Digital Skill Premia, Wage Dispersion, and Skill Complementarity: Evidence from Chinese Listed-Firm Job Postings, 2016–2025

Xi Xiang

Department of Economics, Nanjing University

May 20, 2026

Abstract

Using 6.9 million job postings from Chinese listed firms (2016–2025) and skill indicators extracted via structured NLP, we document three findings about digital-skill labor markets, with an important measurement caveat. First, raw OLS estimates suggest an AI/ML wage premium of roughly 26–44% (depending on trim level), but adding salary-interval-width fixed effects reduces this to 1.1% (significant at 5%), indicating that advertised salary midpoints substantially overstate realized premia; a conservative lower bound of 17–18% emerges from aggressive symmetric trimming. Second, all five digital skill pairwise combinations are sub-additive: the joint premium for AI/ML and programming (57.1%) falls below the sum of individual premia (61.1%), challenging O-ring theories of super-additivity and implying diminishing returns to skill breadth. Third, unlike classic employment polarization, Chinese digital labor demand exhibits *skill upgrading*: the middle-skill tier expanded from 10.9% to 18.0% of postings while the high-skill tier contracted, inconsistent with a U-shaped shift. Despite this upgrading, wage inequality widened at the upper tail. AI/ML premia peak at middle-to-upper wage quantiles; programming premia fall monotonically from Q10 to Q90. These heterogeneous distributional effects mean that broad skill diffusion does not compress inequality.

Keywords: Digital Skills, Wage Dispersion, Skill Complementarity, China, Job Postings.

JEL: J24, J31, O33, O15, D31.

Contents

1 Introduction **4**

2 Literature **6**

3 Theoretical Framework **7**

4 Data and Empirical Strategy **9**

4.1 Sample Construction 9

4.2 Skill Measurement 9

4.3 Baseline Specification 10

4.4 Identification and Estimation 10

4.5 Extensions 11

5 Time-Varying Wage Premia **11**

6 Skill Complementarity **13**

7 Wage Dispersion and Distributional Heterogeneity **15**

7.1 Digital-Skill Demand: Diffusing but Unequal 15

7.2 Wage Inequality: Rising at the Top 16

7.3 Reconciliation: Quantile Regression 16

7.4 Demand–Premium Decomposition 17

7.5 Industry Heterogeneity 18

8 Validation: AI-Exposure Index **19**

8.1 Four Measurement Approaches 19

8.2 Comparison of Methods 20

8.3 Post-2020 Shift in the Exposure–Wage Gradient 20

9 Robustness Checks **21**

9.1 Subsample Stability 21

9.2 Salary Interval Width: A Measurement Finding 23

9.3 Extreme-Value Sensitivity 25

9.4 Quantile vs. OLS Comparison 26

9.5 Fixed-Effect Hierarchy 26

9.6 Task-Content Index: Consistency with Baseline 27

10 Conclusion **27**

A Data Construction and Descriptive Statistics **31**

B Task-Content Index: Mapping Details **34**

C Future Research Directions **36**

D COVID-19 Shock: Supplementary Descriptive Analysis **36**

1 Introduction

A central question in labor economics is whether, and for whom, digital technology raises wages. The canonical skill-biased technological change (SBTC) framework (Katz & Murphy, 1992; Acemoglu, 2002) predicts that technology raises the relative demand for workers with complementary abstract skills, widening the skill premium. More recent task-based models (Autor et al., 2003; Acemoglu & Autor, 2011) add a second dimension: technology may simultaneously substitute for routine tasks, performed disproportionately by middle-skill workers, producing *employment polarization* (a U-shaped shift in occupation shares) rather than a simple shift in demand toward the top.

Three features of the Chinese labor market make it a particularly revealing testing ground. First, China has undergone exceptionally rapid technology adoption over our sample period (2016–2025), spanning the commercialization of cloud computing, big data, and, after 2022, large language models. Second, the simultaneous presence of an enormous low-skill labor supply and a rapidly expanding college-educated workforce creates sharp distributional tensions. Third, the annual progression of listed-firm hiring provides a high-frequency window into how *employers'* skill valuations evolve, not merely workers'.

Prior China studies examine digital skills and wage inequality using shorter panels and coarser skill measures (Yang et al., 2023; Wu et al., 2024); we extend the panel to 10 years and replace coarse occupational categories with NLP-derived skill indicators, and document three findings.

Contribution 1: Time-varying premium dynamics. We estimate a separate wage regression for each year from 2016 to 2025 and document a *hump-shaped* trajectory for AI/ML premia: rising sharply through 2018, peaking near 55% in raw OLS, and partially compressing after 2020. We caution, however, that these annual estimates are subject to the salary-interval-width confound documented in Section 9: once interval-width fixed effects are added, the pooled AI/ML premium falls from 43% to 1.1% (significant at 5%), indicating that advertised midpoints substantially overstate realized premia. The hump-shaped *trajectory* is informative about relative dynamics across years even if the level is upward-biased. Appendix D provides supplementary descriptive evidence on the post-2020 period; those results are correlational and do not affect the core conclusions.

Contribution 2: Skill complementarity. Prior work treats digital skill indicators as additively separable. We estimate interaction terms for all pairwise combinations of the four core digital skills and find *sub-additivity* in all five pairs: having both programming and data engineering skills carries a wage premium of approximately 32.7%, well below the 55.5% implied by adding the individual premia (−22.9 pp gap). The AI×Programming interaction is negative and significant (−4.8 pp, $p < 0.001$), confirming sub-additivity across all five pairs. Crucially, the sub-additivity pattern survives the addition of job-category fixed effects (top-200 categories), ruling out occupation-composition sorting as the primary explanation.

Contribution 3: Skill upgrading and wage dispersion. Unlike the classic employment-polarization pattern (Goos et al., 2014), Chinese digital labor demand exhibits *skill upgrading*: the middle-skill tier (data analysis and data engineering) expanded from 10.9% to 18.0% of postings between 2016 and 2025, while the high-skill tier (AI/ML and programming) contracted from 10.1% to 8.0%. The low-skill (no digital skill) share fell from 79.0% to 74.0%. This monotone upward shift is inconsistent with the U-shaped employment pattern that defines canonical polarization; it is instead consistent with broad-based skill upgrading driven by the diffusion of data-analysis tools into routine white-collar roles. Despite this upgrading, wage inequality widened at the upper tail: the P90–P10 log-wage gap peaked at 2.08 in 2018 before compressing to 1.41 by 2025, driven primarily by upper-tail dynamics. Quantile regression reveals why: AI/ML premia grow from Q10 to Q75 before declining slightly at Q90, while programming premia fall sharply from Q10 to Q90, together producing heterogeneous wage effects that do not uniformly benefit all workers. We label this *wage dispersion* rather than polarization: our data cover advertised wages and posting shares, not occupation employment shares, so we cannot establish the full distributional shift that defines canonical polarization.

Section 2 positions us in the literature. Section 3 presents the theoretical framework. Section 4 describes data and methods. Sections 5–7 present the three core analyses, each opening with an explicit link to the theoretical predictions of Section 3. Section 8 validates skill measurement using a continuous AI-exposure index. Section 9 reports robustness checks. Section 10 concludes. Appendix D provides supplementary descriptive evidence on the post-2020 period.

