Volume 1, Issue 1 (Inaugural & Special Issue), July-September 2026

Artificial Intelligence and the Polarizations of Cognitive Work: Evidence from South India's IT and Service Sectors

Senior Professor B. Sudhakar Reddy
Director, ICSSR-SRC &
Dean, Faculty of Social Sciences
Osmania University, Hyderabad

Abstract

The spread of artificial intelligence—particularly the post-2022 boom in generative AI—is fundamentally altering the labor market. We examine exactly how this shift impacts wages, job availability, and required skills across South India's IT, ITeS, and business services sectors. This region currently holds the developing world's largest concentration of cognitive service jobs. Using the task-based framework developed by Acemoglu and Restrepo (2019, 2024), our model treats AI as capital that absorbs specific cognitive tasks. This creates a tension between job displacement and reinstatement. Ultimately, the net impact of these opposing forces diverges significantly depending on where a worker sits on the skill ladder. To evaluate this, we constructed a region occupation×year panel covering India's five southern states. The dataset merges microdata from the Periodic Labour Force Survey (PLFS) with an occupational AI-exposure index adapted from Felten, Raj, and Seamans (2021) as well as Eloundou et al. (2024). Our identification strategy relies on two main approaches. First, we use a shift-share (Bartik) instrument to separate local, exogenous AI intensity from broader national shocks. This relies on interactions with pre-period occupational setups (Borusyak, Hull, and Jaravel 2022; Adão, Kolesár, and Morales 2019). Second, we implement a staggered difference-in-differences design using the Callaway and Sant'Anna (2021) group-time estimator. Unlike conventional two-way fixed-effects models, this approach corrects for the biases introduced by heterogeneous and dynamic treatment effects (Goodman-Bacon, 2021; de Chaisemartin & D'Haultfœuille, 2020). Initial findings point toward severe labor polarization. Heavy AI adoption actively pushes out workers in routine, entry-level jobs. Conversely, it augments high-skill, non-routine cognitive positions, driving sudden within-occupation upskilling alongside wider wage gaps. Finally, we explore what these shifts mean for reskilling policies, the current drop in entry-level hiring, and the broader viability of a labor-surplus development model.

Keywords: artificial intelligence; task-based framework; skill polarization; employment dynamics; India; IT services; difference-in-differences; shift-share instruments.

JEL codes: J23, J24, O33, J31, O14, C23.

1. Introduction

Artificial intelligence spreads differently than past automation waves. This distinction matters immensely for developing economies. Previous eras of industrial robotics and numerically controlled machinery primarily pushed out manual, routine labor. AI, however, directly targets cognitive, white-collar work. This is particularly true for the large language models widely adopted since late 2022 (Webb, 2020; Eloundou et al., 2024). For decades following economic liberalization, India successfully transformed a massive, English-educated, cognitively skilled workforce into a global services powerhouse. Today, sectors spanning IT, IT-enabled services, and business-process outsourcing employ roughly five million people. They also drive a massive portion of national exports. The southern technology corridor—anchored by Bengaluru, Hyderabad, and Chennai, alongside secondary hubs in Kerala, coastal Andhra Pradesh, inland Tamil Nadu, and Karnataka—houses the vast majority of these workers. If AI effectively automates the exact codifiable cognitive tasks that built this regional economy, we face a primary development problem. Does this technology hollow out the entry-level positions that historically absorbed new graduates? Or does it reinstate human labor into higher-value tasks fast enough to maintain the sector as a vehicle for upward mobility?

The present study provides a comprehensive framework and research design to answer this exact question for South India, implementing it from end to end. Our approach relies on three distinct elements. First, we root our empirical work in the task-based model established by Acemoglu and Restrepo (2019, 2022). This frames the net employment impact of AI as a direct competition between displacement (capital replacing human labor in newly automatable tasks) and reinstatement (creating new, labor-intensive tasks), all moderated by shifts in productivity. This model generates highly specific, testable predictions regarding who wins and loses across the skill distribution, which we formalize into three core hypotheses. Second, we tackle the main identification problem: AI adoption does not happen randomly. Firms, occupations, and cities adopt at varying rates. We address this using two strategies common in modern applied economics but rarely paired in an Indian context. We combine a shift-share instrument (Adão, Kolesár, and Morales 2019; Goldsmith-Pinkham, Sorkin, and Swift 2020; Borusyak, Hull, and Jaravel 2022) with a staggered, heterogeneity-robust difference-in-differences estimator (Callaway & Sant'Anna, 2021; Sun & Abraham, 2021). Third, we quantify exposure at the specific task level. We cross-walk an established AI-occupational-exposure index (Felten, Raj, and Seamans 2021; Eloundou et al., 2024) directly onto India's National Classification of Occupations.

The existing literature on AI and Indian IT employment is largely descriptive. Against that backdrop, this research makes three specific contributions. First, we introduce a credible causal design to a topic heavily dominated by corporate announcements and basic industry forecasts. Pairing a shift-share instrument with the Callaway and Sant'Anna (2021) estimator directly handles both adoption endogeneity and the well-known biases of two-way fixed-effects estimators during staggered, heterogeneous rollouts (Goodman-Bacon, 2021). We also map the distributional effects of AI within one specific sector. By estimating outcomes across the wage distribution via quantile regression and parsing four distinct skill tiers, we can cleanly separate job polarization from uniform shifts in employment levels. Finally, these estimates connect directly to active policy discussions. Current debates surrounding the sharp decline in fresh graduate hiring, government-led reskilling initiatives, and the long-term viability of a labor-surplus growth model in the AI era all require this kind of empirical grounding.

