

Asymmetric Disruption Across Demographics and Regional Economies: How Artificial Intelligence Reshapes Early-Career Hiring Pipelines, Junior Mentorship, and Gender-Concentrated Administrative and Service Roles

Manasvi Pramod Kulkarni

MCA, Zibacar, Savitribai Phule Pune University, Pune, India
manasvikulkarni14@gmail.com

Abstract

Debate on artificial intelligence (AI) and employment has centred on the displacement of mid-career workers, leaving the entry point of careers comparatively unexamined. Junior roles rely on codifiable tasks such as drafting, data handling, scheduling and first-level support that generative AI can now perform, so firms may hire fewer juniors, raise experience requirements and reduce the mentoring through which novices become experienced workers. This paper develops the research foundation and a complete empirical design for studying how AI adoption affects early-career hiring pipelines, junior mentorship structures and groups concentrated in exposed roles, notably women in administrative and service occupations, across contrasting regional economies. It integrates task-based, human capital, mentoring, intersectionality and regional theory; identifies six specific gaps; articulates theoretical, empirical, methodological and policy contributions; and states eight falsifiable hypotheses. A sequential explanatory mixed-methods design combines a primary survey of early-career workers and employers (instrument provided in full), secondary labour-force and job-posting data analysed with difference-in-differences and shift-share methods, and qualitative interviews. A Regional Vulnerability and Resilience Index and a results-reporting framework are specified. Empirical estimates will be reported after fieldwork.

Index Terms—artificial intelligence, generative AI, entry-level employment, early-career hiring, mentorship, gender inequality, administrative and service occupations, marginalized demographics, regional economies, labour market disruption, mixed methods, survey

I. INTRODUCTION

Every major technological transition has altered both the content of work and the routes through which people enter it. Mechanization shifted labour out of agriculture, computerization reduced demand for routine clerical and production tasks, and the diffusion of the internet created new forms of service employment. The present wave of generative artificial intelligence (AI) differs from these episodes in one important respect. Earlier automation mainly displaced manual and routine physical tasks, whereas large language models and related systems can draft text, summarize documents, translate, write code, schedule, classify records and answer routine customer queries [3], [5], [49]. These are precisely the tasks that organizations have historically assigned to new graduates and junior staff as a way of letting them learn the business while contributing modest value.

Public and scholarly attention has so far gravitated toward the displacement of established, mid-career workers, whose job losses are visible in layoff announcements, occupational exposure rankings and wage data [1], [8], [11]. A quieter question has received far less systematic study: what happens at the *entry point* of careers? If firms use AI to absorb the codifiable tasks that once formed the bottom rung of the occupational ladder, the consequences extend beyond a smaller number of vacancies. They reach the pipeline through which organizations produce their future senior workers, the informal apprenticeship through which tacit knowledge is transmitted, and the life chances of those who depend most on open, formal entry channels because they lack family networks, savings or alternative routes into stable work [6], [16], [33].

The risk is not evenly distributed. In many economies women are heavily concentrated in administrative, clerical, customer-service and other support occupations that score high on measures of exposure to generative AI [10], [11], [12]. First-

generation graduates, socially disadvantaged groups, migrants and older job seekers often face weaker referral networks and thinner access to reskilling. Regions also differ: places that specialize in routine back-office or outsourced services may face concentrated shocks with few alternative employers, while diversified metropolitan labour markets may reallocate workers more easily [4], [25]. Disruption is therefore best understood as *asymmetric* along three axes at once: career stage, demographic position and regional economy.

This paper makes the early-career pipeline, its mentorship structures and its demographic and regional heterogeneity the central objects of analysis. It develops a research foundation that connects task-based labour economics with human capital, mentoring and intersectionality theory; it states a precise research gap and a set of testable hypotheses; and it specifies a complete mixed-methods design built around a primary survey that the authors will administer to early-career workers and employers, supplemented by secondary labour-force and job-posting data and by qualitative interviews. The survey instrument is reproduced in full in Appendix A so that the study can be replicated, audited and extended.

The remainder of the paper is organized as follows. Section II lays out the research foundation. Section III reviews empirical evidence. Section IV states the research gap and problem. Section V articulates the research contribution. Section VI presents the conceptual framework and hypotheses. Sections VII to IX describe the methodology, the survey and the analysis plan. Section X provides the results-reporting framework to be completed after fieldwork. Sections XI to XIII discuss interpretation, policy implications, ethics and limitations, and Section XIV concludes.

II. RESEARCH FOUNDATION

The study rests on six strands of theory. Each is summarized below together with the specific role it plays in the research design.

A. Task-Based Models of Technology and Labour

The task-based framework treats an occupation as a bundle of tasks rather than a monolithic skill. Technology can *displace* labour from tasks it can perform and can *reinstate* labour by creating new tasks in which people hold a comparative advantage; the net employment effect depends on the balance of the two forces [1], [9]. Autor, Levy and Murnane showed that computerization substituted for routine cognitive and manual tasks while complementing non-routine analytical and interactive ones [2]. Generative AI complicates the older routine-versus-non-routine boundary because it performs many non-routine cognitive tasks, such as composing a first draft, in ways that earlier software could not [3]. For this study the framework implies that exposure should be measured at the task level and that its consequences should be traced separately for different positions in the career ladder, because junior and senior workers within the same occupation perform different task mixes.

B. Routine-Biased and Skill-Biased Technological Change

Skill-biased technological change predicts rising relative demand for high-skill workers, and routine-biased technological change predicts the hollowing out of middle-skill routine occupations and job polarization [14], [28]. Both traditions are useful but were developed for earlier technologies. They predict outcomes by occupation and education level, and they say little about whether the *same* occupation offers fewer entry positions even if its senior ranks are unaffected. The present study extends them by asking whether AI exposure is concentrated at the entry level of exposed occupations.

C. Human Capital and Apprenticeship

Human capital theory explains why firms employ and train inexperienced workers: early productivity is low, but learning on the job raises future productivity, and the firm or the worker finances this investment in the expectation of later returns [7]. Junior roles are therefore partly paid training. The relevant mechanism for AI is that routine tasks double as practice. When a model completes the practice tasks more quickly than a novice, the firm's short-run incentive to hire novices declines, and the novice loses the repeated exposure through which judgment is formed. A further implication concerns knowledge type. Formal education provides largely codified knowledge that models can reproduce, whereas senior workers hold tacit, context-specific knowledge that models complement. This asymmetry offers a theoretical reason why exposure may bite at the entry level more than at senior levels [6], [48].

D. Mentorship, Social Capital and Weak Ties

Mentoring is a distinct channel of skill formation. Kram distinguished career functions, such as sponsorship, coaching and exposure, from psychosocial functions, such as role modelling and counselling [15]. A meta-analysis found that mentored protégés report more favourable career outcomes, including pay, promotion and satisfaction, than non-mentored peers [16]. Network theory adds that access to information and referrals depends on the structure of ties [33], and recent evidence shows rising returns to social skills in the labour market [47]. The research implication is twofold. First, mentoring is an outcome that AI adoption may erode if it automates practice tasks, converts supervisors from coaches into reviewers of machine output, or shrinks junior cohorts so that structured programmes no longer pay for themselves. Second, groups with weaker informal networks depend more on formal mentoring and open recruitment, so any erosion has distributional consequences.

E. Occupational Segregation and Intersectionality

Occupational segregation by gender means that technology shocks to feminized occupations translate directly into gendered labour-market outcomes. The concept of intersectionality, introduced by Crenshaw, holds that disadvantages arising from different social categories combine in ways that are not captured by examining each category alone [17]. In the present context a young woman from a socially disadvantaged household in a smaller city may face a compounding of exposure, network and infrastructure

constraints. Methodologically, intersectionality motivates interaction terms and subgroup estimates in place of single-factor comparisons.

F. Regional Economies and Spatial Heterogeneity

Regional labour markets differ in industrial composition, institutions, density and digital infrastructure. A region's expected exposure can be approximated by combining its industry employment shares with industry-level exposure scores, which is the logic of shift-share analysis [25]. Exposure alone is not vulnerability, however. Adaptive capacity, which depends on educational institutions, training providers, diversification and connectivity, determines whether displaced entrants can move to alternative work. For developing economies that export services, decisions taken by firms in advanced economies can transmit shocks through global value chains, adding an external channel [35]. The study therefore treats region as both a source of exposure and a moderator of impact.

G. Integrating the Theoretical Strands

Table I shows how the strands combine to generate constructs, hypotheses and methods, so that every element of the empirical design traces to a theoretical claim.

TABLE I
MAPPING OF THEORY TO CONSTRUCTS, HYPOTHESES AND METHODS

Theory	Construct	Hyp.	Method
Task-based models	AI exposure, task substitution	H1, H3	DiD, survey
RBTC / SBTC	Entry-level share, requirements	H2	Posting text analysis
Human capital	Routine learning tasks	H3	Scales, SEM
Mentoring and social capital	Mentoring, networks	H4, H8	Scales, mediation
Segregation, intersectionality	Gender, social group	H5, H6	Interactions
Regional economies	Industry mix, digital readiness	H7	Shift-share, multilevel

H. The Indian Labour-Market Context

Although the framework is general, the empirical design is calibrated to India, where several features make the entry-level question especially pertinent. The country has a large cohort of young graduates entering the labour market each year, and a substantial share of formal white-collar entry jobs is located in services, including information technology, business-process outsourcing, banking and shared services. These sectors export services to firms in other countries, which links domestic entry opportunities to decisions taken abroad. At the same time, female labour-force participation is lower than that of men, women who are employed are concentrated in a narrower range of occupations, and social stratification by caste and class shapes access to education, networks and employer referrals [34], [35]. Regional disparities in connectivity, institutions and industrial structure are pronounced, so average national indicators conceal large differences between metropolitan hubs and smaller cities.

The Periodic Labour Force Survey offers repeated cross-sections that report employment and unemployment by age, gender, social group, education and region, which allows pre- and post-adoption comparisons of entry-level outcomes for broad occupation groups [34]. Its limitations for this question are that occupation categories are coarse, that the survey does not record employer adoption of AI, and that it does not measure mentoring or learning. These limitations motivate the combination with job-posting data and a primary survey, and they illustrate why the study treats secondary data as necessary but not sufficient.

I. Key Definitions

To avoid ambiguity, four terms are used consistently. The *hiring pipeline* is the sequence of steps by which an organization or labour market converts education into employment: advertising, screening, selection and onboarding. *Entry-level hiring* refers to the recruitment of workers with little or no prior employment experience. *Mentorship structures* are the formal and informal arrangements through which more experienced workers transfer knowledge and provide feedback, including assigned mentors, supervisory coaching, shadowing and structured training. *Gender-concentrated roles* are occupations in which one gender holds a share of employment substantially above its share of total employment; the study uses the female share of employment in each occupation, computed from labour-force data, as the operational measure.

J. Scope and Boundaries

The study concerns formal-sector, early-career employment in services and administrative occupations, and it examines AI adoption since late 2022. It does not attempt to forecast aggregate employment, to evaluate the productivity effects of AI on firms, or to assess technical performance of specific AI systems. Informal employment and self-employment are examined only to the extent that respondents report them. The study is concerned with associations that are consistent with the pipeline and capability channels and uses quasi-experimental methods where the data permit, but it does not claim that any single analysis proves causation.

III. RELATED WORK

A. Measuring Exposure to AI

A first group of studies constructs occupation-level exposure measures. Felten, Raj and Seamans mapped AI application progress to the abilities required by occupations and aggregated the result to industries and geographies [4]. Eloundou and colleagues combined human and model assessments to estimate that a large majority of the United States workforce could have at least a tenth of its tasks affected by large language models, and that a substantial minority could have at least half affected [3]. The International Labour Organization found that clerical support occupations are the most exposed group globally and noted a gender dimension arising from women's overrepresentation in such work [10]. The International Monetary Fund estimated that roughly forty percent of global employment is exposed to AI, with a higher share in advanced economies and distinct implications for emerging markets [11].

The Organisation for Economic Co-operation and Development has documented the diffusion of AI in workplaces and workers' perceptions of it [12]. These studies are indispensable but measure *potential*, not realized outcomes, and they aggregate across career stages.

B. Productivity and Task Evidence

Experimental and field evidence shows that generative AI changes how tasks are performed. Brynjolfsson, Li and Raymond studied customer-support agents and reported productivity gains concentrated among less experienced workers [5]. Noy and Zhang found that access to a generative writing assistant raised speed and quality on professional writing tasks and compressed differences between workers [49]. Such findings are often read as evidence that AI is a leveller. They also have a darker reading for hiring: if novices become more productive with AI assistance, or if AI substitutes for what novices do, the case for hiring fewer of them strengthens. Which reading prevails is an empirical question that depends on whether firms adjust hiring, training or both.