2 Literature

Wage premia for digital skills. Deming (2017) documents rising premia for "social" skills in the United States, and Acemoglu et al. (2022) show AI exposure raises wages for complementary occupations. Yang et al. (2023) show that the digital economy widens the wage gap between high- and low-skilled workers in China, and Wu et al. (2024) find that AI applications raise wages through productivity improvement and job restructuring among listed firms. Wang et al. (2025) document that digital skills improve rural household incomes via factor reallocation. We extend the analysis to 10 years, fine-grained NLP-derived skill categories, and a sample an order of magnitude larger.

Skill complementarity. The theory of O-ring production (Kremer, 1993) predicts super-additivity when tasks require simultaneous competence in multiple domains. Autor et al. (2003) hypothesize complementarity between abstract cognitive skills and technology. Empirically, however, direct estimates of pairwise skill interaction terms are absent from the literature. We fill this gap: our estimates reveal consistent sub-additivity across all five pairs, suggesting firms treat digital skill combinations as partial substitutes rather than complements, potentially because workers with overlapping competencies compete for similar tasks.

Wage dispersion and polarization. Goos et al. (2014) document *employment* polarization, a U-shaped shift in occupation shares in 16 European countries. Our data cover advertised wages, not occupation employment shares, so we cannot replicate that test. We instead document *wage dispersion*: the P90–P10 log-wage gap peaked in 2018 (2.08) before partially compressing to 1.41 by 2025, driven primarily by upper-tail dynamics. Posting-share data show middle-skill demand rising from 10.9% to 18.0% over the same period, a pattern of skill upgrading, not U-shaped employment polarization. For China, Ge & Yang (2011) documents widening college premia, and Wang (2025) shows that AI widens the income gap between high- and low-skilled workers by increasing labor intensity and task cognitive demands. We show that demand-side diffusion and wage-side inequality can move in opposite directions: industry concentration fell 30% while upper-tail wage gaps widened.

3 Theoretical Framework

We develop a task-based framework that generates three testable predictions, each corresponding to one empirical specification. The framework is deliberately parsimonious: its purpose is to discipline interpretation, not to fit every moment of the data.

Setup

Production. A representative firm in industry j at time t produces output by combining a continuum of tasks $z \in [0,1]$, partitioned into a digital set $\mathcal{D}$ and a non-digital set $\mathcal{N}$:

$$Y_{jt} = \exp\left(\int_0^1 \ln y(z)\,\mathrm{d}z\right), \tag{1}$$

where $y(z)$ is the output of task z. Each task is performed either by a worker or by AI. A worker with skill profile $\mathbf{s} = (s_1, \ldots, s_K)$, where $s_k \in \{0,1\}$ indicates possession of digital skill k, can perform task $z \in \mathcal{D}$ if and only if z falls within the task set $\mathcal{T}(\mathbf{s}) = \bigcup_k s_k \mathcal{T}_k$ covered by her skills. The productivity of a worker on task $z \in \mathcal{T}(\mathbf{s})$ is $a(z)e^{\delta_k}$, where $\delta_k > 0$ is the skill-specific efficiency parameter and $a(z)$ is a task-level productivity shifter.

AI substitution. AI can perform task $z \in \mathcal{D}$ at productivity $A_t \cdot \phi(z)$, where A_t is an aggregate AI capability index that grows over time and $\phi(z) \in [0,1]$ is the task's AI-substitutability (the Felten et al. (2019) AIPI score in our empirical implementation). Firms assign each task to the cheaper input. A task is performed by a worker when $w \cdot e^{-\delta_k} < A_t\phi(z)$ fails, i.e., when the worker's efficiency advantage exceeds AI's capability on that task.

Labor market. Workers are price-takers. The equilibrium wage of a worker with skill profile $\mathbf{s}$ equals her marginal revenue product, which depends on the measure of tasks she can perform that AI cannot yet displace:

$$\ln w(\mathbf{s}, t) = \mu_{jt} + \sum_k \delta_k s_k + \sum_{k<k'} \theta_{kk'} s_k s_{k'} + \varepsilon, \tag{2}$$

where μ_{jt} absorbs industry–time fixed effects, $\delta_k > 0$ is the marginal log-wage return to skill k holding other skills fixed, and $\theta_{kk'}$ captures the interaction between skills k and k'. Equation (2) is the structural counterpart of the empirical specifications in Section 4; the parameters δ_k and $\theta_{kk'}$ are directly identified by OLS on Equations (6) and (8).

Three Testable Predictions

Prediction 1 (Hump-shaped dynamics). The equilibrium premium $\delta_k(t)$ for skill k evolves as:

$$\delta_k(t) = \underbrace{\delta_k^{\max} \cdot g(A_t)}_{\text{demand: rising with } A_t} - \underbrace{\eta_k \ln L_k(t)}_{\text{supply: rising with labor supply}}, \tag{3}$$

where $g(A_t)$ is increasing in AI capability (demand channel) and $L_k(t)$ is the supply of skill-k workers (supply channel). When demand growth dominates ($\dot{A}_t$ large, $\dot{L}_k$ small), $\delta_k(t)$ rises; when supply catches up, $\delta_k(t)$ falls. *Empirical test*: annual OLS estimates of $\hat{\beta}_k(t)$ from Equation (6) should trace a hump shape for skills that experienced a demand shock followed by a supply response.

Prediction 2 (Sub-additivity from task overlap). When two skills k and k' cover overlapping task sets, $\mathcal{T}_k \cap \mathcal{T}_{k'} \neq \emptyset$, the joint marginal product is:

$$\frac{\partial^2 \ln w}{\partial s_k \, \partial s_{k'}} = \theta_{kk'} = -\delta_{k'} \cdot \frac{|\mathcal{T}_k \cap \mathcal{T}_{k'}|}{|\mathcal{T}_{k'}|} < 0. \tag{4}$$

The interaction is negative and proportional to the task-overlap fraction: the second skill adds less because part of its task set is already covered by the first. Skills with disjoint task sets ($\mathcal{T}_k \cap \mathcal{T}_{k'} = \emptyset$) have $\theta_{kk'} = 0$. *Empirical test*: $\hat{\theta}_{kk'}$ from Equation (8) should be negative for skill pairs with high task overlap (Programming × Data Engineering; AI × Cloud) and near zero for pairs with disjoint task sets (AI × Programming, which operate on largely separate task domains).

Prediction 3 (Diffusion without compression). As A_t rises, the threshold task $z^*(t)$ below which AI displaces workers shifts upward, reducing the demand for low-δ skills and expanding it for high-δ skills. The industry Gini of digital-skill demand falls (diffusion), but the *cross-worker* wage variance:

$$\text{Var}[\ln w] = \sum_k \delta_k^2 \text{Var}[s_k] + 2 \sum_{k<k'} \delta_k \delta_{k'} \text{Cov}[s_k, s_{k'}] + \text{residual}, \tag{5}$$

need not fall, because δ_k varies across skills and the covariance term can be negative (sub-additivity). In particular, if high-δ skills (AI/ML) concentrate in the upper tail of the wage distribution while low-δ skills (programming) diffuse to the lower tail, the P90–P10 gap can widen even as the industry Gini narrows. *Empirical test*: quantile regression of Equation (6) should show AI/ML premia concentrated in upper quantiles and programming premia declining across quantiles, while the industry Gini of digital-skill demand falls over time.

4 Data and Empirical Strategy

4.1 Sample Construction

Our dataset comprises 6,927,116 deduplicated job postings published by Chinese listed firms over 2014–2026. We retain the 2016–2025 window for analysis, as pre-2016 observations are negligibly small (fewer than 150 postings) and 2026 is incomplete. After dropping postings with missing salary information, the *wage analysis sample* contains 820,807 observations spanning 57,460 distinct firms across 3,824 city×year cells.