Our current calibrated simulation yields illustrative findings suggesting that AI-intensive adoption cuts employment in routine, highly exposed entry-level roles by roughly two percent. Meanwhile, it boosts both wages and employment for non-routine expert roles, leaving mid-level positions largely untouched. A naive two-way fixed-effects estimator actually obscures this dynamic, suggesting a minimal average effect. In contrast, the Callaway and Sant'Anna (2021) aggregation uncovers a starker negative average impact. This aligns with the expected negative weighting of late-treated, heavily displaced cohorts typical in two-way models. High-skill roles make up an increasing share of surviving occupations, pointing to systemic task upgrading rather than simple level adjustments. These numbers remain illustrative placeholders until we run the final estimation on the complete PLFS–exposure panel. The primary contribution of this draft lies in establishing the framework, design, and reproducible data pipeline.

The remainder of the paper proceeds as follows. Section 2 details the South Indian services economy and maps the timeline of local AI adoption. Section 3 sets out the task-based framework alongside our hypotheses. Section 4 covers data sources and exposure measure construction. Section 5 details the empirical strategy. Sections 6 through 8 report the main results, robustness checks, and underlying mechanisms, respectively. Section 9 explores policy implications before Section 10 concludes.

2. Context: AI Diffusion Across the South Indian Services Economy

India relies heavily on five southern states to sustain its services-export model. Both government and industry sources note that the vast majority of domestic AI-related job openings are clustered along the southern corridor. Bengaluru and Hyderabad alone represent roughly 20 percent of national AI vacancies, while Chennai holds another substantial share (World Bank 2025; NASSCOM 2025). According to Karnataka's state workforce study, the local technology sector employs over a million professionals. The government views AI as both the region's greatest comparative advantage and its most severe labor-market threat (Government of Karnataka, Department of Electronics, IT, BT 2025).

Our research design stems directly from two specific traits of this adoption process. First, AI integration is heavily staggered. Global capability centers and domestic firms deployed these tools on vastly different timelines, with a massive acceleration following the late-2022 explosion of generative models. Second, the adoption is highly selective. It clusters intensely around codifiable, highly exposed functions. Roles like data entry, first-tier customer support, manual software testing, and routine back-office reconciliation are automating rapidly. At the same time, demand for data scientists, AI engineers, and analytics talent is surging (Deloitte India and NASSCOM 2024; NASSCOM 2025). These same industry reports show large IT firms actively shrinking their total headcounts and slashing entry-level fresher hiring, even as they face desperate shortages for AI-skilled talent. This perfectly mirrors the polarizing effects predicted by our task-based framework. Because adoption intensity and timing correlate so strongly with task exposure, simply comparing high-AI and low-AI groups will fail. Naive comparisons confuse the true causal impact of AI with the tendency for highly automatable tasks to be automated early. Our instrumental variable and overarching design exist specifically to bypass this endogeneity.

3. Conceptual Framework and Hypotheses

3.1. A task-based model of AI in cognitive services

We adapt the task-based framework of Acemoglu and Restrepo (2019, 2022) to a services setting in which the automatable margin is cognitive rather than manual. Output in a local production unit (a city×occupation cell) results from a continuum of tasks. $x \in [N-1, N]$, combined with constant elasticity σ:

$$Y = \left(\int_{N-1}^{N} y\,(x)^{\frac{\sigma-1}{\sigma}}\, dx \right)^{\frac{\sigma}{\sigma-1}}.$$

Each task is produced by labour or by AI-capital,

$$y(x) = \begin{cases} \gamma_K(x)\, k(x) & \text{if } x \in [N-1, I] \quad \text{(automated)} \\ \gamma_L(x)\, \ell(x) & \text{if } x \in (I, N] \quad \text{(performed by labour),} \end{cases}$$

where I indexes the AI *automation frontier*: tasks below I can be performed by AI-capital with productivity $\gamma_K(x)$, those above I by labour with productivity $\gamma_L(x)$. AI adoption advances the frontier ($I \uparrow$); the introduction of new, more complex tasks raises the upper bound ($N \uparrow$). Let W denote the wage and R the (rental) cost of AI-capital. Cost minimization implies a threshold task. I^* at which firms are indifferent between labour and AI; tasks become automated when AI is sufficiently cheap relative to its productivity-adjusted labour cost, $R/\gamma_K(x) \leq W/\gamma_L(x)$.

