(6)
High Exposure × Post	-0.009		0.172		-0.009**	
	(0.012)		(0.121)		(0.004)	
Q4 × Post		-0.034**		-0.061		0.005
		(0.014)		(0.116)		(0.005)
Young workers (age 16–21), Log Wage						
High Exposure × Post	-0.048***					
	(0.014)					
Q4 × Post		-0.061***				
		(0.019)				
Young obs.	394,357	153,626				
Pre-trends p-value	[0.147]	[0.551]	[0.250]	[0.339]	[0.108]	[0.512]
Comparison	Median split	Q4 vs Q3	Median split	Q4 vs Q3	Median split	Q4 vs Q3
Observations	5,682,355	1,978,147	6,061,106	2,052,215	3,899,812	1,622,884
R^2	0.427	0.421	0.127	0.106	0.264	0.140

Columns (1), (3), (5) replicate the main median-split DiD on the full sample. Columns (2), (4), (6) restrict the sample to the top two exposure quartiles (Q4 treated, Q3 control), dropping the bottom half of the exposure distribution. Pre-trends p-values are estimated on the individual-level panel using an exposure × linear time trend (the same construction as Appendix Table A.8) and therefore differ from the occupation-period aggregated p-values in Table 1 (e.g., log wage: 0.147 here vs. 0.103 there). Narrower comparison groups are more similar on unobservables, providing a more credible test of the parallel trends assumption. The lower panel restricts to young workers (age 16–21) and re-estimates the log-wage DiD under both specifications; the young-worker signal strengthens under Q4-vs-Q3. All specifications include occupation, period, and state fixed effects and use PNAD sampling weights. Cluster-robust standard errors (4-digit CBO) in parentheses. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

split estimate of -0.048 (p¡0.001) becomes -0.061 (p=0.002) under Q4-vs-Q3. We read these patterns as consistent with a dose-response interpretation: the headline relative effect is concentrated at the top of the exposure distribution, and the median-split coefficients are noisier averages across heterogeneous exposure contrasts. The aggregate formality coefficient does not survive this narrower contrast, which is one of several reasons we flag the formality decline as suggestive rather than robust (Section 4.3).

A.4.6 Intensity Permutation

The alternative-cutoffs exercise (Appendix A.4.4) and the adjacent-quartile exercise (Appendix A.4.5) both probe whether the results depend on the particular cut used to define high exposure. A complementary question is whether the particular *ranking* of occupations along the exposure measure matters. If the coefficients we report reflect the exposure measure itself, randomly permuting the measure across occupations should produce coefficients close to zero; if they instead reflect an arbitrary feature of the occupation distribution, random permutations should reproduce the headline effects roughly as often as not.

We implement this as a randomization-inference exercise: holding the full estimation pipeline fixed, we randomly permute the occupation-level exposure vector

200 times, re-estimate the DiD coefficient on each draw, and compute a two-sided p-value as the share of permuted coefficients at least as large in absolute value as the actual coefficient. Table A.18 reports the results for the four outcomes where this exercise is well defined.

Table A.18: Intensity Permutation: Randomization Inference on Exposure

	Actual coefficient	Pre-trends p-value	RI p-value	Null 95% range (permutations)	Pre-trend pass rate (null)
Log wage (aggregate)	−0.009	0.103	0.475	[−0.025, 0.025]	0.98
Weekly hours (aggregate)	0.172	0.268	0.185	[−0.223, 0.251]	0.96
Formal (aggregate)	−0.009**	0.298	0.005	[−0.007, 0.007]	0.90
Young log wage (16–21)	−0.048***	0.116	0.010	[−0.040, 0.037]	0.95