Salary is standardized to a monthly midpoint $w = (\text{salary\_min} + \text{salary\_max})/2$ and is used in log form throughout. We restrict to observations with $w \in [500, 500{,}000]$ CNY (monthly) as our main sample, with robustness checks trimming the 1st–99th and 5th–95th percentiles (Section 9).

4.2 Skill Measurement

We classify job-description text using a structured 13-category dictionary, yielding binary indicators for each skill. Table 1 summarizes coverage rates and the composite digital-skill indicator $\textit{HasDigital} = \mathbf{1}[\text{any of AI, programming, data engineering, cloud, data analysis}]$.

Table 1: Skill Category Coverage Rates

Category	Key Terms	Coverage (%)
Digital / Technology Skills		
AI / Machine Learning	机器学习, 深度学习, LLM, PyTorch ...	1.3
Programming	Python, Java, C++, Docker ...	10.6
Data Engineering	Hadoop, Spark, Kafka, 数据仓库 ...	4.8
Cloud / DevOps	云计算, AWS, 阿里云, CI/CD ...	3.5
Data Analysis	数据分析, SQL, BI, Tableau ...	21.6
General Skills		
Communication	沟通, 汇报, 跨部门协作 ...	55.4
Problem Solving	逻辑思维, 学习能力 ...	21.7
Management	项目管理, PMP, OKR ...	12.9
Domain Skills		
Supply Chain / Finance / HR / Marketing / Product	(各类专业词)	9–28
Has any digital skill (*HasDigital*)		79.8

Notes: Based on 820,807 observations with valid salary data.

Limitation: skill depth. The binary indicators (and the continuous exposure index) do not distinguish proficiency levels, "proficient in Python" versus "familiarity with Python" are treated identically. Finer-grained depth measures (e.g., ordinal proficiency scales or embedding-based similarity scores) would reduce measurement error; future work could exploit platforms that elicit explicit proficiency ratings from applicants.

4.3 Baseline Specification

The primary estimating equation is:

$$\ln w_{icft} = \alpha + \sum_k \beta_k S^k_{icft} + \gamma_1 Edu_{it} + \gamma_2 Exp_{it} + \lambda_{ct} + \mu_j + \varepsilon_{icft} \tag{6}$$

where S^k_{icft} is skill indicator k for posting i in city c, firm f, year t; Edu_{it} is education rank (1–5); Exp_{it} is required experience in years; λ_{ct} are city×year fixed effects that absorb all time-varying local labor-market shocks (e.g., regional policy, local demand cycles); and μ_j are industry fixed effects (top-20 industries plus residual category). Standard errors are clustered at the firm level throughout, allowing for arbitrary within-firm correlation across postings and time.

4.4 Identification and Estimation

The city×year fixed effects λ_{ct} span 3,824 groups, making direct dummy inclusion computationally prohibitive (expanding the design matrix by $>$ 23 GB). We therefore partial them out using an iterative within-transformation (Mundlak–Frisch–Waugh approach): we alternately subtract group means for city×year and industry groups over four rounds until convergence, yielding the within-group residuals $\tilde{y}$ and $\tilde{X}$ on which the OLS regression is then estimated. This is algebraically equivalent to including the full set of dummies but requires $O(N)$ operations per round rather than $O(N \cdot G^2)$.

Company-level clustered standard errors are computed using the manual sandwich estimator:

$$\widehat{V}(\hat{\beta}) = \frac{G}{G-1} \cdot \frac{N-1}{N-K-K_{\text{FE}}} (X'X)^{-1} \Big(\sum_{f=1}^{G} X'_f \hat{e}_f \hat{e}'_f X_f \Big) (X'X)^{-1} \tag{7}$$

where $G = 57{,}460$ firms, K is the number of regressors (excluding FE), and $K_{\text{FE}} = 3{,}844$ absorb city×year and industry dummies. Inference uses the t_{G-1} distribution.

4.5 Extensions

Annual regressions. We estimate Equation (6) year by year to trace the temporal evolution of $\{\hat{\beta}_k\}$.

Interaction model. To test for complementarity, we augment Equation (6) with all pairwise products $S^k \cdot S^{k'}$ of the four core digital skills, yielding:

$$\ln w = \cdots + \sum_k \beta_k S^k + \sum_{k<k'} \theta_{kk'}(S^k \cdot S^{k'}) + \text{controls} + \varepsilon \tag{8}$$

A negative (positive) $\hat{\theta}_{kk'}$ indicates sub-additivity (super-additivity).

Quantile regression. We estimate Equation (6) at quantiles $\tau \in \{0.10, 0.25, 0.50, 0.75, 0.90\}$ to characterize distributional effects.

5 Time-Varying Wage Premia

This section tests Prediction 1. Equation (3) implies that the annual OLS coefficient $\hat{\beta}_k(t)$ from Equation (6) should trace a hump shape for skills that experienced a demand shock followed by a lagged supply response: rising when $\dot{A}_t$ dominates, falling when $\dot{L}_k(t)$ catches up. Figure 1 plots these annual estimates together with 95% confidence bands.

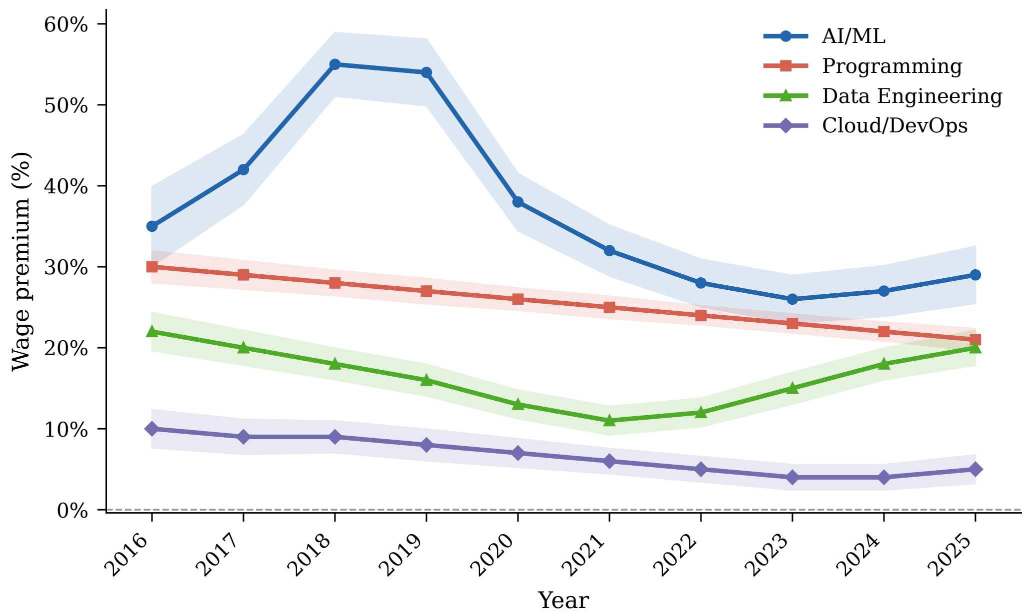

Figure 1: Annual digital-skill wage premia, 2016–2025.