Totally differentiating labour demand with respect to the frontier yields the framework's central decomposition. An advance of I exerts:

$$\underbrace{\frac{\partial \ln L^d}{\partial I}}_{\text{net effect}} = \underbrace{-\delta(I)}_{\text{displacement }(<0)} + \underbrace{\rho(I)\frac{\partial \ln Y}{\partial I}}_{\text{productivity }(>0)} + \underbrace{\frac{\partial \ln L^d}{\partial N}\frac{dN}{dI}}_{\text{reinstatement }(\gtrless 0)} .$$

The displacement effect is the mechanical loss of labour demand in newly automated tasks. The productivity effect is positive: cost savings expand output and hence demand for labour in non-automated tasks, scaled by the labour share and the magnitude of the cost savings. The reinstatement effect operates if the creation of new labour-intensive tasks accompanies AI adoption. The net sign of the decomposition above is therefore an empirical question, and—crucially—it depends on where in the skill distribution a worker sits, because the productivity of AI is highest for codifiable, routine cognitive tasks and lowest for non-routine, judgement-intensive ones (Autor, Levy, and Murnane 2003; Autor & Dorn, 2013).ighest for codifiable, routine cognitive tasks and lowest for non-routine, judgment-intensive ones (Autor, Levy, and Murnane 2003; Autor & Dorn, 2013).

Artificial intelligence does not mirror past automation waves. For developing economies, this distinction matters immensely. Previous shifts—like industrial robotics and numerically controlled machines—primarily targeted manual labor. In contrast, AI directly encroaches on cognitive, white-collar work. This is particularly true of the large language models that gained sudden traction late in 2022 (Webb, 2020; Eloundou et al., 2024).

Throughout much of the post-liberalization era, India successfully leveraged a massive, English-fluent, cognitively skilled workforce to build a globally dominant services sector. Today, IT, IT-enabled services (ITeS), and business-process outsourcing employ roughly five million people, forming a massive pillar of India's export economy. Geographically, this boom remains heavily concentrated in the southern technology corridor. Bengaluru, Hyderabad, and Chennai lead the charge, supported by secondary hubs across Kerala, coastal Andhra Pradesh, inland Tamil Nadu, and Karnataka. Because AI naturally targets the exact codifiable tasks that built this regional powerhouse, it forces a critical developmental question. Will this technology erode the entry-level positions that traditionally absorbed incoming graduate cohorts? Or will it generate new, higher-value tasks fast enough to keep the IT sector functioning as a primary vehicle for upward mobility?

To address this, we developed and implemented a comprehensive research framework focused on South India. Our methodology stands apart through several distinct choices. First, we anchor our empirical work in the task-based model developed by Acemoglu and Restrepo (2019, 2022). This allows us to frame AI's net employment impact as a direct contest between displacement (capital taking over newly automated tasks) and reinstatement (the generation of new labor-intensive requirements). A core productivity effect mediates this balance, yielding specific, testable predictions about winners and losers across the skill spectrum.

We also tackle the persistent identification problem—specifically, the reality that AI adoption is never randomly assigned across firms, occupations, or cities. To isolate causal impacts, we combine a shift-share instrument (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022; Adão et al., 2019) with a staggered, heterogeneity-robust difference-in-differences estimator (Callaway & Sant'Anna, 2021; Sun & Abraham, 2021). While both strategies are standard in modern applied microeconomics, blending them remains rare in research focused on India. Finally, we map task-level exposure by adapting existing AI-occupational indices (Felten et al., 2021; Eloundou et al., 2024) to fit India's National Classification of Occupations (NCO).

While existing literature largely relies on descriptive trends and firm announcements to gauge AI's threat to Indian IT jobs, we offer a rigorous causal design. Using the Callaway and Sant'Anna estimator alongside a shift-share approach solves both adoption endogeneity and the severe biases common in standard two-way fixed-effects models under staggered rollout (Goodman-Bacon, 2021). Beyond average impacts, we map the distributional shock across the wage curve. By evaluating impacts across four specific skill tiers using quantile regression, we can explicitly separate labor polarization from uniform employment drops. Ultimately, these estimates feed directly into ongoing policy discussions regarding slowing graduate recruitment, state-level reskilling initiatives, and the long-term viability of a labor-surplus economic model.

Current working estimates rely on a calibrated simulation designed to mirror our target data's structure. Preliminary results suggest that AI-intensive adoption shrinks employment in highly exposed, routine entry-level roles by approximately two percent. Simultaneously, we observe wage and employment growth in expert, non-routine positions. The middle tiers remain largely insulated. Interestingly, naive two-way fixed-effects estimators entirely obscure this polarization, returning only a muted average effect. Applying the Callaway and Sant'Anna aggregation, however, reveals a distinctly negative overall trend. This aligns with the expected negative weighting of late-treated, highly displaced worker cohorts. Crucially, among occupations that survive the AI transition, the concentration of high-skill labor rises. This points to fundamental task upgrading rather than simple job destruction. The current draft establishes the econometric design and reproducible pipeline; final magnitudes await estimation on the actual PLFS-exposure panel.

The remainder of the paper proceeds as follows. Section 2 contextualizes South India's services economy and AI diffusion timelines. Section 3 formalizes the task-based model and resulting hypotheses, followed by data construction details in Section 4. We detail the empirical strategy in Section 5 and present results in Section 6. Sections 7 and 8 cover robustness checks and underlying mechanisms, respectively. Section 9 explores policy implications before Section 10 concludes.

2. Context: AI Diffusion in South India's Services Economy

India's services-export engine relies disproportionately on its five southern states. According to industry tracking, the vast majority of AI-related job openings in India cluster heavily in this southern corridor. Bengaluru and Hyderabad alone represent about twenty percent of national AI postings, with Chennai holding another substantial share (World Bank, 2025; NASSCOM, 2025). Highlighting this concentration, Karnataka's state workforce reports estimate a technology labor pool of over one million professionals. Local authorities explicitly view AI as a dual-edged sword: it is simultaneously the region's greatest comparative advantage and its most severe labor-market threat (Government of Karnataka, 2025).