Randomization-inference p-values from 200 permutations of the occupation-level continuous AI-exposure score across CBO codes: on each draw, the exposure vector is reshuffled across occupations, re-binarized at the original employment-weighted median to form the high-exposure indicator, and the canonical DiD (occupation, period, and state fixed effects; PNAD survey weights; clustering at CBO) is re-estimated. The "Actual coefficient" column therefore matches the corresponding cells of Table 1 by construction. "RI p-value" is the two-sided share of permuted coefficients at least as large in absolute value as the actual coefficient; the 200-draw floor is 0.005. "Null 95% range" reports the 2.5th and 97.5th percentiles of the permutation distribution of the treatment coefficient. "Pre-trend pass rate (null)" is the share of permutations whose linear pre-trend test passes at the 5% level, indicating how often the pipeline accepts parallel trends under randomly assigned exposure. Significance markers on the actual coefficient reflect the main specification (Table 1), not the permutation p-value. The "Pre-trends p-value" column reports the two-sided Wald test of an exposure-by-linear-trend coefficient, estimated on the pre-treatment individual-level microdata with the canonical occupation, period, and state fixed effects, PNAD survey weights, and CBO clustering; the resulting p-values in this 200-draw run (0.103, 0.268, 0.298, 0.116 for the four outcomes) are close to the aggregated-F construction in Table 1.
***$p<0.01$, **$p<0.05$, *$p<0.10$.

For the young-worker log wage, the randomization-inference p-value is 0.010 (2 of 200 permutations produce an estimate at least as extreme as the actual coefficient in absolute value); for aggregate formality, the p-value is 0.005, at the resolution floor of the 200-draw run (1 of 200 permutations). Both coefficients lie outside the 95% range of the permutation distribution. For the aggregate log wage and weekly hours, the p-values of 0.475 and 0.185 are in line with what random permutations would produce, consistent with the small magnitudes of the actual coefficients and adding no evidence beyond what the main specification already shows. We view this exercise as weakly supportive of the pattern-level reading of the young-worker and formality results, not as an independent identification argument.

A.4.7 Alternative AI Exposure Measures

A concern with any exposure-based design is that results may depend on the particular measure chosen. Table A.19 tests robustness to two alternative AI exposure measures: Anthropic Observed Exposure (Massenkoff and McCrory, 2026), which combines Claude usage data from the Anthropic Economic Index (Handa et al., 2025) with theoretical task exposure from Eloundou et al. (2024) and an automation upweight, and Microsoft Bing Copilot applicability from Tomlinson et al. (2025).

The young-worker (16–21) log wage finding replicates across all three measures: coefficients of −0.053 (Eloundou), −0.047 (Microsoft), and −0.055

Table A.19: Robustness to Alternative AI Exposure Measures

	AI Exposure Measure		
Outcome	Eloundou et al.	Microsoft	Anthropic
Panel A: PNAD Outcomes			
Log Wage	−0.0110	−0.0083	−0.0264
	(0.0126)	(0.0209)	(0.0214)
Pre-trends	Pass	Pass	Fail
Young (16–21)	−0.0526***	−0.0467**	−0.0547**
	(0.0150)	(0.0226)	(0.0216)
Pre-trends	Pass	Pass	Pass
Early career (22–29)	−0.0112	−0.0156	−0.0224
	(0.0140)	(0.0175)	(0.0181)
Pre-trends	Fail	Pass	Pass
Weekly Hours	0.1718	0.2801	0.3365*
	(0.1212)	(0.1710)	(0.1971)
Pre-trends	Pass	Pass	Fail
Formal Employment	−0.0138***	−0.0035	−0.0026
	(0.0041)	(0.0044)	(0.0051)
Pre-trends	Pass	Fail	Pass
Panel B: CAGED Outcomes			
Log Hires	−0.0487***	−0.0100	−0.0317*
	(0.0165)	(0.0164)	(0.0167)
Pre-trends	Fail	Fail	Fail
Log Quits (voluntary)	−0.0484***	−0.0250	−0.0544***
	(0.0182)	(0.0179)	(0.0178)
Pre-trends	Pass	Pass	Pass
Log Layoffs	0.0799***	0.0233	0.0181
	(0.0194)	(0.0186)	(0.0189)
Pre-trends	Fail	Fail	Fail
Education of Hires	−0.0542*	−0.0207	−0.0393
	(0.0296)	(0.0304)	(0.0304)
Pre-trends	Pass	Pass	Fail
Wage of Hires (log)	−0.0128	−0.0167*	−0.0246**
	(0.0089)	(0.0096)	(0.0097)
Pre-trends	Fail	Fail	Fail