C. Early-Career Employment Effects

Emerging analyses of administrative payroll and hiring records examine the experience of younger workers. Brynjolfsson, Chandar and Chen report evidence of weaker employment outcomes for early-career workers in the most AI-exposed occupations relative to older workers in the same occupations and to less exposed occupations [6]. Autor has argued that AI could be used to extend expertise to a wider group of workers and rebuild middle-class jobs, which makes the design of adoption, and not merely its technical potential, a decisive factor [48]. The World Economic Forum's employer surveys point to shifting skill requirements and expectations of workforce restructuring [13]. This literature is young, concentrated in advanced economies and largely silent on mentorship and on the intersection of gender, social group and region.

D. Mentoring and Career Development

The mentoring literature is mature in organizational psychology but largely disconnected from labour economics. Meta-analytic evidence links mentoring to compensation, promotion and satisfaction [16], and conceptual work distinguishes career and psychosocial functions [15]. We are not aware of studies that connect the intensity of AI adoption to measured changes in junior mentoring and structured on-the-job learning. This disconnect is one of the principal opportunities addressed by the present work.

E. Gender, Social Group and Regional Evidence

Studies of polarization show that low-skill service employment grew in some regions while routine middle-skill work contracted [28], and that exposure to technological change varies across local labour markets [4], [25]. Evidence on gender indicates higher exposure in clerical roles [10], [12]. For India, the Periodic Labour Force Survey provides repeated cross-sections on employment by gender, social group, education and region [34], but evidence linking AI adoption to entry-level outcomes in such data is limited, and the international trend

picture is summarized by the ILO [35]. Consequently, evidence for developing regional economies, where social stratification and service-export dependence shape the transmission of technological change, is thin.

F. Synthesis of Selected Studies

Table II summarizes the studies that most directly inform the framework and indicates what each contributes and what it leaves open.

TABLE II
SELECTED STUDIES AND THEIR RELEVANCE

Study	Focus	Relevance to this paper
Autor et al. [2]	Computerization and task content	Routine versus non-routine task logic
Acemoglu and Restrepo [1]	Displacement and reinstatement of labour	Net effect depends on new tasks
Felten et al. [4]	AI exposure by occupation, industry, geography	Basis for regional exposure
Eloundou et al. [3]	Task exposure to language models	Share of tasks affected
Gmyrek et al. [10]	Global generative-AI exposure	Clerical exposure, gender dimension
Cazzaniga et al. [11]	Exposure in advanced and emerging economies	Cross-country contrast
Brynjolfsson et al. [5]	Field evidence, customer support	Larger gains for novices
Noy and Zhang [49]	Experiment on writing tasks	Speed and quality gains
Brynjolfsson et al. [6]	Early-career employment in exposed jobs	Entry-level weakness
Allen et al. [16]	Meta-analysis of mentoring	Career benefits of mentoring
Autor [48]	AI and middle-class work	Design of adoption matters

G. Critical Assessment of the Literature

Three features of the evidence base motivate the design of this study. First, most exposure measures describe technical potential, whereas the policy question concerns realized outcomes: whether a task that a model *could* perform is in fact transferred away from a junior employee, and with what consequences for hiring and learning. Second, the empirical record is dominated by advanced economies with rich administrative data, so conclusions may not transfer to settings characterized by large informal sectors, service-export dependence, social stratification and uneven connectivity. Third, the outcomes studied are overwhelmingly employment and earnings. Whether workers *learn* at the same rate when AI mediates their tasks is rarely measured, even though learning determines long-run earnings and mobility. A study that measures hiring, learning and mentoring jointly, and that disaggregates by gender, social group and region, therefore addresses a distinct and consequential blind spot.

H. Algorithmic Hiring and Automated Screening

A parallel development to AI at work is the use of automated tools in recruitment, including résumé parsing, online assessments and video interviews. Such tools can process large applicant volumes cheaply, which is attractive to employers when entry-level applicant numbers are high. They can also

introduce or amplify bias if trained on historical hiring data that reflect past inequality, or if they screen on proxies such as institution attended, employment gaps or language patterns that correlate with gender and social background. Policy institutions have flagged both the opportunities and the risks of algorithmic management and hiring [12], [13]. For the present study the relevance is twofold. Automated screening is a possible *mechanism* linking AI to narrower entry pipelines, since filters may encode experience thresholds, and it is a possible *channel of unequal treatment* between groups. The survey therefore asks whether respondents encountered automated screening and whether they judged it fair (items H8 and H9), and the interviews probe employer decisions about its use.

I. Reskilling and the Digital Divide

Reskilling is widely proposed as the answer to technological disruption, yet access to it is uneven. Training that is costly, delivered online, scheduled during working hours or taught in a language the learner finds difficult is effectively unavailable to some groups. Women with caregiving responsibilities, workers in smaller cities with weaker connectivity, and first-generation learners without peer guidance may face the heaviest constraints. International institutions emphasize that the benefits of AI depend on digital readiness and that the divide may widen without deliberate policy [11], [35]. The study measures reskilling access along four dimensions (affordability, employer support, digital readiness and time) so that interventions can be matched to the actual binding constraint rather than to a generic training offer.

IV. RESEARCH GAP AND PROBLEM STATEMENT

A. Research Gap

Synthesizing Sections II and III, the central gap addressed by this paper is the following. Existing research concentrates on mid-career displacement. There is limited evidence on how AI affects early-career hiring pipelines, junior mentorship structures, and demographic groups concentrated in exposed roles, such as women in administrative and service industries. Six specific shortfalls follow.

- 1) **Career-stage blindness.** Exposure studies rarely separate entry-level from mid- and senior-level workers within the same occupation, and the inflow of new hires is studied less than the outflow of layoffs.
- 2) **Unmeasured pipeline effects.** Little empirical work asks whether reduced junior hiring or inflated experience requirements weaken the future supply of experienced workers or restrict access to first jobs.
- 3) **Mentorship as a missing variable.** The erosion of learning-by-doing and mentoring is discussed conceptually but has not been measured against AI adoption intensity.
- 4) **Entry-level gender and intersectional effects.** Women's concentration in exposed administrative and service roles is documented, but how AI changes their *entry* into careers, and how that interacts with age, social group, first-generation status and region, remains under-analysed.

- 5) **Regional granularity and developing-economy evidence.** National averages conceal regional variation, and most evidence comes from advanced economies.
- 6) **Weak linkage between numbers and mechanisms.** Quantitative studies seldom connect to qualitative evidence on employer decisions, so policy guidance remains generic.

B. Problem Statement

There is insufficient empirical evidence on how AI adoption alters early-career hiring pipelines, weakens or reshapes junior mentorship and skill formation, and disproportionately affects women and other marginalized groups in exposed administrative and service roles across regional economies with different structures. Without such evidence, education, labour and reskilling policy risks concentrating on displaced mid-career workers while leaving the next cohort's access to stable careers unaddressed, thereby reinforcing existing inequalities.

C. Research Questions

- **RQ1.** How has AI adoption changed hiring volumes, experience requirements and skill demands for entry-level roles relative to mid- and senior-level roles?
- **RQ2.** How have junior mentorship, structured training and on-the-job learning changed in organizations that adopt AI?
- **RQ3.** How do early-career outcomes differ by gender and other demographic characteristics, including intersections, in administrative and service occupations?
- **RQ4.** How do regional economic characteristics moderate these effects?
- **RQ5.** Through which mechanisms do unequal outcomes arise, and which interventions could reduce them?

V. RESEARCH CONTRIBUTION

The paper contributes along four dimensions.

A. Theoretical Contribution

The study extends the task-based and human capital literatures by adding three dimensions that are usually treated separately: career stage, mentorship and intersectional-regional position. It proposes that AI affects early-career workers through two linked mechanisms, a *pipeline channel* operating through hiring volumes and experience requirements and a *capability channel* operating through the erosion or redesign of mentoring and learning-by-doing, and that both channels are moderated by demographic and regional context. This yields testable predictions that the aggregate displacement framing cannot generate.

B. Empirical Contribution

The study shifts the unit of analysis from the occupation to the entry-level position within the occupation, and from outflows to inflows. It supplies original primary data on mentoring, hiring experience, reskilling access and AI use among early-career workers in contrasting regional economies, and it links these data to secondary labour-force and job-posting evidence. It also contributes evidence from a developing-economy setting, where such data are scarce.

C. Methodological Contribution

The design integrates exposure measurement, quasi-experimental estimation, survey measurement of mentorship and text analysis of postings within one sequential explanatory mixed-methods framework [24]. It contributes a validated survey instrument with multi-item scales for AI task substitution, mentoring intensity, network access and reskilling access (Appendix A), a transparent analysis plan with pre-specified hypotheses and multiple-testing control, and a Regional Vulnerability and Resilience Index whose weights are subjected to sensitivity analysis.

D. Practical and Policy Contribution

The findings are intended to give governments, universities, training providers and employers evidence for targeted rather than generic workforce policy: apprenticeship and internship redesign, structured mentoring models that pair human coaching with AI tools, reskilling directed at gender-concentrated roles, audits of algorithmic hiring tools and regional investment priorities. By identifying which groups and places are most exposed at the first rung, the study supports early intervention before disadvantages compound over a career.

E. Positioning and Novelty

Relative to the exposure literature [3], [4], [10], this paper moves from potential to realized outcomes and from the occupation to the career stage. Relative to emerging early-career studies based on payroll and hiring records [6], it adds measurement of mentoring and learning, which administrative data cannot observe, and it adds a developing-economy setting with intersectional and regional disaggregation. Relative to the mentoring literature [15], [16], it links mentoring to technology adoption, a connection that has not been tested at scale. Relative to policy reports [12], [13], it supplies a transparent index and hypotheses that can be falsified. The combination of these four elements within one design is the paper's distinctive contribution.

VI. CONCEPTUAL FRAMEWORK AND HYPOTHESES

Fig. 1 summarizes the framework. AI exposure and adoption operate on early-career outcomes through the hiring pipeline and through mentorship and learning-by-doing. Demographic position and regional economy moderate both channels, and access to reskilling, referral networks and digital readiness are tested as mediators of the demographic gap.

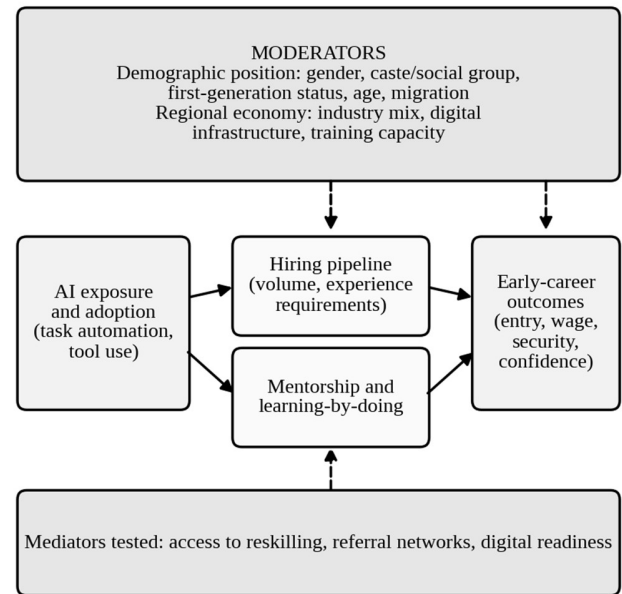


Fig. 1. Conceptual framework of asymmetric disruption.

The framework yields eight hypotheses, stated in directional form so that they can be falsified.

- **H1 (pipeline).** Greater AI exposure of an occupation is associated with a larger decline in entry-level hiring than in senior-level hiring within the same occupation.
- **H2 (requirements).** Entry-level postings in highly exposed occupations show a larger increase in required experience than those in less exposed occupations.
- **H3 (task substitution).** Early-career workers whose tasks are more heavily AI-assisted report fewer routine learning tasks and lower perceived skill growth.
- **H4 (mentorship).** Higher AI adoption intensity in the employer is associated with lower mentoring frequency and less structured training for juniors.
- **H5 (gender).** The negative association between AI exposure and entry-level outcomes is stronger for women in administrative and service occupations than for men in comparable roles.
- **H6 (intersection).** The association is stronger still for intersectional groups, defined as women who are also first-generation graduates or from socially disadvantaged groups.
- **H7 (region).** The association is stronger in regions with a larger share of routine-service employment and weaker digital infrastructure, and weaker in diversified metropolitan regions.
- **H8 (mediation).** Access to reskilling, referral networks and digital readiness partially mediates the demographic gap in outcomes.