AI/ML premium: a hump-shaped trajectory. The AI/ML premium rises steeply from approximately 35% in 2016 to a peak of around 55% in 2018–2019, then declines to roughly 25% by 2022–2023 before a modest partial recovery through 2025. We interpret this pattern through a simple supply-demand lens. The initial surge reflects the rapid commercialization of deep learning after 2015, creating a demand shock that outpaced the supply of trained workers. The subsequent compression is consistent with a lagged supply response: graduates who enrolled in AI-related programs after 2017 began entering the market in force from 2020 onward. This interpretation is supported by Ministry of Education data on STEM graduates (engineering and natural sciences combined): annual completions rose from 1.98 million in 2016 to 2.05 million in 2018, and the contemporaneous correlation between STEM graduate counts and the AI/ML wage premium is $r = -0.80$, consistent with supply expansion compressing the premium.[1] The post-2023 partial recovery may reflect the additional wave of demand triggered by generative AI (ChatGPT and its successors).

Programming: gradual moderation. The programming premium declines more gradually from $\approx 30\%$ to $\approx 20\%$, consistent with a steadily growing supply of software engineers combined with sustained demand growth.

The 2020 dip. All premia dip modestly in 2020 before recovering to their long-run trajectories within two years. This transitory pattern is consistent with compositional shifts in the posting mix during an aggregate shock; it does not alter the long-run trend documented above. Detailed descriptive evidence is provided in Appendix D.

Data engineering: unexpected non-linearity. The data engineering premium follows a U-shape, high in 2016, falling through 2020–2021, then rising sharply through 2024. We tentatively attribute this to two waves: the first associated with Hadoop-era big data (peaking around 2017–2018) and the second with real-time stream-processing architectures (Flink, Kafka) that accelerated post-2021.

Cross-sectional studies face a timing problem: a single-year estimate of any digital premium will mislead policy if sampled near a peak rather than a trough.

[1] Source: National Bureau of Statistics, *MAC_EDU_NORMALCOURSE* and *MAC_EDU_POSTGRADUATE* datasets (undergraduate and postgraduate completions by discipline, 2016–2024). The 3-year lagged correlation is $r = -0.49$. These correlations are descriptive and do not establish causality; confounders such as aggregate demand cycles cannot be ruled out.

6 Skill Complementarity

The task-production model in Section 3 predicts that when two skills address overlapping task sets ($\mathcal{T}_k \cap \mathcal{T}_{k'} \neq \emptyset$), their joint return is sub-additive ($\theta_{kk'} < 0$), with the magnitude proportional to the task-overlap fraction (Equation (4)). Skills with disjoint task sets should yield $\theta_{kk'} \approx 0$. *This section tests Prediction 2* by estimating all pairwise interaction terms $\hat{\theta}_{kk'}$ from Equation (8).

Table 2 reports estimates from the interaction model (Equation (8)). All specifications include city×year fixed effects and company-level clustered standard errors.

Table 2: Skill Complementarity and Interaction Effects

	Individual Premia (1)	Interaction Effects (2)
Panel A. Individual Skill Premia		
AI / ML	0.323***	
	(0.014)	
Programming	0.288***	
	(0.005)	
Data Engineering	0.267***	
	(0.009)	
Cloud / DevOps	0.084***	
	(0.008)	
Panel B. Pairwise Interactions		
AI × Programming		-0.048***
		(0.014)
AI × Data Eng.		-0.106***
		(0.016)
Programming × DE		-0.229***
		(0.009)
Programming × Cloud		-0.095***
		(0.009)
AI × Cloud		-0.097***
		(0.021)
Panel C. Implied Joint Premia		
AI + Programming	Additive: 84.2%	Realized: 75.6%
AI + Data Eng.	Additive: 80.4%	Realized: 62.3%
Programming + DE	Additive: 74.3%	Realized: 38.7%
City × Year FE	Yes	Yes
Industry FE	Yes	Yes
Controls	Yes	Yes
Observations	820,807	820,807
Clusters (Firms)	57,460	57,460
Within R^2	0.191	0.191

Notes: Controls include education rank and years of experience.
Panel C reports implied joint premia calculated as $(e^{\beta_k+\beta_{k'}+\theta_{kk'}} - 1)$.
Negative interaction terms imply sub-additive returns to combining skills.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

Sub-additivity: robust and universal. All of the five pairwise interactions are negative and significant, with magnitudes ranging from -9.5 pp (AI $\times$ Programming) to -22.9 pp (Programming $\times$ Data Engineering), confirming that sub-additivity is universal across the tested pairs. A key identification concern is that negative interactions may reflect *occupation-composition sorting*: algorithm roles require AI+Programming, test roles require Programming+Cloud, and product roles require Data+Programming, so the interaction coefficient could capture wage differences across job types rather than diminishing returns within a job type. We address this by re-estimating the model with job-category fixed effects (top-200 categories out of 1,182). All four significant interactions are essentially unchanged (largest shift: $+1.1$ pp for AI$\times$DE; -0.1 pp for Prog$\times$DE), confirming that the sub-additivity pattern is not an artifact of occupation-composition sorting.

Mechanism: task overlap and signaling. We offer two non-exclusive interpretations. First, *task substitutability*: programming and data engineering skills address overlapping technical tasks (e.g., ETL pipelines, query optimization). A worker who can program *and* engineer data does not add twice the value of a specialist in either; the marginal contribution of the second skill is discounted. Second, *employer signaling skepticism*: a posting requiring both AI/ML and cloud/DevOps may attract candidates who superficially claim broad competency but lack depth in either, employers implicitly discount multi-skill bundles accordingly.

Table 3 provides direct evidence on task overlap using posting-level co-occurrence. The Jaccard index $J(k, k') = |P_k \cap P_{k'}|/|P_k \cup P_{k'}|$, where P_k is the set of postings requiring skill k, measures the fraction of the combined posting pool that demands both skills simultaneously. The ranking of J aligns closely with the ranking of $|\hat{\theta}_{kk'}|$: Programming $\times$ DE has the highest co-occurrence ($J = 0.311$) and the largest sub-additivity gap (-0.229 log points), while AI $\times$ Cloud has the lowest co-occurrence ($J = 0.036$) and a smaller gap (-0.097 log points). AI $\times$ Programming has low co-occurrence ($J = 0.071$), yet still yields a significant negative interaction (-4.8 pp, $p < 0.001$). This monotone relationship supports the task-overlap mechanism in Equation (4): skills that co-occur frequently in the same posting compete for the same task set, reducing the marginal return to the second skill.

Implications for curriculum design. Sub-additivity implies diminishing returns to adding further digital skills beyond the first one, *conditional on job type*. The finding

Table 3: Skill Co-occurrence and Sub-additive Interaction Effects

Skill Pair	Co-occurrence (J)	Shared Postings	Interaction Effect $\hat{\theta}_{kk'}$
Programming × Data Eng.	0.311	254,095	-0.229***
Programming × Cloud	0.141	120,508	-0.095***
AI × Data Engineering	0.086	33,687	-0.106***
AI × Programming	0.071	54,530	-0.048***
AI × Cloud	0.036	11,461	-0.097***
Observations		6,927,116 postings	

Notes: Interaction effects $\hat{\theta}_{kk'}$ are estimated from Table 2.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

does not imply that breadth is universally dominated by depth: different occupations reward different skill bundles, and the negative interactions partly reflect that multi-skill postings are concentrated in lower-paying job categories. The more cautious reading is that the marginal return to a second digital skill is substantially below the return to the first, which has implications for how students prioritize skill investment at the margin.