Notes: Robustness of main findings to alternative AI exposure measures. Panel A: PNAD outcomes (2017Q1–2024Q4) with occupation, state, and period fixed effects. The young (16–21) and early-career (22–29) rows restrict to those age bins before aggregating; the young-worker row operationalizes the BH-surviving headline of §**??**. Panel B: CAGED outcomes (Oct 2020–Dec 2024, excluding Apr–Sep 2021) run at state-occupation-month level with occupation, state, and month fixed effects. Panel B runs on the state-disaggregated panel as a robustness check, not the headline CAGED specification (Table 3); its tighter standard errors reflect sample-size inflation rather than additional identifying variation, since treatment varies only at the occupation level. The Eloundou coefficients in this table are estimated on the occupation-state-period panel (employment-weighted) used for cross-measure comparability, rather than the individual-level unweighted panel used for Table 1 or the occupation-month aggregation used for Table 3; effective cell counts and weighting therefore differ. **Eloundou et al.**: Task-based GPT exposure from O*NET Eloundou et al. (2024). **Microsoft**: Bing Copilot applicability Tomlinson et al. (2025). **Anthropic**: Observed Exposure Massenkoff and McCrory (2026), built on Anthropic Economic Index

(Anthropic), all significant at the 5% level or better, and all three pass pre-trends. This is the strongest cross-measure agreement in the table and corroborates the BH-surviving headline of Section 4. The early-career (22–29) and aggregate log wage coefficients are negative under all three measures but do not reach significance. For aggregate formality, the Eloundou estimate is -0.014 (significant at the 1% level, passing pre-trends), while Microsoft and Anthropic produce small coefficients that are not statistically distinguishable from zero. For CAGED outcomes, voluntary quits are negative under all three measures but only Eloundou clears conventional significance: -0.037 log points (p=0.027, pre-trends pass), against -0.022 for Microsoft (p=0.216) and -0.027 for Anthropic (p=0.113), with pre-trends failing for the latter two. Layoffs are positive under all three measures, with Eloundou 0.070 (p=0.001) and Anthropic 0.038 (p=0.030) reaching conventional significance and Microsoft 0.029 (p=0.118) not; pre-trends fail for layoffs under all three measures. Hiring wages are negative under all three measures, with the Anthropic and Microsoft estimates significant at the 5% and 10% levels respectively, although pre-trends fail under each. Education of hires is negative under all three measures but reaches conventional significance only under Eloundou. Cross-measure agreement is strongest for young-worker wages, where all three measures yield negative and significant coefficients with passing pre-trends. For voluntary quits and layoffs all three measures share signs but only Eloundou (and, for layoffs, Anthropic) clears both significance and pre-trends, so we read these as directionally consistent rather than as cross-measure replications in the sense the reading protocol requires.

A.5 Heterogeneity Analysis

This section examines whether AI effects vary across worker characteristics. Heterogeneity analysis helps identify which groups bear adjustment costs and provides evidence on the mechanisms through which AI affects labor markets.

A.5.1 Age Heterogeneity

Table A.20 presents extended age heterogeneity results, including the early career (22–29) age group omitted from the main text for space and reported here. Young workers (ages 16–21) show wage declines of 4.7% in high-exposure occupations, with passing pre-trends. Prime-age workers show positive hours coefficients but fail pre-trends; older workers show positive but insignificant hours effects with passing pre-trends.