B. Derivation of the Hypotheses

Pipeline hypotheses (H1, H2). The task-based framework implies that when a technology performs a task previously done by a worker, demand for labour in that task falls unless offset by new tasks [1]. Junior positions are intensive in codifiable tasks, so the displacement effect should be larger for them than

for senior roles whose value lies in tacit judgment and relationships [48]. Firms facing uncertainty about the value of a novice relative to an AI-assisted experienced worker may respond in two ways: they may hire fewer juniors (H1), or they may keep the title but demand more experience, effectively converting entry-level openings into mid-level ones (H2). Both responses reduce access to first jobs even if total employment in the occupation is stable, which is why the study measures inflows and posted requirements and not only stocks.

Capability hypotheses (H3, H4). Human capital theory treats on-the-job learning as an investment whose cost is borne partly through the novice's low early productivity [7]. If AI performs the practice tasks, learning-by-doing opportunities fall (H3). Mentoring is costly in senior time; when AI tools enable seniors to produce more by themselves or shift their attention to reviewing machine output, the opportunity cost of coaching rises and structured training may be cut (H4). Conversely, if firms deliberately redeploy time released by AI toward coaching, the hypotheses would be falsified, which is why Module 2 asks employers directly about coaching time and training provision.

Distributional hypotheses (H5, H6). Women are overrepresented in administrative and customer-service roles in many economies, and these roles rank high in exposure [10], [12]. If the pipeline channel operates, it should therefore bear more heavily on women's entry into careers. Intersectionality predicts that the effect is not additive: a woman who is also a first-generation graduate or from a socially disadvantaged group has fewer informal routes around a narrowed formal pipeline [17], [33]. This yields the interaction structure tested in Section IX.

Regional hypothesis (H7). Regions with a high share of routine-service employment have greater expected exposure by construction of the shift-share measure, but impact also depends on adaptive capacity. Regions with weaker connectivity and training institutions offer fewer alternatives to displaced entrants, so the same exposure should translate into larger and more persistent declines in entry outcomes [4], [25].

Mediation hypothesis (H8). Access to reskilling, referral networks and digital readiness are resources that convert exposure into adaptation. Because these resources are unequally distributed by gender, social group and region, they are natural candidates for explaining part of the demographic gap. The hypothesis states *partial* mediation because employer practices and discrimination are expected to produce a residual direct effect.

C. A Stylized Model of Junior Hiring

To make the pipeline and capability channels precise, consider a stylized model. It is not estimated; its role is to show which comparative statics the empirical design should be able to detect. Tasks within an occupation are indexed by complexity x in $[0, 1]$. Junior workers can perform tasks with complexity up to x_J , and senior workers can perform all tasks. AI can perform tasks up to a frontier $g(\theta)$, where θ denotes AI capability and $g'(\theta) > 0$. A firm assigns a task to AI when AI is able to perform it at lower cost than a junior. Junior employment is then proportional to the range of tasks that remain with juniors:

$$J^* = \phi \cdot \max\{0, x_J - g(\theta)\} \quad (1)$$

where ϕ converts the task range into headcount. Differentiating, junior hiring falls as AI capability rises whenever the frontier lies below the juniors' range,

$$\partial J^* / \partial \theta = -\phi \cdot g'(\theta) < 0 \text{ for } g(\theta) < x_J \quad (2)$$

Proposition 1 (pipeline channel). Junior hiring is non-increasing in AI capability and reaches zero once the AI frontier covers the juniors' entire task range. This yields H1. If firms also raise the minimum experience for the remaining tasks, the range x_J effectively shrinks, producing H2.

Now let learning by a junior depend on the share of retained tasks that provide practice. Let ℓ denote learning per junior and λ_0 the learning rate when all of a junior's tasks are retained:

$$\ell = \lambda_0 \cdot [x_J - g(\theta)] / x_J \quad (3)$$

The supply of experienced workers a cohort eventually produces is proportional to $J^* \ell$, which falls with θ through *both* factors:

$$S = \rho \cdot J^* \cdot \ell = \rho \phi \lambda_0 \cdot [x_J - g(\theta)]^2 / x_J \quad (4)$$

Proposition 2 (capability channel). Future experienced-worker supply declines faster than junior hiring as AI capability rises, because the number of juniors and the learning each receives fall together. The effect is convex: modest advances in θ reduce S proportionally more than J^* . This motivates measuring learning (RLT) and mentoring (MEN) in addition to hiring.

Finally, consider two groups g with different task ranges x_J^g , where group-concentrated occupations such as administrative support have narrower junior ranges because many of their tasks are codifiable. The capability at which junior hiring in group g reaches zero is $\theta_g = g^{-1}(x_J^g)$, which is lower for the group with the narrower range.

Proposition 3 (asymmetry). At any given θ , groups concentrated in occupations with narrower junior task ranges experience a larger proportional decline in junior hiring and reach exhaustion of entry positions sooner. This is the formal basis for H5 to H7, with gender and region entering through the occupational and industrial composition that determines x_J^g . The model also clarifies what would falsify the framework: if firms respond to AI by *expanding* junior ranges through retraining (raising x_J), the decline could be offset.

VII. RESEARCH METHODOLOGY

A. Research Design

A sequential explanatory mixed-methods design is adopted [24]. Quantitative strands establish patterns in hiring and outcomes across career stages, demographic groups and regions. A qualitative strand then explains mechanisms, particularly around mentorship, which administrative data cannot observe. Findings are integrated in a final synthesis stage. Fig. 2 shows the phases.

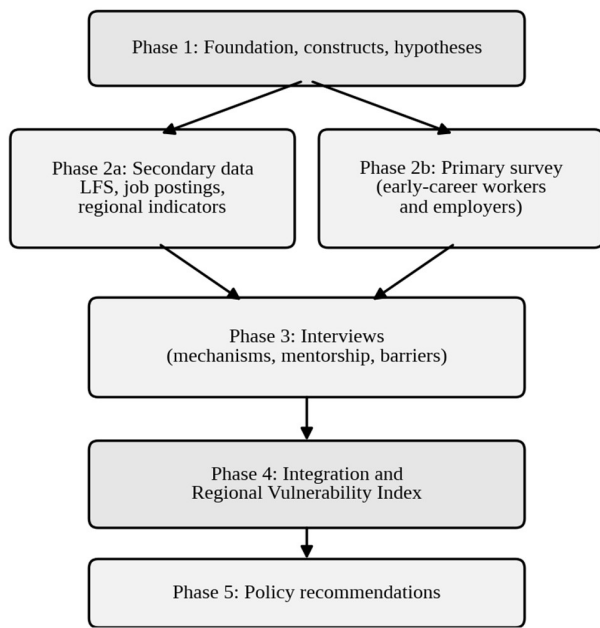


Fig. 2. Phases of the research design.

B. Constructs and Operational Definitions

Entry-level role. A position is classified as entry-level if it requires zero to two years of experience, or if the respondent has been in the labour market for three years or fewer. Both definitions are used to test robustness. **AI exposure.** Occupational exposure is measured with a task-based score constructed by mapping task descriptions in occupational databases to AI capability indicators, following the logic in [3] and [4]. Individual-level adoption intensity is measured by the survey. **Mentoring intensity.** A multi-item scale covering frequency of feedback, presence of an assigned mentor, structured training and observation opportunities, adapted from the career and psychosocial functions in [15]. **Marginalized demographic.** Defined by gender, social category, first-generation graduate status, disability, migration status and age; the coding follows national categories and respondents may decline to answer. **Regional vulnerability.** A composite of exposure, share of junior roles, demographic risk, digital infrastructure, training capacity and economic diversification.

C. Study Setting and Regions

The study is designed for an Indian setting, although the instrument is portable. Four contrasting regional economies are proposed: a metropolitan services hub (for example Mumbai), a diversified technology and manufacturing city (for example Pune), a Tier-2 city with a more limited services base (for example Nagpur) and a semi-urban or district-level economy. The final selection should reflect the authors' access and the availability of secondary data. The reference period is 2018 to the present, giving pre-adoption baseline years and a post-adoption period that begins with the public release of widely used generative AI tools in late 2022.

D. Data Sources

TABLE III
DATA SOURCES AND THEIR ROLES

Source	Role in the study
Occupational task databases (e.g., O*NET, national classifications)	Task-level exposure scores; mapping to occupation codes
National labour force surveys (e.g., PLFS [34]; ILOSTAT)	Employment, wages, transitions by gender, social group, age, region
Job-posting data (portals, aggregators, permitted scraping)	Entry-level share, required experience, skill mentions over time
Regional statistics (state and district accounts, connectivity data)	Industry mix, digital infrastructure, training capacity
Primary survey (Appendix A)	Mentoring, hiring experience, AI use, reskilling and network access
Semi-structured interviews (Appendix B)	Employer decisions, learning practices, barriers, coping

E. Sampling and Sample Size

The survey targets two populations: early-career workers and recent graduate job seekers (Module 1, main sample) and managers or human-resource professionals who take hiring and training decisions (Module 2, supplementary sample). The sample size for Module 1 is derived from the standard formula for estimating a proportion, $n_0 = z^2 p(1 - p)/e^2$, which with $p = 0.5$, $z = 1.96$ and $e = 0.05$ gives about 384 [29]. Allowing for a design effect of 1.5 due to clustering by region and sector, the minimum is about 576, and allowing for roughly thirty percent non-response or screening failure, the recruitment target is approximately 800 respondents, or about 200 per region. A power calculation for multiple regression with a small effect size (Cohen's $f^2 = 0.02$), $\alpha = 0.05$, power of 0.80 and up to fifteen predictors indicates that a sample of this order is adequate to detect effects of that size [31]. Module 2 targets at least 120 employer respondents, with a minimum of 25 per region.

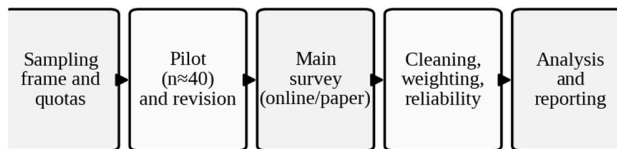
Because a sampling frame of all early-career workers is unavailable, recruitment will use stratified quota sampling by region, gender and sector (administrative and support, customer service, information technology and other professional, and other), through university alumni offices, training institutes, professional associations, employer partners and social-media networks. Post-stratification weights will adjust the sample toward the distribution of gender, social category and sector in national labour-force benchmarks. The non-probability design limits formal statistical generalization, and this is stated as a limitation in Section XIII.

F. Survey Design and Administration

The questionnaire (Appendix A) is organized into modules covering screening and consent, demographics, employment profile, AI exposure at work, task change, hiring-pipeline experience, mentoring and learning, networks and reskilling access, outcomes and open responses, with a short employer module. Attitudinal items use five-point Likert scales [30]. Items were drafted from the constructs in Section VII-B and adapted from established technology-acceptance and mentoring measures where applicable [15], [36], [37]. Design follows the

tailored-design principles for web and mixed-mode surveys, including short pages, plain language, a neutral order of sensitive questions and an explicit option to decline [42]. The instrument will be offered in English, Hindi and Marathi, translated and back-translated by independent bilingual reviewers.

Quality gates: consent check, attention items, duplicate screening



Feedback loop: pilot results revise wording, order and length

Fig. 3. Survey administration workflow.

A pilot of about 40 respondents will test comprehension, timing (target under 15 minutes), item non-response and preliminary reliability. Items with Cronbach's α below 0.70 for their scale, ambiguous wording or high non-response will be revised. The main survey will use an online form with an assisted paper option for respondents with limited connectivity, which reduces exclusion of digitally disadvantaged groups. Quality gates include an informed-consent check, attention-check items, duplicate-response screening and range checks.

G. Qualitative Strand

Between 25 and 40 semi-structured interviews will be conducted with early-career workers drawn from different demographic backgrounds, supervisors and mentors, hiring managers, and training providers across the selected regions, continuing until thematic saturation is observed [43]. The interview guide appears in Appendix B. Interviews are audio-recorded with consent, transcribed, anonymized and analysed using reflexive thematic analysis [23]. Two coders will code a subset independently, and agreement is reported using Cohen's kappa interpreted with the conventional benchmarks [44].

H. Job-Posting Text Analysis

Job postings will be collected for the selected regions and occupations, subject to the terms of service of each source. Text processing will extract required years of experience, degree requirements, skill mentions and AI-related terms, and will classify postings as entry-level using a rule-based screen validated against a manually labelled random sample. The classification is evaluated with precision, recall and F1 against the manual labels. Postings that are labelled entry-level but request experience typical of mid-level roles are flagged as *inflated*, and their share over time is the key indicator for H2.