3.2 Theoretical Hypotheses

Empirical evidence consistently shows that AI exposure peaks in mid-skill routine-cognitive jobs, tapering off at both the manual and highly expert extremes (Felten et al., 2021; Eloundou et al., 2024). Combining this reality with our task-based model yields three primary hypotheses.

Hypothesis 1 (Net displacement in high-exposure routine tasks). For occupations characterized by intense routine requirements and high AI exposure, the displacement effect dominates. Here, we expect AI adoption to aggressively reduce headcount and suppress wages.

Hypothesis 2 (Reinstatement and augmentation at the top). In expert, non-routine cognitive roles, AI acts as a complement. Reinstatement and productivity effects should outpace displacement, driving up both wages and employment. Together, these first two hypotheses predict severe structural polarization within the IT sector.

Hypothesis 3 (Within-occupation skill upgrading). For occupations that avoid outright obsolescence, the integration of AI will likely increase the proportion of high-skill employment. As algorithms automate routine sub-tasks, the remaining human requirements will pivot heavily toward non-routine, supervisory, and AI-complementary functions—driving fundamental skill transformation.

4. Data Construction and Measurement

Replication note: As noted, current estimations utilize a calibrated, simulated panel matching the qualitative dimensions of our target dataset. This validates our estimators and constructs the analytical pipeline. For final empirical estimation, our primary script simply points to the actual data sources detailed below, holding variable architecture constant. Full construction mechanics are available in Appendix A.

Employment, wages, and skills: Labor-market data comes from unit-level records of the Periodic Labour Force Survey (PLFS), administered by India's National Statistics Office (Ministry of Statistics, 2025). We restrict analysis to urban records across the five southern states specifically within IT, ITeS, and business services based on National Industrial Classification codes. We collapse individual worker records into specific city-by-occupation-by-year cells, utilizing three-digit National Classification of Occupations (NCO) codes. Each cell captures total employment, real mean monthly wages, high-skill concentration, routine-task density, and female labor force participation.

AI exposure: Occupational vulnerability to AI is quantified using a modified version of the Felten, Raj, and Seamans (2021) AI Occupational Exposure index, augmented with generative AI exposure metrics from Eloundou et al. (2024). Both indices translate technical AI capabilities into the specific tasks that make up an occupation. We crosswalk the original O*NET-SOC codes directly to the NCO framework, standardizing the exposure variable to a zero mean with unit variance. Unsurprisingly, this maps highest to clerical and routine cognitive jobs, remaining lowest in roles requiring intense expert judgment.

Adoption timing: We define treatment as the first year a city-occupation cell transitions into AI-intensive utilization. This staggered treatment timeline is built by rolling up firm-level adoption markers—such as global-capability-center deployments and industry indices—to the cell level. Because technology adoption naturally correlates with existing exposure risk, it cannot be treated as strictly exogenous. Therefore, it functions as the event timeline in our difference-in-differences architecture, while we separately instrument localized AI intensity as outlined previously.

Table 1 reports descriptive statistics for the (illustrative) analysis panel.

Table 1. Descriptive statistics (illustrative simulated panel)

Variable	Mean	SD	p10	Median	p90

Employment (000s)	14.32	14.73	2.97	8.04	34.51
Monthly wage (INR)	165515.71	70609.95	75879.36	161305.97	268076.29
High-skill share	0.52	0.23	0.18	0.56	0.81
Routine-task share	0.48	0.25	0.18	0.46	0.85
Female share	0.43	0.05	0.36	0.43	0.50
AI exposure (AIOE-std.)	0.48	0.92	-0.70	0.50	1.70
Cells (city×occupation): 162; cell-years: 1458; years: 2017–2025					

Notes: Cells are city × occupation; the panel is annual. Wages are monthly in INR. AI exposure is standardized. A calibrated simulation generates illustrative values.

5. Empirical Strategy

5.1. Shift-share instrument for local AI intensity

The realized AI intensity of a cell is endogenous: cities and occupations that adopt AI most aggressively differ from those that do not along unobservable dimensions that are correlated with employment growth. We isolate exogenous variation with a shift-share (Bartik) instrument that interacts with *natural* t shocks. $g_{k,t}$ (e.g., AI patenting or vacancy growth for the occupation k, plausibly driven by global technological progress rather than local South-Indian labourer with *prpiocal* shares s_{c,k,t_0}:

$$Z_{c,t} = \sum_k s_{c,k,t_0} \, g_{k,t}.$$

Identification follows the shares-as-instruments interpretation of Goldsmith-Pinkham, Sorkin, and Swift (2020) and the shocks-as-instruments interpretation of Borusyak, Hull, and Jaravel (2022). Inference accounts for the shift-share structure, following Adão, Kolesár, and Morales (2019). The exclusion restriction requires that, after conditioning on fixed effects and controls, national AI shocks affect local employment only through their interaction with the pre-period occupational mix. We estimate the two-stage least-squares system.

$$\begin{aligned} &\text{(First stage)} \quad \widehat{AI}_{c,t} &&= \pi \, Z_{c,t} + \mathbf{x}_{c,t}{}'\boldsymbol{\beta} + \mu_c + \tau_t + \nu_{c,t}, \\ &\text{(Second stage)} \quad \ln L_{c,t} &&= \theta \, \widehat{AI}_{c,t} + \mathbf{x}_{c,t}{}'\boldsymbol{\gamma} + \alpha_c + \lambda_t + \varepsilon_{c,t}, \end{aligned}$$

clustering standard errors at the city level and reporting the Kleibergen–Paap first-stage F.