Table A.20: Age Heterogeneity in AI Exposure Effects

	Young (16-21)	Early Career (22-29)	Prime (30-54)	Older (55+)
Log Monthly Income	-0.0481^{***}	-0.0098	-0.0044	-0.0072
	(0.0144)	(0.0135)	(0.0128)	(0.0175)
Pre-trends p	[0.116]	[0.038]	[0.330]	[0.095]
Weekly Hours	-0.1518	0.1468	0.1922	0.1830
	(0.1895)	(0.1660)	(0.1272)	(0.1829)
Pre-trends p	[0.439]	[0.874]	[0.041]	[0.224]
Formal Employment	-0.0153^{**}	-0.0130^{**}	-0.0045	-0.0066
	(0.0061)	(0.0055)	(0.0047)	(0.0047)
Pre-trends p	[0.509]	[0.863]	[0.092]	[0.975]

Notes: Binary DiD estimates by age group. Dependent variables: Log Monthly Income (employed workers), Weekly Hours (employed workers), Formal Employment (probability). Sample period: 2017Q1–2024Q4. All specifications include occupation, state, and period fixed effects. Pre-trends p-values test for differential trends before ChatGPT release (November 2022); $p \geq 0.05$ supports parallel trends. Standard errors in parentheses, clustered at occupation level. $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.10$.

Triple Difference Analysis As an alternative to estimating effects separately by age group, triple interactions can directly test whether AI affects young workers differently. This approach requires a stronger identifying assumption: that the age gradient in outcomes was constant across exposure levels before treatment.

A.5.2 Age Triple Difference

The main analysis estimates effects separately by age group. An alternative approach uses triple interactions to directly test whether AI affects young workers differently:

$$\begin{aligned} Y_{ijt} = \alpha_j + \delta_t + \gamma_s + \text{Young}_i + (\text{High}_j \times \text{Young}_i) \\ + \beta_1(\text{High}_j \times \text{Post}_t) + \beta_2(\text{Young}_i \times \text{Post}_t) \\ + \beta_3(\text{High}_j \times \text{Young}_i \times \text{Post}_t) + \varepsilon_{ijt} \end{aligned}$$

where Young indicates ages 16–21 (consistent with the separate young worker regression). The triple difference fails its pre-trends test (p=0.017): the young-older wage gap in high-AI versus low-AI occupations was already widening pre-treatment. The pre-trend violation invalidates causal interpretation of the triple interaction, so we do not lean on its point estimate.[5]

To gauge the magnitude of this drift, we project the pre-period triple-trend

[5] For completeness, the raw triple-interaction coefficient is -0.184 (SE 0.061, p=0.003); we report it but do not interpret it given the pre-trend failure.

slope (log points per quarter) forward to the average post-treatment quarter. The projection accounts for % of the actual triple-D static coefficient (projected versus actual). A linear extension of pre-existing differential drift therefore explains a meaningful share of the post-treatment age gradient before any AI shock, which is the substantive content of the pre-trend failure: the cross-age contrast was already in motion.

The separate-regression approach we fall back to is a weaker test, not a safer one. It requires parallel trends *within* each age group, which the data support (p=0.116 for ages 16–21; p=0.160 for ages 22+; see Table A.3 for the pooled 22+ row and individual age-group rows). But the age gradient interpretation the paper uses ("young hurt, prime and older untouched, therefore barrier-to-entry") is a cross-age comparison, and the triple difference is precisely the specification designed to test that comparison under a unified identifying assumption. The pre-trend failure in the triple difference is evidence that the age-wage gradient in high- versus low-AI occupations was already evolving differentially before ChatGPT: whatever is driving the post-treatment age gradient in the separate regressions was partly in motion pre-treatment as well, and the separate-regression design does not absorb that cross-age drift the way the triple difference would have if its own pre-trends had passed. We therefore read the age pattern as documentary rather than causally identified: the estimated magnitudes in each age-stratified regression are defensible as relative contrasts within that age group, but the comparison across age groups that motivates the barrier-to-entry reading rests on an assumption (constant pre-treatment age gradient across exposure levels) that the data explicitly reject. The finer age breakdown reinforces this caution: the pooled 22+ group passes pre-trends, but within it the early career band (22–29) fails the wage pre-trends test (p=0.038; Table A.20), so the pooled result is papering over heterogeneity in the pre-treatment dynamic.