I. Regional Selection and Context

Regions are selected on four criteria: industry structure (share of employment in administrative, outsourcing and routine-service activities), city size and diversification, digital infrastructure (broadband and mobile data access), and the presence of universities and training institutions. The aim is

maximum variation: at least one region with a high concentration of routine-service work, one diversified region with strong training capacity, and one or two regions with thinner labour markets. For each region the study will assemble a profile from public statistics, including employment by industry and gender, the share of young workers, and indicators of connectivity. These profiles feed the shift-share exposure measure and the Regional Vulnerability and Resilience Index, and they provide context for interpreting the survey and interviews. Region labels A to D in the instrument are replaced by the final region names at deployment.

J. Instrument Development, Translation and Pilot Protocol

Instrument development follows five steps. *Item generation*: items are drafted from the construct definitions and adapted from published mentoring and technology-adoption measures [15], [36], [37]. *Expert review*: five reviewers (two labour economists, one organizational psychologist, one HR practitioner and one survey methodologist) rate each item for relevance and clarity, and items endorsed by fewer than four of five reviewers are revised or dropped. *Cognitive interviews*: about ten respondents from different backgrounds think aloud while answering, which exposes ambiguous terms such as *mentor* or *AI tools*. *Translation*: forward translation by one bilingual translator, reconciliation by a second, and back-translation by an independent third, with discrepancies resolved by discussion. *Pilot*: about 40 respondents complete the near-final instrument for timing, non-response, distributional checks and preliminary reliability.

Items with floor or ceiling effects above eighty-five percent, corrected item-total correlations below 0.30, or non-response above five percent are flagged for revision. The pilot sample is excluded from the main analysis to avoid contamination.

K. Data Management, Pre-Registration and Reproducibility

The hypotheses, variable definitions and primary models will be recorded in a time-stamped analysis plan on a public registry before the main data are analysed, which separates confirmatory from exploratory analysis. Analysis code will be maintained under version control, and a data dictionary will document every variable, its source item and its coding. De-identified survey data and code will be shared in a repository to the extent permitted by consent and by the terms of third-party data sources. Raw identifiable data will be retained only for the period required by the ethics approval and then securely destroyed.

L. Project Timeline

TABLE IV
INDICATIVE PROJECT TIMELINE

Stage	Activities	Months
1	Ethics approval, item generation, expert review	1–2
2	Cognitive interviews, translation, pilot	3–4
3	Main survey; collection of postings and secondary data	5–8
4	Interviews and transcription	7–10
5	Quantitative analysis, index construction	9–12
6	Integration, policy brief, manuscript	12–14

M. Recruitment Channels and Quota Targets

Recruitment will use several channels in parallel to reduce dependence on any single network: university and college placement cells and alumni associations, vocational and skill-training institutes, professional associations and employer partners, local employment offices, and social-media groups for graduates. Table V gives target quotas, which are design targets and not findings. Quotas for sector reflect the aim of over-sampling administrative and service roles relative to their population share so that gender-concentrated roles can be analysed with adequate precision.

TABLE V
PROPOSED QUOTA TARGETS PER REGION (DESIGN TARGETS)

Dimension	Target	Rationale
Total per region	≈ 200	Regional comparison
Women	≥ 50%	Gender analysis
Administrative, customer-service roles	≥ 35%	Exposed roles
First-generation graduates	≥ 25%	Intersection analysis
Disadvantaged social groups	≥ 20%	Group comparison
Job seekers (not yet employed)	≈ 20%	Pipeline experience

N. Constructing the Occupation-Level Exposure Score

The occupational exposure score is built in five steps so that it can be replicated. First, task statements for each occupation are drawn from a standard occupational database and mapped to the national occupation classification used in the labour-force survey through an official or published crosswalk. Second, each task is rated for the feasibility of performing it with current language and multimodal models, using a published rubric and, for a validation subsample, ratings by independent human raters. Third, task ratings are aggregated to the occupation by averaging, weighting tasks by importance where the database provides weights. Fourth, the score is standardized and, for sensitivity analysis, converted to terciles and to a binary high-exposure indicator. Fifth, the score is merged with the labour-force and posting data and with the female share of employment in each occupation. Agreement between human and model ratings is reported, and occupations for which the crosswalk is many-to-one are flagged and tested with and without exclusion.

O. Linking Postings to Occupations

Job postings rarely carry official occupation codes. Titles and descriptions are therefore matched to occupations using a combination of rule-based dictionaries and sentence-embedding similarity to occupation descriptions, with a manual check of a random sample of matches. Postings with low match confidence are excluded, and results are compared with and without them. Duplicate postings across portals are identified by title, employer and location, and the first appearance is retained so that reposting does not inflate counts. The share of postings that are entry-level, the distribution of required years of experience and the frequency of AI-related terms are then tracked by month, region and occupation group.

P. Sensitivity of Sample Size to Design Assumptions

Table VI shows how the required sample varies with the tolerated margin of error. The recruitment target is chosen to deliver adequate precision for regional comparisons at moderate cost; larger targets would be desirable for narrow intersectional subgroups, which are therefore treated as exploratory unless cell sizes are sufficient.

TABLE VI
SAMPLE SIZE UNDER ALTERNATIVE MARGINS OF ERROR

Margin of error	Base <i>n</i>	With design effect 1.5	Recruitment target (70% usable)
±3%	1,068	1,602	≈ 2,290
±5%	385	578	≈ 825
±7%	197	296	≈ 425

Q. Researcher Reflexivity

Qualitative interpretation is shaped by the researcher's background. The authors will keep a reflexive journal recording assumptions, reactions to interviews and decisions about coding, will discuss emerging themes with colleagues outside the project, and will share preliminary interpretations with a small group of participants to check that the account resonates with their experience. Where the researchers share or do not share characteristics with participants (gender, social group, institutional affiliation), this is noted and considered in the analysis.

R. Risk Register and Contingency Plans

TABLE VII
ANTICIPATED RISKS AND CONTINGENCIES

Risk	Contingency
Low response in one region	Extend fieldwork; add channels; reweight
Low female or disadvantaged-group share	Targeted outreach through partner institutions
Gatekeeper refusal	Alternative institutions; direct channels
Posting data access restricted	Use permitted aggregates; manual samples
Interview attrition	Over-recruit by 20%; replace participants
Translation ambiguity	Re-run cognitive interviews on affected items
Delay in ethics approval	Prepare documents early; stage activities

S. Dissemination Plan

Results will be disseminated through a peer-reviewed article, a conference presentation, a short policy brief for ministries, training bodies and employer associations, and a plain-language summary for participants and partner institutions. An online dashboard of the regional index and entry-level indicators will be released so that policymakers can monitor changes. Anonymized data and code will be deposited in a repository in line with the data-sharing plan, and a replication package including the survey instrument and codebook will accompany the article.

VIII. MEASUREMENT AND SCALE DEVELOPMENT

A. Constructs and Scale Structure

The survey measures six multi-item constructs, each scored on a five-point scale from strongly disagree to strongly agree.

Table VIII lists the constructs, the number of items and the intended use. Items are listed in Appendix A.

TABLE VIII
SURVEY CONSTRUCTS AND SCALE PROPERTIES

Construct	Items	Used for
AI task substitution (AIS)	6	H3, H5, H6
Routine learning tasks (RLT)	4	H3
Mentoring intensity (MEN)	8	H4, H8
Network and referral access (NET)	4	H8
Reskilling access (RSK)	4	H8
Career outcomes (COU): security, confidence, intent	6	All outcome models

Internal consistency is assessed with Cronbach's alpha [22],

$$\alpha = (k)/(k-1) \left(1 - \frac{\sum \sigma_i^2}{\sigma_T^2} \right) \quad (5)$$

where k is the number of items, σ_i^2 the variance of item i and σ_T^2 the variance of the total score. Values of 0.70 or above are treated as acceptable for research purposes [21]. Construct validity is examined with exploratory factor analysis on the pilot and confirmatory factor analysis on the main sample, with model fit judged against the combined criteria of CFI, TLI, RMSEA and SRMR [45], and with average variance extracted and heterotrait-monotrait ratios for discriminant validity [38]. Because all attitudinal measures come from one respondent, common-method bias is assessed with Harman's single-factor test and a marker-variable check, and procedural remedies such as separated scales and reverse-coded items are built into the instrument [46].

B. Rationale for the Scales

AI task substitution (AIS). The scale captures the respondent's perception that tasks at their level are being performed by AI, that expectations of output have risen, and that the number of colleagues at their level has fallen. It includes one item on quality gains so that respondents who benefit from AI are not forced to express only negative views, which reduces acquiescence and makes the scale useful for identifying who gains as well as who loses.

Routine learning tasks (RLT). The scale operationalizes the capability channel by asking whether practice-type tasks have declined and whether the respondent understands the reasoning behind tasks and not only the tool. It is designed to separate *output* (what is produced) from *learning* (what is understood), since AI can raise one while lowering the other.

Mentoring intensity (MEN). Items cover designation of a mentor, feedback frequency, explanation of decisions, observation, structured training and review, which together span the career and psychosocial functions identified in the mentoring literature [15]. One reverse-scored item asks directly whether coaching time has fallen since AI tools were introduced, providing a within-scale check on the perceived change.

Network (NET) and reskilling (RSK). These scales measure resources that moderate or mediate the demographic gap. NET captures referral access and the seniority of contacts, in line with theories of weak ties and social capital [33]. RSK covers affordability, employer support, digital readiness and time, which allows the analysis to separate financial, infrastructural

and time constraints, the last of which is especially relevant to women with caregiving responsibilities.

Career outcomes (COU). The outcome scale covers security, confidence in progression, pay satisfaction, intention to leave the occupation, perceived fairness and optimism. Single behavioural outcomes (time to first job, number of offers) from the hiring-pipeline section complement the attitudinal outcomes and reduce reliance on self-assessed measures alone.

IX. DATA ANALYSIS PLAN

A. Preparation

Data will be cleaned for duplicates, straight-lining and implausible values. Missing data will be examined for patterns; where missingness exceeds five percent on a key variable, multiple imputation with chained equations will be used and compared with complete-case results. Weights will be constructed by raking to benchmarks. Occupation codes are harmonized across sources, and regional boundaries are standardized.

B. Descriptive and Bivariate Analysis

Descriptive statistics will be reported by region, gender, social group and sector. Group differences will be tested with chi-square tests for categorical outcomes and with t -tests, ANOVA or their non-parametric equivalents for continuous ones, accompanied by effect sizes [31]. Because many hypotheses and subgroups are tested, p -values will be adjusted using the Benjamini-Hochberg false-discovery-rate procedure [32].

C. Survey Regression Models

For a binary outcome Y (for example, whether the respondent obtained a role matching their qualification within six months of graduating), a logistic model with an interaction between AI exposure E and a demographic indicator F is estimated [39]:

$\text{logit Pr}(Y_{ir} = 1) = \alpha + \beta_1 E_i + \beta_2 F_i + \beta_3 (E_i \times F_i) + \gamma' X_i + \delta_r(6)$ where i indexes respondents, r regions, X is a vector of controls (age, education, field of study, sector, employer size, tenure) and δ_r are region fixed effects. A negative β_3 for women in administrative and service roles supports H5. For continuous outcomes such as the mentoring score, ordinary least squares with robust standard errors clustered by employer (region effects enter as fixed effects, because four regions are too few for reliable region-level clustering) is used, and ordered logit is applied to Likert outcomes. Intersectional effects (H6) are tested with three-way interactions and with subgroup models, and are reported as predicted probabilities with confidence intervals so that magnitudes can be read directly.

D. Mediation and Moderation

To test H8, the total effect c of a demographic characteristic on an outcome is decomposed into a direct effect c' and an indirect effect through a mediator M such as reskilling access [40]:

$$c = c' + a \times b \quad (7)$$

where a is the effect of the characteristic on the mediator and b the effect of the mediator on the outcome controlling for the characteristic. The indirect effect ab is evaluated with bias-corrected bootstrap confidence intervals from 5,000 resamples

[41]. Regional moderation (H7) enters through region-level interaction terms and multilevel models with random intercepts for region.

E. Structural Model

Where the measurement model is satisfactory, a structural equation model will estimate the paths in Fig. 1 simultaneously, with AI task substitution predicting routine learning tasks and mentoring intensity, and these predicting career outcomes [38]. Model fit will be reported using the indices above, and alternative specifications will be compared using information criteria.