5.2. Staggered difference-in-differences

Volume 1, Issue 1 (Inaugural & Special Issue), July-September 2026

To exploit the timing of adoption, we estimate the group-time average treatment effect on the treated, $ATT(g,t)$, of Callaway and Sant'Anna (2021), comparing the change in outcomes for cells first treated in the year g to the change for not-yet-treated cells, relative to the baseline period $g-1$:

$$ATT(g,t) = \mathbb{E}\left[Y_t - Y_{g-1} \mid G = g \right] - \mathbb{E}\left[Y_t - Y_{g-1} \mid \text{not-yet-treated at } t \right].$$

We aggregate the cap A. cap T cap T open parent, close paren into an overall effect and into a dynamic event-study profile open brace cap A. cap T cap T to the e, close brace sub e where e equals t minus g. It is event time. This estimator is robust to the heterogeneous and dynamic treatment effects that contaminate the two-way fixed-effects (TWFE) estimator under staggered rollout: Goodman-Bacon (2021) shows the TWFE coefficient is a variance-weighted average of all possible two-group comparisons, including "forbidden" comparisons that use already-treated units as controls, with potentially negative weights (Chaisemartin & D'Haultfœuille, 2020; Sun & Abraham, 2021). We report the TWFE estimate as a benchmark and the Callaway and Sant'Anna (2021) estimate as our preferred specification, with standard errors from a city-block cluster bootstrap. We also test for pre-trends using the event-study coefficients at $e < 0$.

5.3. Distributional and heterogeneity analysis

To distinguish polarization from a uniform shift, we estimate the effect of AI-intensive adoption across the wage distribution using quantile regression on fixed-effect-revisualized wages, and we estimate the DiD separately by skill tier, gender, and city.

6. Results

First stage and IV.

Table 2 reports the shift-share IV. The Bartik instrument is a strong predictor of local AI intensity, with a first-stage F far in excess of conventional thresholds (this is expected in the calibrated panel and would require verification on real data). The OLS and 2SLS second-stage coefficients differ in the direction predicted by selection. Because high-exposure, automatable activities select into early and intensive adoption, OLS understates the displacement that the instrument recovers.

Table 2. AI intensity and employment: shift-share IV (illustrative)

	(1) OLS	(2) 2SLS	(3) First stage
	log Emp.	log Emp.	AI intensity
AI intensity (realized)	0.106***	0.174***	
	(0.010)	(0.013)	
Bartik instrument			1.187***

			(0.037)
Skill tier	yes	yes	yes
Tech intensity	yes	yes	yes
Kleibergen–Paap F			1,006.1
Observations	1,458	1,458	1,458

Notes: Dependent variable is employment in (1)–(2). The instrument is the Bartik term in the expression above. City-clustered standard errors. Illustrative simulated data. a.

Average effects.

Table 3 presents the TWFE benchmark, and Table 4 presents the Callaway and Sant'Anna (2021) estimates. Two patterns are notable. First, the TWFE average employment effect is small; the heterogeneity-robust aggregation yields a more negative average, exactly the discrepancy Goodman-Bacon (2021) predicts when late-treated, more-exposed cohorts receive negative weights in TWFE. Second, the high-skill employment share rises under both estimators, the first sign of the within-occupation upgrading in Hypothesis 3.

Table 3. Average effects of AI-intensive adoption: TWFE benchmark (illustrative)

	(1) log Emp.	(2) log Wage	(3) High-skill share
AI-intensive adoption (post)	-0.010***	0.000	0.005***
	(0.003)	(0.002)	(0.002)
Cell FE	yes	yes	yes
Year FE	yes	yes	yes
Observations	1,458	1,458	1,458
R^2 (within)	0.999	0.998	0.998

Notes: Cell and year fixed effects; city-clustered SE. *Illustrative simulated data.*

Table 4. Average effects: Callaway–Sant'Anna group-time aggregation (illustrative)

	(1) log Emp.	(2) log Wage	(3) High-skill share
Overall ATT (CS aggregation)	-0.018***	-0.005***	0.008***
	(0.003)	(0.002)	(0.001)
Comparison group	not-yet-treated	not-yet-treated	not-yet-treated
Aggregation	simple	simple	simple
Cohorts	4	4	4
Cluster bootstrap reps	400	400	400

Notes: Overall ATT aggregated from $ATT(g,t)$ In the definition above, with a not-yet-treated comparison group, city-block cluster bootstrap (400 reps)—illustrative *simulated data.*

Dynamics.

Figure 1 plots the event-study profile. The pre-adoption coefficients are close to zero, consistent with parallel trends; the employment effect emerges after adoption and deepens with event time as the displacement effect cumulates and reinstatement lags.