A.5.3 AI Exposure and Interest-Rate Sensitivity

Section 3.5 flags Brazil's aggressive Selic tightening cycle (2021–2022) as a possible confound for the aggregate-wage results. The confound is plausible only if interest-rate-sensitive sectors and AI-exposed occupations overlap. Brynjolfsson et al. (2026) document that they do not in the United States, where the Pearson correlation between AI exposure and interest-rate sensitivity is negative.

We perform an analogous check for Brazil. We do not have a Brazilian Zens-style occupation-level interest-rate-exposure measure, so we use a coarser proxy: occupations in CBO 2002 major groups 7 (industrial and production manual workers, including construction trades) and 8 (industrial machine operators, including auto and durable-goods manufacturing) are classified as high interest-rate sensitivity. Brazilian Central Bank monetary-policy documents consistently identify construction and durables as the Selic-sensitive sectors, with the policy rate transmitted through residential financing and vehicle credit. The proxy is binary and sectoral rather than continuous and credit-weighted; the question is whether the conclusion about the sign of the correlation is robust to this coarseness.

Across CBO 2002 occupations, the Pearson correlation between AI exposure and the high-IR indicator is (p=). Mean AI exposure is for occupations classified as high-IR-sensitive and for the remainder, a difference of . The Brazilian pattern matches the US pattern Brynjolfsson et al. (2026) report: AI exposure concentrates in administrative, sales, and customer-support occupations (CBO 4 and 5), while interest-rate sensitivity concentrates in construction and durable-goods manufacturing (CBO 7 and 8). The Selic cycle should therefore have hit *low*-AI-exposure occupations harder than high-AI-exposure ones, attenuating rather than amplifying the wage gap our DiD measures. The 2022Q1–Q2 date placebos for aggregate wages remain a real concern, but a Selic mechanism does not naturally explain them; we discuss alternative explanations (election-year wage dynamics, sectoral recovery from COVID-19) in Section 2.

A.5.4 Digital Occupation Heterogeneity

AI effects may depend on an occupation's pre-existing digital intensity. Occupations where workers already interact extensively with computers could experience stronger AI effects (because AI tools integrate more readily into existing digital workflows) or weaker effects (because digitally experienced workers adapt more effectively). I construct a digital intensity measure from O*NET, averaging "Interacting With Computers" (Work Activities) and "Computers and Electronics" (Knowledge) importance scores, mapped to CBO using the same crosswalk as the AI exposure measure (Section A.2.1). The correlation between digital intensity and AI exposure is moderate ($r = 0.61$), providing sufficient independent variation for the triple-difference specification. Of 438 mapped occupations, 224 are classified as high-digital (above-median intensity) and 214 as low-digital.[6]

[6] The digital-intensity mapping universe (438 occupations) exceeds the 434 COD codes that appear in the PNAD estimation sample because the crosswalk is constructed from the full ISCO-08/COD classification, whereas the estimation sample is restricted to COD codes with strictly positive worker counts over 2017Q1–2024Q4 after excluding military occupations (COD major group 0). Four crosswalk-mapped codes have no estimation-sample observations and therefore do not appear in the PNAD DiD, but are retained in the digital-intensity counts reported here because the digital-intensity measure is defined at the occupation-classification level. Restricting the digital-intensity split to the 434 occupations in the estimation sample does not materially change the results reported in Table A.21.

Table A.21 presents two approaches. Panel A estimates separate DiDs for high- and low-digital occupations: the treatment effect (High AI Exposure × Post) is estimated within each digital intensity group. Panel B implements a triple difference, where β_1 (High AI × Post) captures the AI effect in non-digital occupations, β_2 (High Digital × Post) absorbs any digital-specific post-treatment trends, and β_3 (High AI × High Digital × Post) tests whether AI effects concentrate in digitally intensive occupations.