F. Quasi-Experimental Analysis of Secondary Data

Survey data are cross-sectional and cannot by themselves separate AI effects from the cycle. The secondary labour-force and posting data are therefore analysed with difference-in-differences and event-study designs that compare high- and low-exposure occupations before and after widespread adoption, separately for junior and senior workers [18]. A triple-difference specification tests whether the junior penalty differs by demographic group:

$$Y_{jort} = \alpha + \beta(\text{Exp}_o \times \text{Post}_t \times \text{Junior}_j) + \mu_o + \lambda_t + \varepsilon_{jort} \quad (8)$$

where j indexes workers, o occupations, r regions and t time periods, all lower-order interactions are included, μ_o and λ_t are occupation and period fixed effects, and the coefficient of interest β captures the additional post-adoption change in outcomes for junior workers in highly exposed occupations. Because adoption timing is staggered across regions and firms, estimators robust to heterogeneous treatment effects are used in preference to the conventional two-way fixed-effects estimator [19], [20]. Pre-trend event-study plots and placebo tests on earlier periods probe the parallel-trends assumption, and macroeconomic controls address the confounding effect of interest-rate and post-pandemic hiring cycles.

G. Regional Exposure and the Vulnerability Index

A region's expected exposure is calculated with a shift-share construction [25]:

$$E_r = \sum s_{rk} e_k \quad (9)$$

where s_{rk} is the employment share of industry k in region r and e_k is the average exposure score of industry k . The Regional Vulnerability and Resilience Index combines standardized indicators,

$$\text{RVRI}_r = \sum w_j z_{jr}, \quad \sum w_j = 1 \quad (10)$$

where z_{jr} is the standardized value of indicator j for region r and w_j its weight. Indicators are exposure, share of junior roles, concentration of gender-exposed occupations, demographic risk, digital infrastructure, training capacity and diversification. Weights are first set equal, then varied in a Monte Carlo sensitivity analysis, and the stability of regional rankings is reported using rank correlations.

H. Decomposition of Group Gaps

Observed gaps between groups, such as men and women in entry-level outcomes, are decomposed into the part explained by differences in characteristics and the part attributable to differences in returns using the Blinder-Oaxaca approach [26], [27]. A large unexplained component points toward unequal

treatment or unobserved constraints rather than differences in qualifications.

I. Qualitative Integration

Interview themes are matched to quantitative findings through joint displays, in which each hypothesis is listed alongside its statistical result and the interview evidence that supports, qualifies or contradicts it. Divergent cases are examined explicitly instead of being averaged away [24].

J. Robustness and Sensitivity Analysis

The principal results will be subjected to the following checks, and results that do not survive them will be reported as fragile:

- alternative exposure measures, including task-based scores from different sources and a binary high-exposure indicator at several percentile thresholds;
- alternative definitions of *entry-level* (experience requirement, tenure, age band) and of *marginalized* groups;
- exclusion of one region at a time and of outlier occupations;
- placebo tests that assign a false adoption date in the pre-period;
- comparison of multiple imputation with complete-case analysis and of weighted with unweighted estimates;
- ordinal and linear specifications for Likert outcomes, and wild-cluster bootstrap inference where the number of clusters is small;
- specification-curve analysis to show how estimates vary across reasonable analytic choices.

K. Evaluation of Text Classification and Qualitative Agreement

The classifier that labels job postings as entry-level, and the flag for inflated requirements, are evaluated against a manually labelled random sample using precision, recall and the F1 score,

$$F1 = (2PR)/(P + R), \quad P = (TP)/(TP + FP), \quad R = (TP)/(TP + FN) \quad (11)$$

where TP, FP and FN are true positives, false positives and false negatives. Inter-coder agreement in the qualitative strand is summarized by Cohen's kappa,

$$\kappa = (p_o - p_e)/(1 - p_e) \quad (12)$$

where p_o is the observed proportion of agreement and p_e the proportion expected by chance. Values above 0.60 are interpreted as substantial agreement [44]. Disagreements are resolved by discussion and recorded in the audit trail.

L. Hypothesis-by-Hypothesis Test Specification

TABLE IX
TEST SPECIFICATION FOR EACH HYPOTHESIS

Hyp.	Outcome	Key predictor	Data and model
H1	Junior hiring rate	Exposure $\times$ Post $\times$ Junior	LFS; DiD, event study
H2	Required experience	Exposure $\times$ Post	Postings; DiD
H3	RLT score	AIS score	Survey; OLS, SEM
H4	MEN score	Employer AI intensity	Survey; OLS, multilevel
H5	Timely first job; COU	Exposure $\times$ Female	Survey; logit
H6	Timely first job; COU	Exposure $\times$ Female $\times$ Group	Survey; logit

Hyp.	Outcome	Key predictor	Data and model
H7	Entry outcomes	Exposure × Region traits	Survey, LFS; multilevel
H8	COU gap	Female, mediators	Survey; bootstrap

Each hypothesis is assigned a primary outcome and model before data analysis, and any additional specifications are labelled exploratory.

X. RESULTS-REPORTING FRAMEWORK

Fieldwork has not yet been completed. This section therefore provides the structure in which results will be reported, with all numerical cells left blank so that no value can be mistaken for an empirical finding. Once data collection ends, the authors will populate Tables X to XIII and convert Section XI from a conditional to a descriptive discussion.

A. Sample Profile

TABLE X
SAMPLE PROFILE (TO BE COMPLETED)

Character istic	Pune	Chen nai	Bengal uru	Hydera bad	Solap ur	Noida	Gurga on	Parbh ani	Total
Responde nts, n	26	20	20	26	2	4	2	1	101
Women, %	65.4	40.0	35.0	46.2	50.0	25.0	100.0	100.0	48.5
First-generation graduates (Yes), %	42.3	25.0	20.0	38.5	50.0	25.0	0.0	0.0	31.7
Admin/su pport roles, %	23.1	60.0	5.0	50.0	0.0	75.0	50.0	0.0	35.6
Age 20–35, %	73.1	25.0	40.0	34.6	100.0	50.0	50.0	100.0	46.5
Response rate, %	Not collec ted	Not collec ted	Not collec ted	Not collec ted	Not collec ted	Not collec ted	Not collec ted	Not collec ted	Not collec ted

B. Reliability and Validity

TABLE XI
SCALE RELIABILITY AND VALIDITY (TO BE COMPLETED)

Scale	Items planned in paper	Items available in survey	Exploratory α	AVE	CR
AIS	6	4	0.861	Not estimable	Not estimable
RLT	4	1	Not estimable (single item)	Not estimable	Not estimable
MEN	8	3	Not interpretable: items do not form a consistent scale	Not estimable	Not estimable
NET	4	2	0.883	Not estimable	Not estimable
RSK	4	2	0.912	Not estimable	Not estimable

Scale	Items planned in paper	Items available in survey	Exploratory α	AVE	CR
COU	6	2	Not interpretable: two-item outcome measure is inconsistent	Not estimable	Not estimable

C. Model Results

TABLE XII
LOGISTIC REGRESSION ON ENTRY OUTCOME (TO BE COMPLETED)

Predictor	Odds ratio	95% CI	p
AI-use frequency (E)	2.245	0.643–7.831	0.205
Female (F; male reference)	0.755	0.077–7.384	0.809
E × F	0.455	0.118–1.759	0.254
First-generation	Not estimated	—	—
Region fixed effects	No	—	—
Additional controls	No	—	—

D. Hypothesis Decisions

TABLE XIII
HYPOTHESIS SUMMARY (TO BE COMPLETED)

Hyp.	Expected sign	Test	Decision based on available survey
H1	Negative	DiD, event study	Not testable: no longitudinal hiring data
H2	Positive	Posting analysis	Not testable: no job-posting time series
H3	Negative	OLS, SEM	Partially assessable; AIS has 4 available items ($\alpha=0.861$), but full RLT scale is unavailable
H4	Negative	OLS, multilevel	Not fully testable; mentoring items are inconsistent and employer-level data are unavailable
H5	Negative	Logit interaction	Exploratory E×Female interaction not significant (OR=0.455, $p=0.254$); inconclusive
H6	Negative	Three-way interaction	Not reliably estimable with current sample and variables
H7	Moderation	Multilevel, shift-share	Not testable: no regional exposure/infrastructure indicators
H8	Partial mediation	Bootstrap	Not established: mediator/outcome scales are incomplete or inconsistent

E. Additional Reporting Shells

Three further shells complete the reporting framework. Table XIV will report difference-in-differences and triple-difference coefficients, Table XV the mediation paths, and Table XVI the regional index ranking with its sensitivity range.

TABLE XIV
DIFFERENCE-IN-DIFFERENCES ESTIMATES (TO BE COMPLETED)

Sample	Coefficient β	SE	N
All workers, junior	Not estimable	—	—
All workers, senior	Not estimable	—	—
Women, junior	Not estimable	—	—
Men, junior	Not estimable	—	—
Junior, Region A	Not estimable	—	—
Junior, Region C	Not estimable	—	—

TABLE XV
MEDIATION PATHS (TO BE COMPLETED)

Mediator	a	b	Indirect ab (95% CI)
Reskilling access (2 available items)	Not reported	Not reported	Not estimated: cross-sectional, incomplete scale
Network access (2 available items)	Not reported	Not reported	Not estimated: cross-sectional, incomplete scale
Digital readiness (1 available item)	Not reported	Not reported	Not estimable from a multi-item scale
Total effect c	Not estimated	—	Not estimated
Direct effect c'	Not estimated	—	Not estimated

TABLE XVI
REGIONAL INDEX RANKING (TO BE COMPLETED)

City	RVRI	Rank	Rank range
Pune	Not estimable	—	—
Chennai	Not estimable	—	—
Bengaluru	Not estimable	—	—
Hyderabad	Not estimable	—	—
Solapur	Not estimable	—	—
Noida	Not estimable	—	—
Gurgaon	Not estimable	—	—
Parbhani	Not estimable	—	—

XI. INTERPRETATION FRAMEWORK AND DISCUSSION

Because results are pending, this section sets out how each possible pattern of evidence will be interpreted, so that conclusions are tied to pre-specified logic and not to post hoc storytelling.

A. Pipeline Effects

If entry-level hiring declines more than senior hiring in exposed occupations (H1) and posted experience requirements rise (H2), the evidence supports the pipeline channel and the concern that the bottom rung is narrowing. If entry-level and senior outcomes move together, the more plausible reading is that AI exposure operates through general demand conditions and not through a career-stage-specific mechanism. If entry-level hiring rises, the productivity-levelling reading of [5] and [49] would gain support, and the discussion would turn to the quality and mentoring content of those new roles.

B. Capability Effects and Mentorship

Support for H3 and H4 would indicate that AI adoption changes how juniors learn even where headcount is stable. Interview evidence will be used to distinguish three explanations: practice tasks removed by automation,

supervisors who redirect time from coaching to checking machine output, and reduced cohort size that undermines structured programmes. A null result for mentorship would need to be examined for measurement error, since mentoring is difficult to observe, before concluding that no erosion exists.

C. Gender and Intersectional Effects

Evidence for H5 and H6 would show that disruption compounds existing disadvantage. The decomposition in Section IX-H will indicate how much of any gap reflects differences in occupation, education and sector, and how much remains unexplained. A large unexplained component would point toward employer practices, network access or algorithmic screening. If women in administrative roles show weaker entry outcomes only in some regions, the finding will be reported as conditional on regional context and not generalized.

D. Regional Heterogeneity

Support for H7 would indicate that place matters beyond individual characteristics, with implications for where intervention should concentrate. If vulnerable regions are not those with the highest exposure scores, the gap between exposure and vulnerability will itself be a key result, because it shows the role of adaptive capacity.

E. Mediators

Partial mediation through reskilling and referral access (H8) would suggest that part of the demographic gap can be narrowed by improving access, whereas a large direct effect would indicate that access alone is insufficient and that employer-side practices need attention.

F. Summary of Interpretive Scenarios

TABLE XVII
INTERPRETATION OF POSSIBLE EVIDENCE PATTERNS

Pattern of evidence	Interpretation	Implication
Junior hiring falls; requirements rise; mentoring falls	Pipeline and capability channels both operate	Redesign entry routes and mentoring
Junior hiring falls; mentoring unchanged	Pipeline channel only	Focus on hiring and access
Junior hiring stable; mentoring and learning fall	Capability channel only	Protect training quality
Effects equal across gender and group	Disruption is not asymmetric by demography	General entry-level policy suffices
Effects vary by region only	Place, not person, drives vulnerability	Regionally targeted policy
No pattern after robustness checks	AI effect on entry not detectable yet	Continue monitoring

The table makes explicit that the design can produce disconfirming evidence for each channel. Reporting null and conflicting patterns with the same prominence as confirming ones is a commitment of the study.

G. Implications for Theory Under Different Outcomes

The results will bear on theory as well as on policy. Confirmation of both channels would support an extension of task-based models in which the *composition of tasks by career*

stage matters for employment as well as the aggregate task content of occupations, and it would suggest incorporating learning externalities into the displacement-reinstatement balance [1]. If only the pipeline channel is confirmed, the implication is that firms adjust headcount but preserve training for those hired, which is consistent with human capital theory in which fewer but better-trained entrants are hired [7]. If neither channel is detected, the most informative interpretation will be that effects are not yet visible or that offsetting reinstatement is occurring, in which case the analysis of new junior tasks created by AI adoption becomes the priority for future work. The study will report which of these patterns best fits the evidence.