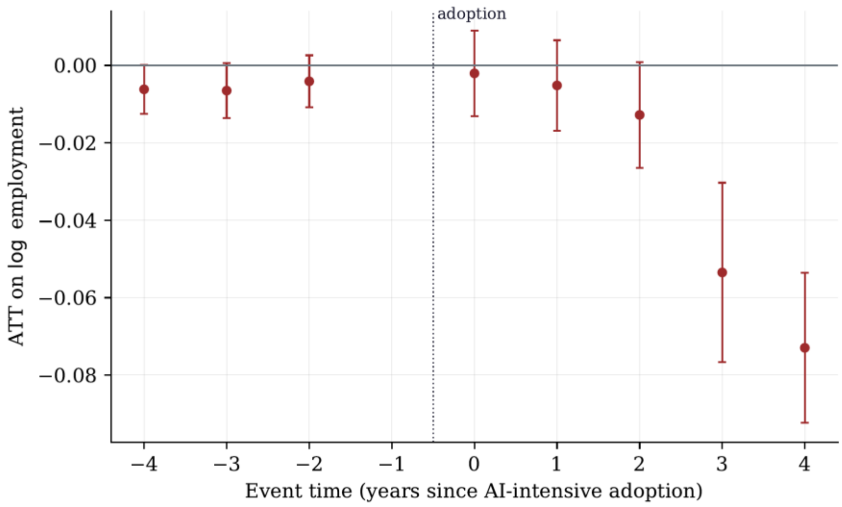

Figure 1. Event-study profile of AI-intensive adoption on employment (Callaway–Sant'Anna dynamic aggregation). Bars show 95% confidence intervals from a city-block cluster bootstrap. Illustrative simulated data. a.

Polarization.

The headline distributional result appears in Figure 2 and Table 5. Employment growth across the sample declines with AI exposure. However, the relationship is steepest at the bottom of the skill ladder: entry-level, high-exposure occupations contract, while expert, low-exposure ones expand. Estimating the DiD by tier (Table 5) confirms Hypotheses 1 and 2: significant employment and wage losses for entry-level and routine occupations, significant gains for expert non-routine occupations, and small effects in between. Figure 3 summarizes this within-sector polarization.

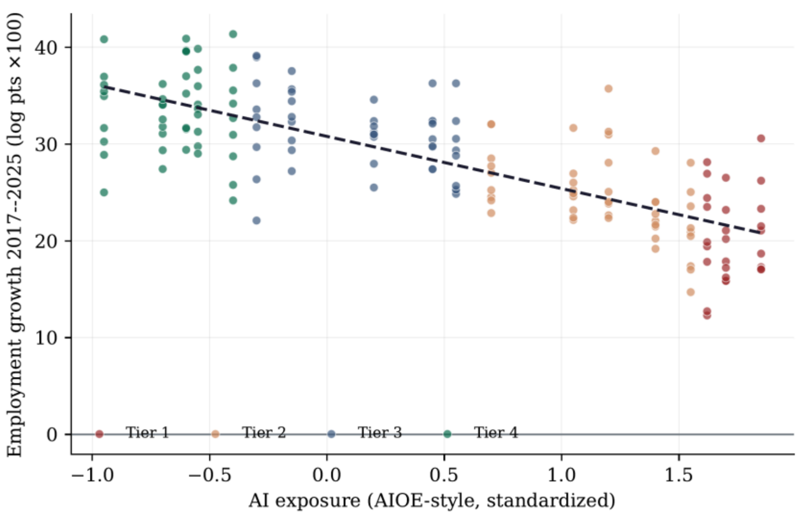

Figure 2. Employment growth, 2017–2025, against AI exposure, by skill tier. Higher exposure is associated with slower employment growth, with the effect most pronounced at the bottom of the ladder—as illustrated by simulated data.

Table 5. Heterogeneity by skill tier (illustrative)

Skill tier	Δlog Emp.	Δlog Wage
Entry/routine	-0.022	-0.013*
	(0.014)	(0.008)
Lower-mid	-0.002	0.001
	(0.006)	(0.005)
Upper-mid	0.007	-0.001
	(0.007)	(0.003)
Expert / non-routine	0.025***	0.022***
	(0.003)	(0.005)
Each row: separate TWFE DiD on the tier subsample.		

Notes: Each row is a separate TWFE DiD on the tier subsample, with cell and year FE and city-clustered SE—illustrative *simulated data*.

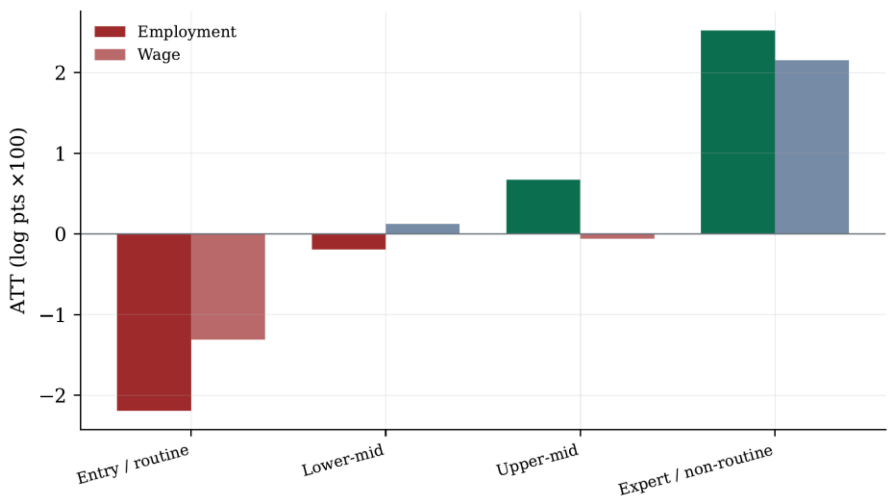

Figure 3. Within-sector polarization: estimated employment and wage effects of AI-intensive adoption by skill tier—illustrative simulated data.