Table A.21: Digital Occupation Heterogeneity

	Log Wage (1)	Weekly Hours (2)	Formal (3)
Panel A. Split-Sample DiD			
High Digital Occupations	0.0125	0.1263	-0.0128^{***}
	(0.0096)	(0.1186)	(0.0046)
Pre-trends p-value	[0.929]	[0.253]	[0.376]
Observations	1,702,365	1,748,888	1,432,983
Low Digital Occupations	0.0004	0.1260	-0.0057
	(0.0192)	(0.1898)	(0.0053)
Pre-trends p-value	[0.327]	[0.400]	[0.321]
Observations	3,964,604	4,296,680	2,453,059
Panel B. Triple Difference			
High AI × Post	0.0009	0.1294	-0.0056
	(0.0193)	(0.1882)	(0.0052)
High Digital × Post	-0.0380^{***}	0.1159	0.0015
	(0.0104)	(0.1321)	(0.0056)
High AI × High Digital × Post	0.0119	-0.0015	-0.0071
	(0.0216)	(0.2223)	(0.0069)
Pre-trends p-value	[0.079]	[0.358]	[0.317]
Observations	232,558	233,968	213,359

Notes: Panel A estimates separate DiDs for high- and low-digital occupations. Digital intensity is the average of O*NET "Interacting With Computers" (Work Activities) and "Computers and Electronics" (Knowledge) importance scores, mapped to CBO via the same crosswalk as the AI exposure measure; occupations above the median are classified as high-digital. High AI exposure = above median of Eloundou et al. exposure score. Panel B implements a triple difference: the triple interaction tests whether AI effects concentrate in digitally-intensive occupations. All specifications include occupation, period, and state fixed effects, with standard errors clustered at occupation level in parentheses. Sample: PNAD-C 2017Q1–2024Q4; both panels are estimated on the same underlying sample. Observation counts differ between panels because Panel A reports individual-quarter observations (the raw PNAD-C cell count), whereas Panel B reports occupation × period × state cells (the estimating panel after collapsing), not a narrower sample. ***p<0.01, **p<0.05, *p<0.10.

Aggregate Results In the split-sample analysis (Panel A), high-digital occupations show a statistically significant decline in formal employment of 1.3 percentage points ($p = 0.006$, pre-trends $p = 0.376$), while low-digital occupations show no significant effect on any outcome. Neither group exhibits significant wage or hours effects. In the triple difference (Panel B), the triple interaction is statistically insignificant for all three outcomes (wages $p = 0.581$, hours $p = 0.995$, formality $p = 0.306$), indicating no evidence that aggregate AI effects concentrate differentially in digitally intensive occupations. The coefficient on High Digital $\times$ Post is significant for wages (-0.0380, $p = 0.001$), reflecting a general post-treatment wage decline in digital occupations unrelated to AI exposure.

Young Workers Restricting the sample to workers aged 16–21 inverts the aggregate digital-intensity contrast. In the split-sample, young workers in *low*-digital occupations exhibit a 5.0% wage decline ($p = 0.010$, pre-trends $p = 0.234$) and a 1.5 percentage point formality decline ($p = 0.022$, pre-trends $p = 0.374$). Young workers in high-digital occupations show no significant wage effect ($p = 0.939$, pre-trends $p = 0.657$). The young worker triple-interaction coefficient is +5.2 log points ($p = 0.076$, pre-trends $p = 0.147$; untabulated, estimated on the same panel as Table A.21 Panel B): the AI wage penalty documented in Section A.5.1 for young workers is offset in high-digital occupations, leaving the penalty concentrated among young workers in occupations with low pre-existing digital intensity.

This pattern admits several interpretations. Young workers in non-digital occupations (retail salespeople, cashiers, general laborers) may lack the capacity to leverage AI tools productively, leaving them exposed to indirect competitive effects without offsetting productivity gains. Alternatively, the occupations driving the young worker wage penalty (entry-level service and clerical roles) tend to have low digital intensity by construction, so this finding may reflect the occupational composition of the affected group rather than a distinct mechanism. The concentration of wage penalties in low-digital occupations is nonetheless informative: it suggests that the early labor market effects of generative AI are not confined to workers who interact directly with computers, consistent with broader task reallocation affecting occupations adjacent to AI-automated functions.