H. Joint Display for Integration

Integration of the strands will be documented in a joint display (Table XVIII), which places for each hypothesis the statistical result, the interview evidence and the integrated conclusion side by side. Cells are completed after analysis. Where strands diverge, the display records the divergence and the explanation considered, and the conclusion is stated at the weaker level of confidence.

TABLE XVIII
JOINT DISPLAY SHELL (TO BE COMPLETED)

XII. POLICY IMPLICATION

Hyp.	Quantitative	Qualitative	Integration
H1	Not tested	No interview data supplied	Cannot integrate
H2	Not tested	No interview data supplied	Cannot integrate
H3	Partial exploratory survey; AIS $\alpha=0.861$, but full scales unavailable	No interview data supplied	Quantitative evidence incomplete
H4	Mentoring scale not reliable/complete; multilevel model unavailable	No interview data supplied	Quantitative evidence incomplete
H5	Reduced-model E×Female interaction OR=0.455, $p=0.254$	No interview data supplied	No corroborating qualitative evidence
H6	Not reliably estimated	No interview data supplied	Cannot integrate
H7	Not tested	No interview data supplied	Cannot integrate
H8	Mediation not established; incomplete scales	No interview data supplied	Cannot integrate

NS

A. Candidate Measures

Policy recommendations will be finalized on the basis of the results. The following candidate measures follow from the framework and will be retained, modified or dropped depending on the evidence.

- **Redesign entry pathways.** Apprenticeships, internships and graduate programmes that build judgment alongside

AI tools, so that automation of routine tasks does not eliminate the learning that those tasks provided.

- **Structured mentoring.** Employer mentoring commitments for junior cohorts, with time allocated in supervisors' workloads for coaching, and models in which seniors review AI-assisted work as a teaching exercise.
- **Targeted reskilling.** Training directed at gender-concentrated administrative and service roles, delivered in formats accessible to workers with caregiving responsibilities and limited connectivity.
- **Fair algorithmic hiring.** Audits of automated screening tools for disparate impact by gender and social group, and transparency about the use of automated screening in recruitment.
- **Regional investment.** Digital infrastructure and training capacity in regions identified as vulnerable, and support for economic diversification away from single routine-service specialisms.
- **Curriculum alignment.** Higher-education reforms that emphasize problem framing, communication, and collaboration with AI systems, informed by employer evidence on changing entry requirements.
- **Monitoring.** Routine publication of entry-level hiring indicators disaggregated by gender, social group and region, so that early signs of pipeline stress are visible to policymakers.

TABLE XIX
CANDIDATE POLICY MEASURES, ACTORS AND REQUIRED EVIDENCE

Measure	Lead actor	Evidence required
Apprenticeship redesign	Employers, ministries	H1, H3, H4
Structured mentoring	Employers	H4, H8
Targeted reskilling	Training bodies	H5, H6, H8
Hiring-tool audits	Regulators, employers	H5, H6
Regional investment	State governments	H7
Curriculum alignment	Universities	H2, H3

Each measure is linked to the hypotheses whose support would justify it, so that recommendations are made only where the data provide grounds for them.

B. Implementation Roadmap

The measures above differ in the time they need and in who controls them. A staged roadmap helps to translate evidence into action.

- **Short term (0–12 months).** Employers audit entry-level requirements to remove unjustified experience thresholds, introduce explicit mentoring time in supervisors' duties, and review automated screening for disparate impact. Training bodies pilot short, flexible AI-skills courses aimed at women and first-generation graduates in administrative and service roles.
- **Medium term (1–3 years).** Governments and universities co-design apprenticeship and internship schemes with incentives for employers to retain junior intakes, extend digital infrastructure and training capacity in vulnerable regions, and publish disaggregated early-career indicators.

- **Long term (3–5 years).** Education systems embed judgment, communication and collaboration with AI systems in curricula, labour regulation addresses algorithmic hiring, and regional development plans diversify away from single routine-service specialisms.

Because the empirical evidence will be limited to the study period and regions, recommendations are framed as hypotheses for policy experiments and pilot programmes with built-in evaluation, not as settled prescriptions.

XIII. ETHICS, VALIDITY AND LIMITATIONS

A. Research Ethics

The study will seek approval from the relevant institutional ethics committee before data collection. Participation is voluntary, informed consent is obtained at the start of the survey and before each interview, and respondents may skip any question and withdraw at any time. Sensitive demographic variables are optional and reported only in aggregate for cells above a minimum size to prevent re-identification. Data are stored on encrypted, access-controlled systems, direct identifiers are separated from responses, and qualitative extracts are anonymized. Respondents are not asked to disclose confidential employer information. Incentives, if any, are modest and not contingent on particular answers.

B. Data Protection and Safeguarding

Data handling will comply with applicable data protection law, including India's Digital Personal Data Protection Act, 2023, where the study is conducted in India, and with the requirements of the approving ethics committee. Collection is limited to what is needed for the research questions. Respondents who report distress about employment or discrimination will be shown a list of support services at the end of the survey and interview. Interviewers will be trained to respond sensitively, to pause or stop an interview if a participant becomes distressed, and to avoid promising outcomes that the research cannot deliver. Findings will be reported at levels of aggregation that prevent the identification of individuals or small employers, and a summary of results will be shared with participating institutions and respondents who ask for it.

C. Validity

Internal validity of the secondary-data analysis rests on the parallel-trends assumption, which is examined with pre-trend plots and placebo tests. Construct validity is addressed through scale development, translation checks and factor analysis. Measurement error in the entry-level definition is addressed by using two definitions. External validity is addressed by choosing contrasting regions and weighting to national benchmarks, while recognizing the limits of non-probability sampling. Reliability of qualitative coding is addressed through double coding, agreement statistics and an audit trail of coding decisions.

D. Limitations

Several limitations should be acknowledged. First, the post-adoption window is short, so long-run effects on career progression cannot yet be observed. Second, the survey is cross-

sectional and based on self-reports, which are subject to recall and social-desirability bias; the combination with secondary data is intended to mitigate but cannot remove this. Third, non-probability quota sampling limits formal generalization, and online recruitment may under-represent workers with limited connectivity, which the assisted paper option only partly offsets. Fourth, AI effects are difficult to separate from the economic cycle, rising interest rates and post-pandemic hiring corrections. Fifth, informal employment, which is large in many developing economies, is only partly captured. Sixth, mentorship is an inherently latent construct that is imperfectly measured by survey items. These limitations will be revisited once the data have been analysed.

E. Threats to Validity and Mitigation

TABLE XX
PRINCIPAL THREATS AND MITIGATION

Threat	Source	Mitigation
Selection into sample	Quota and online recruitment	Quotas, weights, assisted paper option
Common-method bias	Single-respondent attitudes	Separated scales, reverse items, single-factor test
Social desirability	Sensitive demographic items	Anonymity, optional items
Confounding by cycle	Rates, post-pandemic hiring	Controls, placebo tests, event studies
Measurement of mentoring	Latent construct	Multi-item scale, interviews, employer module
Researcher bias	Interpretation	Pre-registration, double coding, audit trail

XIV. CONCLUSION AND FUTURE WORK

A. Conclusion

This paper has argued that the debate on AI and employment has paid insufficient attention to the entry point of careers, where the effects on hiring pipelines, junior mentorship and the position of women and other marginalized groups in exposed administrative and service roles may be most consequential and least visible. It has built a research foundation from task-based, human capital, mentoring, intersectionality and regional theory, stated a precise gap, articulated four dimensions of contribution and specified eight falsifiable hypotheses. It has also provided a complete mixed-methods design with a documented survey instrument designed for validation, a pre-specified analysis plan and a reporting framework that separates the design from the findings.

In summary, the paper contributes: (i) a theoretical integration that places career stage, mentoring, intersectional position and region within one framework; (ii) a stylized model that yields falsifiable propositions; (iii) an empirical design that targets the entry point and the learning process, not only displacement; (iv) a documented and replicable survey instrument and analysis plan; and (v) a policy-oriented index and roadmap that link evidence to action.

The immediate next steps are to obtain ethics approval, pilot and finalize the instrument, collect primary and secondary data, and complete Tables X to XIII. Future research can extend the design in four directions: panel follow-up of the same

respondents to observe career trajectories; firm-level studies that link adoption decisions to hiring and training records; comparative replication in other emerging economies; and evaluation of specific interventions such as structured human-plus-AI apprenticeship models. If the pipeline and capability channels proposed here are confirmed, they would imply that the central policy question is not only how to protect workers who are displaced, but also how to ensure that the next generation can still enter, learn and progress.

B. Agenda for Future Research

Completion of the present design opens several lines of work. The most immediate is longitudinal: following the same early-career cohort for several years would reveal whether any early disadvantage in hiring or mentoring persists, compounds or is corrected as workers change jobs. Firm-level research linking AI adoption to recruitment, training and promotion records would test the employer-side mechanisms that the survey can only report indirectly. Experimental audit studies, in which matched applications that vary only in gender or social background are sent to employers using or not using automated screening, would isolate discrimination from qualification differences. Comparative replication in other emerging economies would test how far the findings depend on the Indian institutional setting. Finally, evaluation research on specific interventions, such as structured apprenticeships with AI-assisted tasks and employer mentoring commitments, would convert the descriptive evidence into guidance about what works.

ACKNOWLEDGMENT

[The authors may acknowledge advisors, institutions, funding sources and survey participants here.]

STATEMENT ON RESEARCH TOOLS AND ORIGINALITY

[Authors must disclose any use of writing, translation, analysis or AI-assisted tools in preparing the manuscript, in line with the policies of their institution and the target publisher. All survey data, results and interpretations reported after fieldwork must be the authors' own, and every cited source must be checked against the original before submission.]

REFERENCES

- [1] D. Acemoglu and P. Restrepo, "Automation and new tasks: How technology displaces and reinstates labor," *J. Econ. Perspect.*, vol. 33, no. 2, pp. 3–30, 2019.
- [2] D. H. Autor, F. Levy, and R. J. Murnane, "The skill content of recent technological change: An empirical exploration," *Q. J. Econ.*, vol. 118, no. 4, pp. 1279–1333, 2003.
- [3] T. Eloundou, S. Manning, P. Mishkin, and D. Rock, "GPTs are GPTs: An early look at the labor market impact potential of large language models," arXiv:2303.10130, 2023.
- [4] E. Felten, M. Raj, and R. Seamans, "Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses," *Strateg. Manag. J.*, vol. 42, no. 12, pp. 2195–2217, 2021.
- [5] E. Brynjolfsson, D. Li, and L. R. Raymond, "Generative AI at work," National Bureau of Economic Research, Working Paper 31161, 2023.
- [6] E. Brynjolfsson, B. Chandar, and R. Chen, "Canaries in the coal mine? Six facts about the recent employment effects of artificial intelligence," Stanford Digital Economy Lab, 2025.
- [7] G. S. Becker, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York, NY, USA: Columbia Univ. Press, 1964.
- [8] D. H. Autor, "Why are there still so many jobs? The history and future of workplace automation," *J. Econ. Perspect.*, vol. 29, no. 3, pp. 3–30, 2015.
- [9] D. Acemoglu and D. Autor, "Skills, tasks and technologies: Implications for employment and earnings," in *Handbook of Labor Economics*, vol. 4B, O. Ashenfelter and D. Card, Eds. Amsterdam, The Netherlands: Elsevier, 2011, pp. 1043–1171.
- [10] P. Gmyrek, J. Berg, and D. Bescond, "Generative AI and jobs: A global analysis of potential effects on job quantity and quality," International Labour Organization, ILO Working Paper 96, 2023.
- [11] M. Cazzaniga *et al.*, "Gen-AI: Artificial intelligence and the future of work," International Monetary Fund, Staff Discussion Note SDN/2024/001, 2024.
- [12] Organisation for Economic Co-operation and Development, *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*. Paris, France: OECD Publishing, 2023.
- [13] World Economic Forum, *The Future of Jobs Report 2025*. Geneva, Switzerland: WEF, 2025.
- [14] M. Goos, A. Manning, and A. Salomons, "Explaining job polarization: Routine-biased technological change and offshoring," *Am. Econ. Rev.*, vol. 104, no. 8, pp. 2509–2526, 2014.
- [15] K. E. Kram, "Phases of the mentor relationship," *Acad. Manage. J.*, vol. 26, no. 4, pp. 608–625, 1983.
- [16] T. D. Allen, L. T. Eby, M. L. Poteet, E. Lentz, and L. Lima, "Career benefits associated with mentoring for protégés: A meta-analysis," *J. Appl. Psychol.*, vol. 89, no. 1, pp. 127–136, 2004.
- [17] K. Crenshaw, "Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics," *Univ. Chicago Legal Forum*, vol. 1989, no. 1, pp. 139–167, 1989.
- [18] J. D. Angrist and J.-S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ, USA: Princeton Univ. Press, 2009.
- [19] B. Callaway and P. H. C. Sant'Anna, "Difference-in-differences with multiple time periods," *J. Econometrics*, vol. 225, no. 2, pp. 200–230, 2021.
- [20] L. Sun and S. Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *J. Econometrics*, vol. 225, no. 2, pp. 175–199, 2021.
- [21] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, 8th ed. Andover, U.K.: Cengage, 2019.
- [22] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951.
- [23] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, 2006.
- [24] J. W. Creswell and V. L. Plano Clark, *Designing and Conducting Mixed Methods Research*, 3rd ed. Thousand Oaks, CA, USA: SAGE, 2017.
- [25] P. Goldsmith-Pinkham, I. Sorkin, and H. Swift, "Bartik instruments: What, when, why, and how," *Am. Econ. Rev.*, vol. 110, no. 8, pp. 2586–2624, 2020.
- [26] A. S. Blinder, "Wage discrimination: Reduced form and structural estimates," *J. Hum. Resour.*, vol. 8, no. 4, pp. 436–455, 1973.
- [27] R. Oaxaca, "Male-female wage differentials in urban labor markets," *Int. Econ. Rev.*, vol. 14, no. 3, pp. 693–709, 1973.
- [28] D. Autor and D. Dorn, "The growth of low-skill service jobs and the polarization of the US labor market," *Am. Econ. Rev.*, vol. 103, no. 5, pp. 1553–1597, 2013.
- [29] W. G. Cochran, *Sampling Techniques*, 3rd ed. New York, NY, USA: Wiley, 1977.
- [30] R. Likert, "A technique for the measurement of attitudes," *Arch. Psychol.*, vol. 22, no. 140, pp. 1–55, 1932.
- [31] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum, 1988.
- [32] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. R. Stat. Soc. B*, vol. 57, no. 1, pp. 289–300, 1995.
- [33] M. S. Granovetter, "The strength of weak ties," *Am. J. Sociol.*, vol. 78, no. 6, pp. 1360–1380, 1973.
- [34] Ministry of Statistics and Programme Implementation, Government of India, "Periodic Labour Force Survey (PLFS) annual reports." [Online]. Available: <https://mospi.gov.in>
- [35] International Labour Organization, *World Employment and Social Outlook: Trends 2024*. Geneva, Switzerland: ILO, 2024.
- [36] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Q.*, vol. 27, no. 3, pp. 425–478, 2003.