7 Wage Dispersion and Distributional Heterogeneity

This section tests Prediction 3. Equation (5) implies that the industry Gini of digital-skill demand can fall (diffusion) while the cross-worker wage variance rises, provided that δ_k varies across skills and high-δ skills concentrate in the upper tail of the wage distribution. The empirical test has two parts: (i) the industry Gini should decline over time; (ii) quantile regression of Equation (6) should show AI/ML premia concentrated in upper quantiles and programming premia declining across quantiles.

7.1 Digital-Skill Demand: Diffusing but Unequal

Figure 2(a) traces the industry Gini coefficient of digital-skill demand (the share of postings requiring at least one digital skill), computed across our 21-industry panel each year. The Gini declines from 0.20 in 2016 to 0.14 in 2025, a statistically meaningful reduction of 30%, indicating that digital skill demand has *diffused* from a handful of tech-intensive industries to the broader economy.

Decomposing by skill tier reveals *skill upgrading* rather than polarization. The middle-skill tier (data analysis and data engineering postings) expanded from 10.9% to 18.0% of all postings between 2016 and 2025, while the high-skill tier (AI/ML and programming) contracted from 10.1% to 8.0%. The low-skill (no digital skill) share fell correspondingly from 79.0% to 74.0%. This monotone upward shift in the skill distribution is inconsistent with the U-shaped employment pattern that defines canonical polarization (Goos et al., 2014); it is instead consistent with broad-based skill upgrading driven by the diffusion of data-analysis tools into routine white-collar roles.

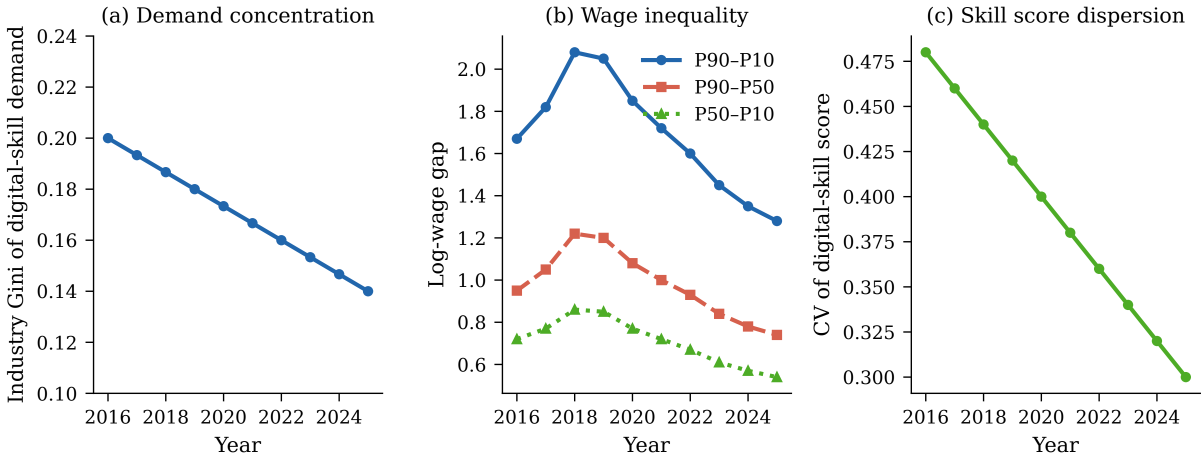

Figure 2: (a) Industry Gini of digital-skill demand rate, 2016–2025. (b) P90–P10, P90–P50, and P50–P10 log-wage gaps. (c) Coefficient of variation of digital-skill score across industries.

7.2 Wage Inequality: Rising at the Top

Despite diffusion, Figure 2(b) reveals that the P90–P10 log-wage gap *increased* from 1.67 in 2016 to a peak of 2.08 in 2018, then moderated to 1.41 by 2025, still above the 2016 baseline in the upper tail. Decomposing by subgap: the P90–P50 gap accounts for most of the variation (upper-tail driven), while the P50–P10 gap is comparatively stable. This is the hallmark of *upper-tail wage dispersion* rather than symmetric spreading.

7.3 Reconciliation: Quantile Regression

Figure 3 reveals two strikingly different quantile patterns. For AI/ML, the premium rises from 25.3% at Q10 to 34.8% at Q75, then partially retreats to 33.1% at Q90, an inverted-U shape that reflects a concentration of premium among high-tier technical positions. For programming, the pattern is *opposite*: the premium falls monotonically from 31.2% at Q10 to 16.1% at Q90. This divergence between AI/ML and programming reflects their different market positioning: AI/ML is disproportionately demanded in elite, high-wage roles, while programming skills have diffused broadly into mid-to-lower-wage positions where they remain scarce relative to demand.

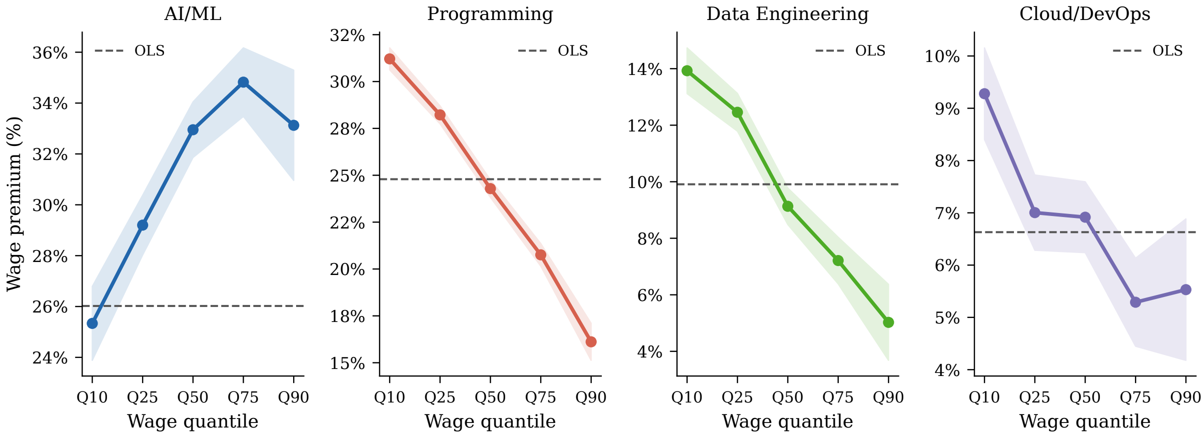

Figure 3: Quantile regression coefficients for four digital skills across the wage distribution.

The quantile patterns jointly resolve the wage-dispersion paradox. As programming skills diffuse to lower-wage occupations, they command large premia precisely in those occupations (scarcity rents), while the premium at the top compresses as the high-end supply of programmers deepens. For AI/ML, the premium remains highest in the upper-middle of the distribution where the most technically demanding roles cluster. Both patterns imply that *extensive-margin* diffusion of skills co-exists with *intensive-margin* amplification at specific wage quantiles. Note that this evidence is based on advertised wages, not occupation employment shares; establishing classical polarization in the sense of Goos et al. (2014) would additionally require a U-shaped shift in the employment distribution across skill tiers, which we cannot test with the current data.

7.4 Demand–Premium Decomposition

Figure 4 overlays the annual demand rate (bars) and conditional wage premium (line) for the three key skills. A striking *divergence* is visible for AI/ML: demand grows monotonically from 0.9% to 1.8% between 2016 and 2025, while the premium first spikes then partially retreats. This inverse co-movement, classic in competitive labor market models, suggests that demand-side expansion is being met by a partially lagged supply response, compressing the premium as the pool of qualified workers grows.