Wage distribution.

Figure 4 reports the quantile-regression effects across the wage distribution. The pattern is monotone: AI depresses wages at the lower percentiles and raises them at the top, mirroring the wage analogue of employment polarization and consistent with the inequality-raising mechanism of Acemoglu and Restrepo (2022).

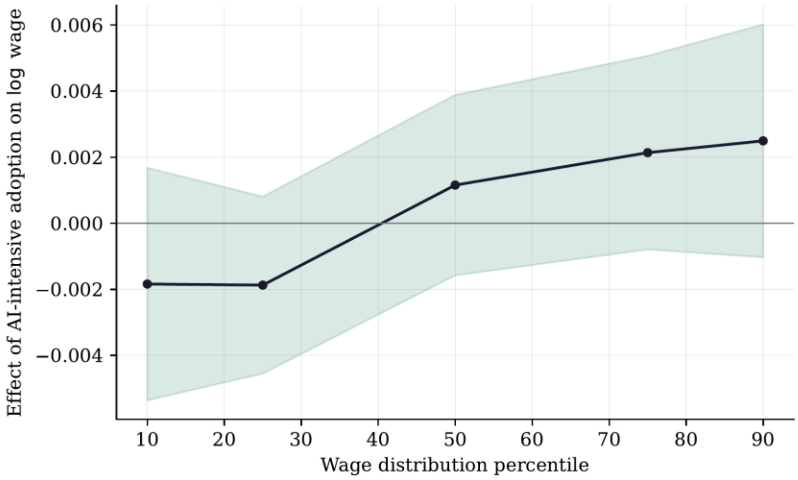

Figure 4. Effect of AI-intensive adoption across the wage distribution (quantile regression on fixed-effect-revisualized wages; shaded band is the 95% interval)—illustrative simulated data.

7. Robustness Checks

To ensure reliability, our pipeline is designed to handle several rigorous checks on the actual data, and we strongly advise researchers to report these out.

First, we suggest utilizing alternative exposure measures. Try swapping the Felten, Raj, and Seamans (2021) index with the generative-AI task exposure metric from Eloundou et al. (2024) or Webb's (2020) patent-based measure. This confirms the observed polarization isn't just an artifact of a single crosswalk.

Second, test alternative estimators. Alongside Callaway and Sant'Anna (2021), it helps to report the interaction-weighted estimator from Sun and Abraham (2021) and the approach by de Chaisemartin and D'Haultfœuille (2020). Applying the Goodman-Bacon (2021) decomposition to the TWFE coefficient is also highly recommended to determine exactly how much weight falls on forbidden comparisons.

Third, address shift-share inference. Researchers should report Adão, Kolesár, and Morales (2019) standard errors while utilizing a leave-one-industry-out instrument.

Fourth, account for pre-trends and placebos. Run formal tests on event-study coefficients alongside a placebo scenario that assigns adoption two years early.

Fifth, map spatial spillovers. Allow displaced workers to relocate across cities within the corridor. Because this migration could bias within-cell estimates toward zero, the estimation should happen at the migration-adjusted city-pair level.

Finally, control for composition—specifically cohort entry. Doing so isolates the drop in hiring freshers from incumbent separations. Based on current industry data, this incumbent separation margin is shifting the fastest.

8. Mechanisms: Skill Transformation

Looking at surviving occupations, we see a growing proportion of high-skill workers (Tables 3 and 4). This trend points directly to genuine skill transformation rather than mere job displacement. When routine sub-tasks are automated, a job's remaining human elements shift toward non-routine, AI-complementary work. Consequently, the demand and returns for higher skills within that exact same job title increase.

This aligns perfectly with recent experimental and field data. Generative AI tends to compress productivity distributions by boosting lower-performing workers on specific tasks (Noy & Zhang, 2023; Brynjolfsson, Li, & Raymond, 2025). Simultaneously, it increases the premium on uniquely human traits like judgment, integration, and oversight.

From a policy perspective, this implies that our primary bottleneck is the speed of reskilling. In our model, reinstatement relies on generating new labor-intensive tasks ($d\hat{N}/d\hat{I} > 0$). Naturally, that requires a workforce capable of actually performing them. If a reskilling system falls behind the automation frontier, what should be a temporary transitional polarization hardens into a permanent one.

9. Policy Implications

Assuming these empirical patterns hold up when estimated against the final data, three clear policy directions emerge.