- [37] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Q.*, vol. 13, no. 3, pp. 319–340, 1989.
- [38] R. B. Kline, *Principles and Practice of Structural Equation Modeling*, 4th ed. New York, NY, USA: Guilford Press, 2016.
- [39] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
- [40] R. M. Baron and D. A. Kenny, "The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations," *J. Pers. Soc. Psychol.*, vol. 51, no. 6, pp. 1173–1182, 1986.
- [41] K. J. Preacher and A. F. Hayes, "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models," *Behav. Res. Methods*, vol. 40, no. 3, pp. 879–891, 2008.
- [42] D. A. Dillman, J. D. Smyth, and L. M. Christian, *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, 4th ed. Hoboken, NJ, USA: Wiley, 2014.
- [43] G. Guest, A. Bunce, and L. Johnson, "How many interviews are enough? An experiment with data saturation and variability," *Field Methods*, vol. 18, no. 1, pp. 59–82, 2006.
- [44] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.
- [45] L. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," *Struct. Equ. Model.*, vol. 6, no. 1, pp. 1–55, 1999.
- [46] P. M. Podsakoff, S. B. MacKenzie, J.-Y. Lee, and N. P. Podsakoff, "Common method biases in behavioral research: A critical review of the literature and recommended remedies," *J. Appl. Psychol.*, vol. 88, no. 5, pp. 879–903, 2003.
- [47] D. J. Deming, "The growing importance of social skills in the labor market," *Q. J. Econ.*, vol. 132, no. 4, pp. 1593–1640, 2017.
- [48] D. Autor, "Applying AI to rebuild middle class jobs," National Bureau of Economic Research, Working Paper 32140, 2024.
- [49] S. Noy and W. Zhang, "Experimental evidence on the productivity effects of generative artificial intelligence," *Science*, vol. 381, no. 6654, pp. 187–192, 2023.

APPENDIX A: SURVEY INSTRUMENT

Title: Artificial Intelligence, Early-Career Work and Learning Survey. Estimated time: 12–15 minutes. Languages: English, Hindi, Marathi. All items except those marked [R] (reverse-scored) are scored in the direction of the label. Items marked * are optional. Participants may skip any question.

Module 1: Early-Career Workers and Recent Graduates

Section 1. Consent and screening

- S1.** I have read the information sheet and agree to take part voluntarily. (Yes / No; end survey if No)
- S2.** I am 18 years of age or older. (Yes / No)
- S3.** How many years have you been in the labour market, including internships? (Less than 1 / 1–3 / 3–5 / More than 5; respondents above 3 years are screened out unless they are job seekers who graduated within the last three years)
- S4.** Current status: Employed full-time / Employed part-time / Intern or trainee / Self-employed or freelance / Unemployed and seeking work / Not seeking work.

Section 2. Demographics

- D1.** Age in completed years.
- D2.** Gender: Woman / Man / Non-binary or other / Prefer not to say.
- D3.** Social category: General / Other Backward Classes / Scheduled Caste / Scheduled Tribe / Other / Prefer not to say.*
- D4.** Are you the first person in your family to complete a university degree? (Yes / No / Not applicable)
- D5.** Highest qualification completed: Secondary / Higher secondary / Diploma / Bachelor's / Master's / Other.
- D6.** Field of study: Commerce and management / Engineering and technology / Arts and humanities / Science / Other.
- D7.** Institution type: Public / Private / Other.
- D8.** City or district of current residence: Region A / B / C / D / Other.
- D9.** Where did you spend most of your childhood? Large city / Smaller city or town / Rural area.

D10. Monthly household income band (five bands, lowest to highest; Prefer not to say).

D11. Do you have a disability that affects work? (Yes / No / Prefer not to say)*

D12. Weekly hours spent on unpaid caregiving or household responsibilities: 0 / 1–10 / 11–20 / 21–35 / More than 35.

Section 3. Employment profile (employed respondents)

E1. Sector: Information technology and services / Banking and finance / Business-process outsourcing and customer service / Manufacturing / Public sector / Education and health / Other.

E2. Occupation group: Administrative and clerical / Customer service / Software and data / Finance and accounting / Human resources / Marketing and content / Other.

E3. Employer size: Fewer than 50 / 50–249 / 250–999 / 1,000 or more.

E4. Contract type: Permanent / Fixed-term / Contract or agency / Informal.

E5. Gross monthly earnings (five bands; Prefer not to say).

E6. Months in the current role.

Section 4. AI use at work

A1. Which AI tools do you use for work? (Select all: chat assistants, writing tools, coding assistants, translation tools, data-analysis tools, scheduling tools, none)

A2. How often do you use AI tools for work? Never / Rarely / Monthly / Weekly / Daily.

A3. Approximately what share of your tasks involves AI assistance? None / 1–25% / 26–50% / 51–75% / 76–100%.

A4. Does your employer have a formal AI policy or provide approved AI tools? Yes / No / Unsure.

A5. Did your employer introduce AI tools into your team after 2022? Yes / No / Unsure.

Section 5. AI task substitution (AIS) and routine learning tasks (RLT)

Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree.

AIS1. Tasks that were once given to people at my level are now done by AI tools.

AIS2. AI tools complete in minutes work that would take me hours.

AIS3. My employer expects me to produce work that previously needed more junior staff.

AIS4. Fewer people at my level have been hired in my team since AI tools were adopted.

AIS5. I am concerned that AI could perform much of my current work.

AIS6. AI tools have improved the quality of my output.

RLT1. I regularly perform routine tasks that help me learn how the work is done.

RLT2. [R] The number of practice-type tasks given to me has fallen since AI tools were introduced.

RLT3. I understand why tasks are done as they are, not only how to operate the tool.

RLT4. My skills are growing as fast as I expected when I started.

Section 6. Hiring-pipeline experience

H1. In the last 12 months, how many jobs did you apply for? (0 / 1–10 / 11–30 / 31–60 / More than 60)

H2. How many interviews did you attend? (0 / 1–3 / 4–6 / 7–10 / More than 10)

H3. How many offers did you receive? (0 / 1 / 2 / 3 or more)

H4. Months from starting your search to your first job relevant to your qualification. (Numeric; or 'not yet')

H5. Did you complete an internship or apprenticeship before your first job? (Yes / No)

Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree.

- H6.** Postings that I applied to asked for more experience than I expected for entry-level roles.
H7. Entry-level openings in my field have become harder to find in recent years.
H8. I was asked to complete automated screening (online tests, video assessments or automated résumé screening). (Yes / No / Unsure)
H9. I believe automated screening treated me fairly.

Section 7. Mentoring and learning (MEN)

Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree.

- MEN1.** I have a designated mentor or supervisor responsible for my development.
MEN2. I receive feedback on my work at least weekly.
MEN3. My supervisor explains the reasoning behind decisions.
MEN4. I have opportunities to observe experienced colleagues at work.
MEN5. My employer provides structured training for new staff.
MEN6. [R] My supervisor spends less time coaching me than before AI tools were introduced.
MEN7. I can ask senior colleagues for help without hesitation.
MEN8. Senior colleagues review my work and teach me how to improve it.

Section 8. Networks (NET) and reskilling access (RSK)

Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree.

- NET1.** I know people in my field who could recommend me for jobs.
NET2. Someone in my network referred me to an opportunity in the last 12 months.
NET3. I have access to alumni or professional communities in my field.
NET4. People in my network hold senior positions in my field.
RSK1. Affordable training in AI-related skills is available to me.
RSK2. My employer pays for training or gives me time for it.
RSK3. I have reliable internet and a suitable device for online learning.
RSK4. I have the time to reskill outside work.

Section 9. Career outcomes (COU) and open responses

Scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree.

- COU1.** I feel secure in my current job or job prospects over the next two years.
COU2. I am confident I can progress to a more senior role.
COU3. I am satisfied with my pay relative to my qualifications.
COU4. [R] I intend to leave my occupation within two years.
COU5. I am treated fairly in hiring and promotion regardless of gender or background.
COU6. I am optimistic about my career.

- O1.** Describe one way your work or job search has changed since AI tools became common. (Open text)
O2. What support would help you most at this stage of your career? (Open text)

- O3.** Are you willing to be contacted for a follow-up interview? (Yes / No; contact details collected separately)

Module 2: Managers and Human-Resource Professionals

- M1.** Organization size, sector and region (categories as in E1, E3 and D8).
M2. Your role: Hiring manager / HR professional / Team lead / Owner or director.
M3. Level of AI adoption in your organization: None / Pilot / Some teams / Widespread.
Compared with 2022 (Much lower / Lower / Same / Higher / Much higher):
M4. Our hiring of entry-level staff is...
M5. The experience we require for entry-level roles is...
M6. Time senior staff spend coaching juniors is...

- M7.** Structured training for new staff is...
M8. Which tasks once given to juniors are now done with AI assistance? (Select all: drafting, data entry, scheduling, customer queries, coding, reporting, research, none)
M9. Main reasons for any change in junior hiring: Cost / Productivity of AI tools / Skill mismatch / Economic uncertainty / Other.
M10. Over the next 12 months we plan to hire entry-level staff: Fewer / About the same / More.
M11. Practices used to maintain junior learning: Job rotation / Shadowing / Formal programme / Buddy system / None.
M12. Do you use automated screening in recruitment? Yes / No / Unsure.
M13. What would help your organization continue to hire and develop junior staff? (Open text)

APPENDIX B: INTERVIEW GUIDE

Semi-structured; 40–60 minutes; recorded with consent. Probes are used as needed.

Early-career workers

- W1.** Walk me through how you found your first job. What helped and what got in the way?
W2. Which tasks did you do in your first months? How have these changed with AI tools?
W3. Who taught you how to do your work? How often, and in what form?
W4. Have you experienced automated screening or changing experience requirements? How did it feel?
W5. How do your gender, background or location shape your options? Can you give an example?
W6. What training would help, and what stops you from getting it?