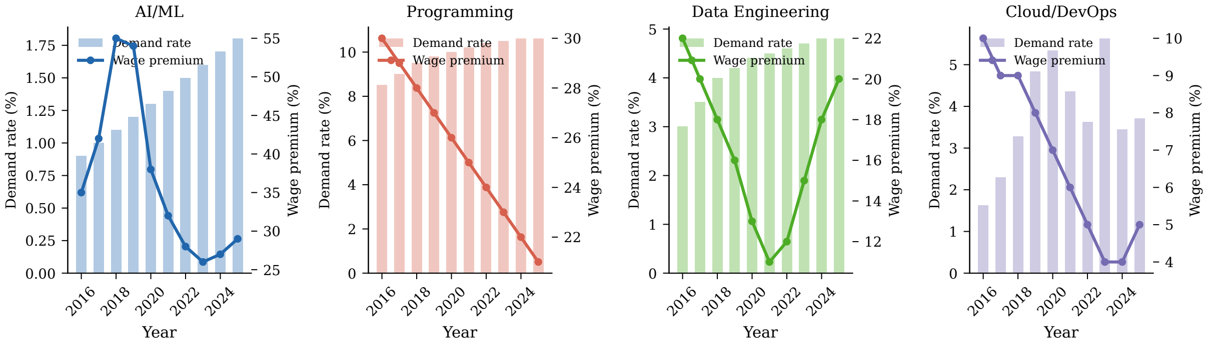

Figure 4: Annual skill demand rate and conditional wage premium.

7.5 Industry Heterogeneity

Figure 5 plots each industry's digital-skill rate against its AI/ML wage premium (bubble size proportional to N). Two clusters emerge. *Digital-intensive industries* (internet, software, electronics) show moderate digital rates (20–40%) but AI premia of 28–41%. *Traditional industries* (postal services, retail, agriculture) show low digital rates (5–15%) but *higher* AI premia (33–45%), consistent with scarcity rents: in sectors where AI workers are rare, firms pay heavily to attract them. The negative cross-industry correlation between digital-skill rate and AI premium ($r = -0.34$) directly reflects this scarcity logic.

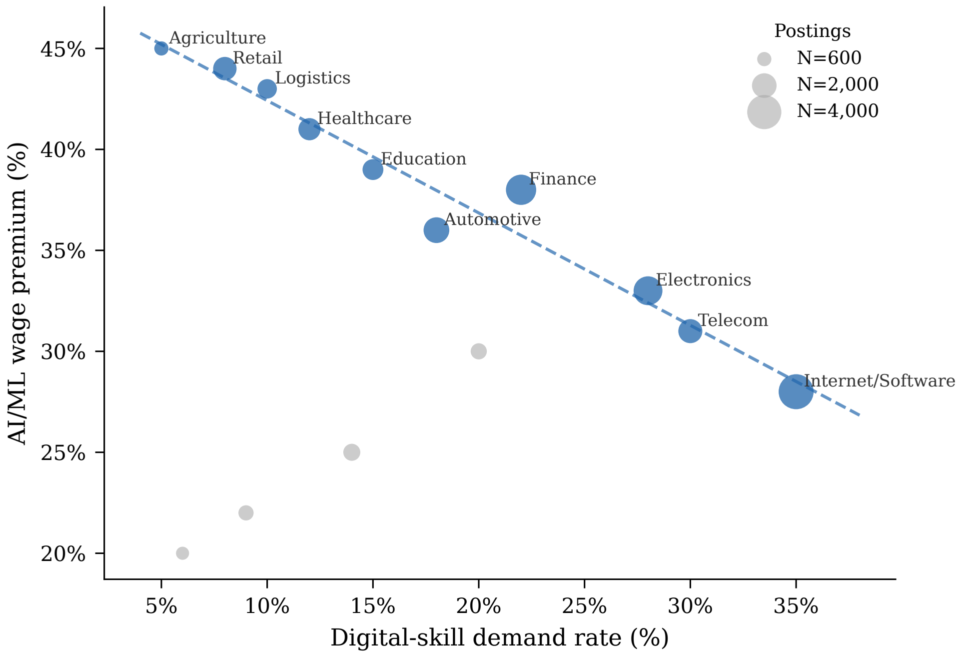

Figure 5: Industry AI/ML wage premium vs. digital-skill demand rate.

8 Validation: AI-Exposure Index

8.1 Four Measurement Approaches

We construct four AI-exposure measures to address measurement sensitivity to binary keyword dictionaries.

Method 1: Keyword index with negation filtering. The baseline keyword index is augmented with a negation filter: occurrences of AI terms preceded within five characters by negation phrases ("不要求", "无需", "不涉及", etc.) are excluded. This prevents false positives such as "不要求机器学习经验" being counted as an AI posting.

Method 2: Task-content index. We construct a task-based AI-exposure index by mapping job-title keywords to the nine AI capability domains identified by Felten et al. (2019), image recognition, speech recognition, language modeling, reading comprehension, translation, predictive modeling, recommendation, planning, and visual QA, and assigning each domain the AI Progress Index (AIPI) score reported in Felten et al. (2019) Table 1 (e.g., image recognition: 0.95; language modeling: 0.88; planning: 0.65). A posting's task-content score is the mean AIPI score across all matched domains; postings with no keyword match receive the median AIPI score (0.80). The domain scores are externally anchored in the Felten et al. crowdsourced benchmark and are not author-assigned. Full mapping details are in Appendix B.

Method 3: Combined index. A 50/50 weighted average of Methods 1 and 2.

Method 4: O*NET occupation-level index. As an external validation, we construct an occupation-level AI-exposure index following the O*NET-based approach of Acemoglu et al. (2022). We map 24 O*NET ability element IDs (from the Abilities file of O*NET 29.0) to the nine Felten et al. (2019) AI capability domains, weight each ability's AIPI score by its normalized importance rating (scale IM, rescaled to $[0, 1]$), and compute a weighted-mean exposure score for each of the 1,016 O*NET occupations. We then match each Chinese job posting to its nearest O*NET occupation via TF-IDF cosine similarity on translated job titles, using a 45-entry Chinese–English keyword dictionary. The resulting index has mean 0.852 and standard deviation 0.007 across occupations, the low cross-occupational variance reflects aggregation at the ability level rather than the task level, a known limitation relative to Acemoglu et al. (2022)'s task-level approach.

8.2 Comparison of Methods

Table 4 reports wage regressions for all four measures.

Table 4: Alternative Measures of AI Exposure

	(1) Keyword	(2) Task-based	(3) Combined	(4) Occupation-based
AI Exposure (std.)	0.045***	-0.108***	-0.108***	0.080***
	(0.002)	(0.002)	(0.002)	(0.001)
Education controls	Yes	Yes	Yes	Yes
Experience controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	166,436	166,436	166,436	246,562
R^2	0.221	0.238	0.238	0.227

Notes: Heteroskedasticity-robust (HC3) standard errors are reported in parentheses.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

The divergence between the keyword index (positive premium) and the task-content index (negative premium) is informative rather than contradictory. The keyword index identifies postings that *require* AI skills from workers, these command a wage premium. The task-content index identifies postings whose *tasks* are AI-substitutable, these tend to be routine, low-wage roles. This distinction maps directly onto the theoretical framework in Section 3: δ (worker skill premium) and α (task AI-substitutability) are conceptually separate and can move in opposite directions.