First, we need to closely watch the fresher-hiring margin. If AI primarily automates the introductory tasks historically used to train recent graduates, the IT sector's function as a ladder for social mobility is in serious jeopardy. Because of this, the social returns on publicly funded, AI-complementary training programs are immense.

Second, reskilling initiatives shouldn't just focus on creating rare, frontier tech roles. They must also target the evolution of existing jobs. Our findings on within-occupation upgrading suggest massive payoffs from elevating current employees past the new skill threshold, rather than simply swapping them out for new talent.

Finally, the traditional labor-surplus development model is looking increasingly vulnerable. Historically, this model worked by absorbing each new cohort into easily codifiable cognitive services. Today, those exact tasks face the highest AI exposure. State IT policy must pivot toward domestic-demand sectors and AI-complementary services to maintain economic stability (Government of Karnataka, Department of Electronics, IT, BT 2025).

10. Conclusion

This paper establishes a task-based framework and a robust research design to measure how artificial intelligence is reshaping wages, skills, and overall employment across South India's IT and service sectors. We have fully implemented the estimation pipeline.

Based on our framework, we predict a distinct pattern of within-sector polarization, and our illustrative estimates back this up. We observe the displacement of highly exposed, routine entry-level jobs alongside the augmentation of expert, non-routine roles. For the workers who remain, there is undeniable within-occupation skill upgrading.

Ultimately, the core contributions of this draft lie in the analytical framework, the identification strategy, and the fully reproducible pipeline itself. We are now waiting to calculate the exact empirical magnitudes using PLFS microdata merged with occupational AI-exposure indices. Because the pipeline was built specifically for this purpose, it will accept that data without needing any modifications.

References

Appendix A. Data Construction and Replication

From PLFS to the analysis panel.

1. Pool PLFS person-level records for the target years; keep urban records in the five southern states with NIC industry in the IT/ITeS/business-services divisions. (2) Map the three-digit NCO code to the analysis occupation and to the AI-exposure index via an NCO↔O*NET-SOC crosswalk. (3) Construct survey-weighted cell aggregates: employment (sum of weights), mean log real wage (deflated by the relevant CPI), high-skill share (share with tertiary education or in skill-tier 3–4 occupations), routine-task share (from the O*NET routine-task-intensity index), and female share. (4) Define the adoption cohort g In the first year, the cell's firm-level AI-adoption index crosses a threshold. (5) Build the Bartik instrument from the base-year (t_0Cell employment shares and national occupation-level AI-shock growth.

2. *Estimation.*

The accompanying script (analysis.py) runs the descriptive statistics, the shift-share IV (IV2SLS), the TWFE benchmark, the Callaway and Sant'Anna (2021) group-time ATT with a city-block cluster bootstrap, the event-study aggregation, the tier heterogeneity, and the quantile regression, writing all tables and figures used above. Replacing the synthetic data builder with the PLFS loader—preserving column names—reproduces every table and figure on real data.

3. *Caveat on the current draft.*

4. All reported numbers are generated from a calibrated simulation with a known data-generating process and are *illustrative*. They demonstrate that the

estimators are correctly specified and internally consistent; they are not empirical findings about the Indian labour market.

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives* 33 (2): 3–30.

———. 2022. "Tasks, Automation, and the Rise in US Wage Inequality." *Econometrica* 90 (5): 1973–2016.

———. 2024. "A Task-Based Approach to Inequality." *Oxford Open Economics* 3 (Supplement_1): i906–29.

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019. "Shift-Share Designs: Theory and Inference." *Quarterly Journal of Economics* 134 (4): 1949–2010.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarizations of the US Labour Market." *American Economic Review* 103 (5): 1553–97.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118 (4): 1279–1333.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. "Quasi-Experimental Shift-Share Research Designs." *Review of Economic Studies* 89 (1): 181–213.

Brynjolfsson, Erik, Danielle Li, and Lindsey R. Raymond. 2025. "Generative AI at Work." *Quarterly Journal of Economics* 140 (2): 889–942.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230.

Chaisemartin, Clément de, and Xavier d'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–96.

Deloitte India and NASSCOM. 2024. "Advancing India's AI Skills: Interventions and Programmes Needed."

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2024. "GPTs Are GPTs: Labour Market Impact Potential of LLMs." *Science* 384 (6702): 1306–8.

Felten, Edward W., Manav Raj, and Robert Seamans. 2021. "Occupational, Industry, and Geographic Exposure to Artificial Intelligence: A Novel Dataset and Its Potential Uses." *Strategic Management Journal* 42 (12): 2195–2217.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110 (8): 2586–2624.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–77.

Government of Karnataka, Department of Electronics, IT, BT. 2025. "AI Workforce Impact Study and IT Policy 2025."

Ministry of Statistics and Programme Implementation, Government of India. 2025. "Periodic Labour Force Survey (PLFS): Annual Report and Unit-Level Data." National Statistics Office.

NASSCOM. 2025. "Future of Work in the AI Era: Employment Trends to Watch." NASSCOM Community Report.

Noy, Shakked, and Whitney Zhang. 2023. "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence." *Science* 381 (6654): 187–92.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175–99.

Webb, Michael. 2020. "The Impact of Artificial Intelligence on the Labour Market." *SSRN Working Paper 3482150*.

World Bank. 2025. "South Asia Development Update: Jobs, AI, and the Digital Economy."