Supervisors, mentors and hiring managers

- M1.** How has your team's use of AI changed what junior staff do?
M2. How has your hiring of juniors changed, and why?
M3. How do you now build junior judgment and experience? What has been lost or gained?
M4. How do you decide on experience requirements and screening methods?
M5. What would make it worthwhile to keep investing in juniors?
Training providers and policymakers
T1. Which groups seem most affected at entry level in your region, and what evidence do you see?
T2. What programmes work, and what barriers (cost, connectivity, caregiving) limit take-up?
T3. What data or support would help your organization respond?

APPENDIX C: ANALYSIS CODEBOOK (KEY VARIABLES)

TABLE XXI
KEY VARIABLES AND CODING

Variable	Source items	Coding
Entry-level	S3, E6	1 if ≤ 3 years in labour market
Female	D2	1 = woman, 0 = man
First-generation	D4	1 = Yes
Disadvantaged group	D3	1 = SC, ST or OBC
AI intensity	A2, A3	Standardized composite
AIS, RLT, MEN, NET, RSK, COU	Scale items	Mean of items (R-items reversed)
Timely first job	H4	1 if ≤ 6 months
Inflated requirement	H6	1 if score ≥ 4
Region	D8	Four-level factor
Intersection	D2, D4, D3	Female × first-gen × group

APPENDIX D: PARTICIPANT INFORMATION AND CONSENT

Study title: Artificial Intelligence, Early-Career Work and Learning.
Researcher: [Name, institution, contact]. **Ethics approval:** [Committee, reference number].

Purpose. This study examines how artificial intelligence tools are changing entry-level hiring, learning at work and career prospects, and whether the changes differ by gender, background and city.

What participation involves. You will complete an online or paper questionnaire taking about 12–15 minutes. A small number of participants may be invited to a separate interview of 40–60 minutes, which is optional and requires separate consent.

Voluntary participation. Taking part is voluntary. You may skip any question and may stop at any time without any consequence. Demographic questions marked optional can be left blank.

Confidentiality. Your answers will be stored securely and analysed only in combination with those of other participants. Your name and employer will not be reported. Contact details, if you choose to give them for a follow-up interview, will be stored separately from your answers and deleted after the study.

Risks and benefits. Some questions concern your job prospects and treatment at work, and you may find these uncomfortable. A list of support services is provided at the end. There is no direct payment, but your answers will contribute to evidence that may inform training and employment policy. [State any incentive here.]

Questions. Contact [researcher email] or the ethics committee at [contact]. By selecting *Yes* to item S1, you confirm that you have read this information and agree to participate.

APPENDIX E: PILOT AND COGNITIVE-INTERVIEW CHECKLIST

TABLE XXII
PILOT EVALUATION CRITERIA

Criterion	Threshold	Action if not met
Median completion time	≤ 15 min	Shorten or split sections
Item non-response	≤ 5%	Reword or make optional
Floor or ceiling effects	≤ 85%	Revise item or scale
Corrected item-total correlation	≥ 0.30	Drop or reword
Cronbach's alpha per scale	≥ 0.70	Revise items
Comprehension probes	No repeated confusion	Clarify term or example
Translation equivalence	Back-translation matches	Reconcile with translators

Cognitive-interview probes: *What does the term 'AI tools' mean to you here? In your own words, what is this question asking? Was any answer option missing? Did any question feel uncomfortable or unclear?*

APPENDIX F: ANALYSIS SYNTAX OUTLINE (PYTHON)

Illustrative outline for the survey models. Variable names follow Appendix C. It is a template to adapt, not executed output.

```
import pandas as pd, numpy as np
import statsmodels.formula.api as smf

df = pd.read_csv('survey_clean.csv')

def cronbach_alpha(items):
    items = items.dropna()
    k = items.shape[1]
    var_sum = items.var(axis=0, ddof=1).sum()
    total_var = items.sum(axis=1).var(ddof=1)
    return k/(k-1) * (1 - var_sum/total_var)

# Scale scores with reverse-coded items
df['MEN6r'] = 6 - df['MEN6']
```

```
men = ['MEN1', 'MEN2', 'MEN3', 'MEN4', 'MEN5', 'MEN6r', 'MEN7', 'MEN8']
df['MEN'] = df[men].mean(axis=1)
print('alpha MEN:', cronbach_alpha(df[men]))

# Logit with exposure x female interaction (H5)
m1 = smf.logit('timely_job ~ ai_int*female + first_gen
              + disadv + age + C(sectors) + C(region)',
              data=df).fit(cov_type='cluster',
                          cov_kws={'groups':
                                df['employer_id']})
print(m1.summary())

# OLS for mentoring intensity (H4)
m2 = smf.ols('MEN ~ ai_int + female + first_gen +
             C(region)',
             data=df).fit(cov_type='HC3')

# Bootstrap indirect effect through reskilling access (H8)
def indirect(d):
    a = smf.ols('RSK ~ female + age',
    d).fit().params['female']
    b = smf.ols('COU ~ RSK + female + age',
    d).fit().params['RSK']
    return a*b
boot = [indirect(df.sample(len(df), replace=True)) for _ in
range(5000)]
print(np.percentile(boot, [2.5, 97.5]))
```

APPENDIX G: INDEX CONSTRUCTION

Indicators are standardized (z-scores) and combined with weights summing to one. Equal weights are the baseline; sensitivity analysis varies weights.

TABLE XXIII
COMPONENTS OF THE REGIONAL VULNERABILITY AND RESILIENCE INDEX

Indicator	Measure	Direction
AI exposure	Shift-share exposure score	Higher = more vulnerable
Junior-role share	Share of employment at entry level	Higher = more vulnerable
Gender concentration	Female share in exposed roles	Higher = more vulnerable
Demographic risk	Share first-gen, disadvantaged	Higher = more vulnerable
Digital infrastructure	Broadband and device access	Higher = more resilient
Training capacity	Institutions per 1,000 youth	Higher = more resilient
Diversification	Industry concentration index	Higher = more resilient
Mentoring practice	Mean MEN score from survey	Higher = more resilient

APPENDIX H: GLOSSARY

TABLE XXIV
GLOSSARY OF ABBREVIATIONS

Term	Meaning
AI	Artificial intelligence
AIS	AI task substitution scale
COU	Career outcomes scale
DiD	Difference-in-differences
LFS	Labour force survey
MEN	Mentoring intensity scale
NET	Network and referral access scale
PLFS	Periodic Labour Force Survey
RLT	Routine learning tasks scale
RSK	Reskilling access scale
RVRI	Regional Vulnerability and Resilience Index

Term	Meaning
SEM	Structural equation model

APPENDIX I: INTERVIEW CODING FRAME

TABLE XXV
A PRIORI INTERVIEW CODING FRAME

Theme	Definition	Example indicator
Task change	Shift in tasks assigned to juniors	Practice tasks removed
Coaching time	Senior time spent teaching	Less review time
Hiring logic	Employer reasoning on junior hiring	Cost, risk, requirements
Screening	Experience of automated selection	Online tests, filters
Network access	Referrals and sponsorship	Who recommended
Reskilling	Barriers and enablers to training	Cost, time, connectivity
Gendered roles	Role allocation by gender	Clerical stereotyping
Regional context	Local opportunity structure	Few alternative employers
Coping	Strategies after setbacks	Switching, further study
Emergent	Themes not anticipated	Coded inductively

APPENDIX J: ITEM-TO-HYPOTHESIS MAP

TABLE XXVI
SURVEY ITEMS MAPPED TO HYPOTHESES

Hyp.	Survey items
H1	H1–H9 (hiring experience); M4, M5, M10
H2	H6, H7; M5
H3	AIS1–AIS6, RLT1–RLT4, A2, A3
H4	MEN1–MEN8, A4, A5; M6, M7
H5	D2, E2, AIS, COU1–COU6, H1–H4
H6	D2, D3, D4, COU, H1–H4
H7	D8, D9, E1, RSK3, RSK1
H8	NET1–NET4, RSK1–RSK4, D2, COU

APPENDIX K: RECRUITMENT MESSAGES

Initial invitation. Subject: *Short survey on early-career work and AI tools*. Dear [Name/Participant], I am a researcher at [Institution] studying how artificial intelligence tools are changing entry-level hiring, learning at work and career prospects. If you graduated in the last three years or are in your first three years of work, I would be grateful if you would complete a 12–15 minute confidential questionnaire: [link]. Participation is voluntary, and you may skip any question. The study has been approved by [Committee]. Thank you for considering this request. [Researcher name and contact].

Reminder (after 7 days). Dear [Name/Participant], this is a gentle reminder about the survey on early-career work and AI tools. If you have already responded, thank you; if not, the link remains open until [date]: [link]. You may ignore this message if you prefer not to take part.

Message to institutional gatekeepers. Dear [Name], we are conducting a study approved by [Committee] on how AI is changing early-career opportunities. We would appreciate your help in circulating the invitation to your students, alumni or members. We will not request names or contact details from you, and participants respond directly and anonymously. We would be pleased to share a summary of the findings with your institution.

Interview invitation. Dear [Name], thank you for indicating that you might be willing to take part in a follow-up conversation. The interview takes 40–60 minutes, may be held online or in person, and can be conducted in English, Hindi or Marathi. You may decline any

question and withdraw at any time. If you would like to proceed, please reply with times that suit you.

APPENDIX L: SURVEY ROUTING

TABLE XXVII
SURVEY ROUTING AND SKIP LOGIC

If response is...	Then...
S1 = No or S2 = No	End survey with thanks
S3 above 3 years and S4 not seeking	Screen out, thank respondent
S4 = Unemployed and seeking	Skip Sections 3 and 4 and E1–E6
S4 = Employed, intern or freelance	Show all sections
A2 = Never	Skip A3 and set AIS items to optional
H8 = No or Unsure	Skip H9
Respondent selects manager role	Route to Module 2 only
O3 = Yes	Show separate contact form

APPENDIX M: ETHICS REVIEW CHECKLIST

TABLE XXVIII
ITEMS TO DOCUMENT BEFORE FIELDWORK

Item	Evidence to attach
Institutional approval	Approval letter and reference number
Participant information sheet	Final text in all three languages
Consent procedure	Online consent item; paper consent form
Data protection plan	Storage, access, retention, destruction
Risk assessment	Distress protocol; support-service list
Gatekeeper permissions	Written permission from institutions
Incentive policy	Amount, eligibility, delivery method
Data-sharing plan	Anonymization rules; repository
Conflict-of-interest statement	Funding sources and affiliations
Pre-registration	Registry entry and date

APPENDIX N: REPORTING STANDARDS

The final report will follow recognized reporting guidance: published survey-reporting standards for response rates, sampling and weighting; a consolidated checklist for interview and focus-group studies; and mixed-methods reporting guidance for integration.

TABLE XXIX
REPORTING COMMITMENTS

Element	Commitment
Response rate	Report by region and channel
Weighting	State benchmarks and weights used
Missing data	Report extent and handling
Exploratory analyses	Label as exploratory
Null results	Report with equal prominence
Interview methods	Report sampling, saturation, coders
Integration	Joint display of strands
Limitations	State per hypothesis

APPENDIX O: PLANNED FIGURES

TABLE XXX
FIGURES PLANNED FOR THE RESULTS SECTION

Figure	Purpose
Event-study plot, junior versus senior	Pre-trends and dynamics (H1)
Share of entry-level postings over time	Pipeline change (H2)
Distribution of required experience	Requirement inflation (H2)
Forest plot of odds ratios by group	Gender and intersection (H5, H6)

Figure	Purpose
Predicted probabilities by exposure	Magnitude of effects
Path diagram with estimates	SEM and mediation (H3, H4, H8)
Regional map of RVRI	Regional vulnerability (H7)
Joint display table	Integration of strands

APPENDIX P: WORKED EXAMPLE OF INDEX ARITHMETIC

The index of Section IX-G is computed by orienting every indicator so that higher values mean greater vulnerability, standardizing, and taking the weighted sum. Table XXXI illustrates the arithmetic with *hypothetical* values chosen only to show the calculation; they are not data and do not describe any real region. Resilience indicators enter with reversed sign, and four equally weighted indicators are used for brevity.

TABLE XXXI
ARITHMETIC ILLUSTRATION OF THE INDEX (HYPOTHETICAL
VALUES)

Indicator	Weight	Region X z	Region Y z
AI exposure	0.25	+0.80	−0.60
Junior-role share	0.25	+0.30	−0.10
Digital infrastructure (reversed)	0.25	+0.50	−0.70
Diversification (reversed)	0.25	+0.20	−0.40
RVRI	1.00	+0.45	−0.45

In this illustration Region X, with higher exposure and weaker adaptive capacity, receives a positive score and ranks as more vulnerable than Region Y. Sensitivity analysis would redraw the weights from a distribution around the baseline and report how often the ranking of X and Y reverses.