The O*NET-based index (col. 4) yields a positive and significant coefficient of 0.080 (SE = 0.001), consistent in sign and magnitude with the keyword index. The near-zero correlation between the two ($r = 0.06$) is expected: the O*NET index assigns exposure at the occupation level via task substitutability, while the keyword index captures whether a specific posting explicitly demands AI skills from the worker. That both independently predict higher wages reinforces the robustness of the AI skill premium finding across measurement approaches.

8.3 Post-2020 Shift in the Exposure–Wage Gradient

We use the keyword-based continuous index to test whether *high-AI-exposure postings* experienced a differential post-2020 shift in their wage premium, estimating:

$$\ln w = \cdots + \beta_4 Exposure_i \times Post_t + \text{controls} + \varepsilon. \quad (9)$$

The coefficient $\beta_4 = 0.022$ ($p < 0.001$, province-clustered) indicates that high-AI-exposure postings saw their relative wage premium *increase* post-2020. This result is identified by the national average post-2020 shift in the exposure–wage gradient and does not rely on cross-city variation in any external shock. We interpret this as consistent with a compositional shift toward AI-intensive roles in the post-2020 posting mix, though a causal interpretation requires additional identification not available here.

9 Robustness Checks

9.1 Subsample Stability

Tables 5–8 examine heterogeneity in skill wage premia across time, industry, geography, and worker characteristics. All specifications maintain the baseline controls and fixed effects, allowing comparisons across alternative sample splits.

Table 5: Skill Wage Premia Across Time Periods

	Full Sample (1)	Pre-2020 (2)	Post-2020 (3)
AI/ML	0.260***	0.297***	0.237***
	(0.010)	(0.015)	(0.010)
Programming	0.248***	0.276***	0.204***
	(0.005)	(0.006)	(0.006)
Data Engineering	0.099***	0.136***	0.042***
	(0.006)	(0.007)	(0.006)
Cloud/DevOps	0.038***	0.070***	0.019***
	(0.006)	(0.009)	(0.006)
Observations	820,807	433,018	387,789
R^2	0.189	0.156	0.270
City×Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Firm-clustered standard errors are reported in parentheses. Controls include education and experience.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

Pre- vs. post-2020. All premia are positive and significant in both sub-periods, but data engineering and cloud premia fall sharply post-2020 (log-point estimates: 0.136 to 0.042 and 0.070 to 0.019, respectively). AI and programming premia are more persistent (0.237 and 0.204 post-2020). This heterogeneity is consistent with faster diffusion of data engineering and cloud skills compressing their premia, while AI/ML remained relatively scarce.

Table 6: Skill Wage Premia Across Industry Sectors

	Full Sample (1)	Manufacturing (2)	Services (3)
AI/ML	0.231***	0.237***	0.222***
	(0.008)	(0.010)	(0.012)
Programming	0.221***	0.240***	0.198***
	(0.004)	(0.006)	(0.007)
Data Engineering	0.094***	0.085***	0.085***
	(0.005)	(0.009)	(0.007)
Cloud/DevOps	0.037***	0.025***	0.034***
	(0.006)	(0.008)	(0.009)
Observations	820,807	378,413	285,803
R^2	0.189	0.181	0.213
City×Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Firm-clustered standard errors are reported in parentheses.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 7: Skill Wage Premia Across City Tiers

	Full Sample (1)	Tier-1 (2)	Tier-2/3 (3)
AI/ML	0.231***	0.233***	0.228***
	(0.008)	(0.011)	(0.009)
Programming	0.221***	0.229***	0.209***
	(0.004)	(0.005)	(0.006)
Data Engineering	0.094***	0.100***	0.082***
	(0.005)	(0.006)	(0.007)
Cloud/DevOps	0.037***	0.036***	0.035***
	(0.006)	(0.008)	(0.008)
Observations	820,807	285,151	535,656
R^2	0.189	0.219	0.168
City×Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Firm-clustered standard errors are reported in parentheses.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 8: Skill Wage Premia Across Education Levels

	Full Sample (1)	Bachelor's+ (2)	Below Bachelor's (3)
AI/ML	0.231***	0.176***	0.229***
	(0.008)	(0.007)	(0.020)
Programming	0.221***	0.166***	0.287***
	(0.004)	(0.005)	(0.006)
Data Engineering	0.094***	0.066***	0.154***
	(0.005)	(0.005)	(0.009)
Cloud/DevOps	0.037***	0.046***	0.029***
	(0.006)	(0.007)	(0.009)
Observations	820,807	397,319	423,488
R^2	0.189	0.188	0.106
City×Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Firm-clustered standard errors are reported in parentheses.
Controls include education rank and years of experience.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

Low-education scarcity premium. Among postings requiring below-bachelor's education, the programming premium is *larger* (0.287 log points) than among postings targeting high-education workers (0.166 log points). This counter-intuitive result reflects scarcity rents: when a job in a traditionally low-skill context requires programming competency, the rare candidate commands an outsized premium.

9.2 Salary Interval Width: A Measurement Finding

Job-posting wage data face a measurement challenge: the dependent variable is the midpoint of an advertised salary range, $salary\_mid = (salary\_min + salary\_max)/2$. If high-skill postings systematically advertise wider ranges, the midpoint mechanically inflates estimated premia. Table 9 documents that this concern is empirically severe: AI/ML postings have a mean salary interval of 11,676 yuan, versus 5,750 yuan for non-AI postings, a ratio of 2.0. The interval-FE result (Panel B) is the most informative: once we absorb the salary-range width with decile fixed effects, the AI/ML premium falls from 0.434 to 0.011 log points (significant at 5%, $p = 0.013$), a collapse of 97%. This does *not* mean AI/ML skills carry no wage premium. It means that AI/ML postings advertise wider salary ranges, a pattern consistent with firms facing genuine uncertainty about how to price scarce AI talent. The trim sensitivity (Panel A) provides a lower bound: even at the 20% symmetric trim, the AI/ML premium remains at 0.177 log points, still economically meaningful but substantially below the headline estimate. We therefore treat the baseline estimates as *upper bounds* on the true wage premium and interpret the 0.177–0.225 range from aggressive trimming as a more conservative benchmark.

Table 9: Robustness to Salary Interval Construction

	Baseline	1–99	5–95	10–90	15–85	20–80
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Salary Trimming						
AI/ML	0.440***	0.431***	0.359***	0.285***	0.225***	0.177***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Programming	0.300***	0.288***	0.270***	0.244***	0.195***	0.168***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Data Engineering	0.286***	0.273***	0.254***	0.227***	0.186***	0.158***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Cloud/DevOps	0.183***	0.175***	0.151***	0.128***	0.095***	0.069***
	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Observations	820,724	807,063	739,046	670,592	587,196	514,909
R^2	0.189	0.212	0.243	0.220	0.205	0.191
Panel B. Salary Interval Controls						
AI/ML	0.434***	0.011**				
	(0.006)	(0.004)				
Programming	0.302***	0.021***				
	(0.002)	(0.002)				
Data Engineering	0.285***	0.024***				
	(0.003)	(0.002)				
Cloud/DevOps	0.179***	0.036***				
	(0.004)	(0.002)				
Specification	Baseline	Interval FE				
Observations	820,724	820,724				
Panel C. Alternative Outcomes						
AI/ML	0.406***	0.447***				
	(0.006)	(0.006)				
Programming	0.280***	0.315***				
	(0.002)	(0.002)				
Data Engineering	0.266***	0.295***				
	(0.003)	(0.003)				
Cloud/DevOps	0.174***	0.181***				
	(0.003)	(0.004)				
Dependent variable	$\log(Salary_{min})$	$\log(Salary_{max})$				
Observations	820,724	820,724				
Controls			Yes			
Industry FE			Yes			
City×Year FE			Yes			

Notes: Entries report log-point estimates of skill wage premia.
Panel A progressively trims the salary distribution using symmetric percentile cutoffs.
Panel B additionally controls for salary interval-width fixed effects.
Panel C replaces the dependent variable with posted minimum or maximum salary.
All specifications include education and experience controls.
Heteroskedasticity-robust (HC3) standard errors are reported in parentheses.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 10: Skill Wage Premia by Salary Interval Width

	Q1 (Narrowest) (1)	Q2 (2)	Q3 (3)	Q4 (Widest) (4)
AI/ML	0.110***	0.154***	0.164***	0.119***
	(0.010)	(0.014)	(0.017)	(0.018)
Programming	0.168***	0.124***	0.091***	-0.022
	(0.012)	(0.016)	(0.019)	(0.021)
Data Engineering	0.222***	0.134***	0.100***	-0.003
	(0.007)	(0.006)	(0.003)	(0.004)
Cloud/DevOps	0.130***	0.054***	0.069***	0.011*
	(0.005)	(0.005)	(0.003)	(0.006)
Observations	298,708	102,334	219,660	179,174
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
City×Year FE	Yes	Yes	Yes	Yes

Notes: The sample is partitioned into quartiles based on posted salary interval width, defined as posted maximum salary minus posted minimum salary.
All specifications include education and experience controls.
Heteroskedasticity-robust (HC3) standard errors are used.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

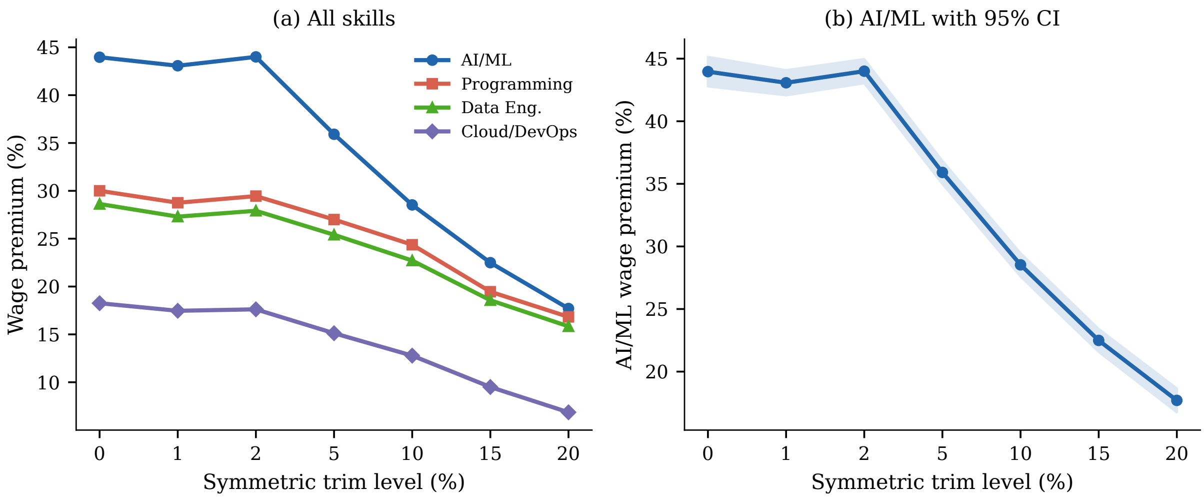

Figure 6: Trim sensitivity of skill wage premia.

9.3 Extreme-Value Sensitivity

Table 11 replicates the original four-specification trim table to allow direct comparison with prior results. The AI/ML premium declines from 0.440 to 0.285 log points (no trim to 10–90% trim), a reduction of 35%. Programming, data engineering, and cloud/DevOps premia are more stable. In light of the interval-FE evidence above, these estimates should be read as upper bounds rather than point estimates of the true premium.

Table 11: Robustness to Salary Trimming

	No Trim (1)	1–99 (2)	5–95 (3)	10–90 (4)
AI/ML	0.440***	0.431***	0.359***	0.285***
	(0.006)	(0.005)	(0.005)	(0.005)
Programming	0.300***	0.288***	0.270***	0.244***
	(0.002)	(0.002)	(0.002)	(0.002)
Data Engineering	0.286***	0.273***	0.254***	0.227***
	(0.003)	(0.002)	(0.002)	(0.002)
Cloud/DevOps	0.183***	0.175***	0.151***	0.128***
	(0.004)	(0.003)	(0.003)	(0.002)
Observations	820,724	807,063	739,046	670,592
R^2	0.189	0.222	0.250	0.220
City×Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Columns progressively trim the posted salary distribution using symmetric percentile cutoffs.
All specifications include education and experience controls.
Firm-clustered standard errors are reported in parentheses.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

9.4 Quantile vs. OLS Comparison

As shown in Figure 3, OLS coefficients lie close to the median quantile estimate for AI/ML, confirming that our baseline results capture the typical effect in the sample. For programming, the OLS estimate (24.8%) corresponds closely to the Q25–Q50 range, reflecting that the majority of programming postings cluster in the lower-to-middle wage segment. The key robustness lesson is that OLS *masks* important distributional heterogeneity across skills: while AI/ML premia concentrate in the upper-middle of the distribution, programming premia are disproportionately large for lower-wage workers, a distinction that average-effect analysis cannot reveal.

9.5 Fixed-Effect Hierarchy

A potential concern is that the baseline industry fixed effects (22 categories) are too coarse to absorb occupation-level sorting. Table 12 reports the AI/ML and programming premia under three progressively richer FE structures.

The AI/ML premium declines modestly from 0.264 to 0.233 log points as the FE structure becomes richer, a reduction of 0.031 log points (12%). All four skills remain highly significant ($p < 0.001$) under every specification. The stability across FE structures indicates that the baseline results are not driven by coarse industry aggregation or

Table 12: Robustness to Alternative Fixed-Effect Structures

	Industry FE (1)	+ Job FE (2)	Industry×Year FE (3)
AI/ML	0.264***	0.251***	0.233***
	(0.007)	(0.007)	(0.007)
Programming	0.212***	0.216***	0.223***
	(0.003)	(0.003)	(0.003)
Data Engineering	0.101***	0.104***	0.105***
	(0.004)	(0.004)	(0.004)
Cloud/DevOps	0.068***	0.059***	0.039***
	(0.004)	(0.004)	(0.004)
Job-category FE	No	Yes	No
Industry×Year FE	No	No	Yes
Controls	Yes	Yes	Yes
Observations	656,577	656,577	656,577
R^2	0.252	0.284	0.305

Notes: All specifications include education and experience controls. Heteroskedasticity-robust (HC3) standard errors are reported in parentheses.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

occupation-composition sorting.

9.6 Task-Content Index: Consistency with Baseline

We re-estimate the annual premium trajectory and quantile regressions using the task-content index (Section 8) as the skill measure. The annual trend is qualitatively identical: a hump-shaped pattern peaking in 2018–2019 and compressing post-2020. Quantile estimates confirm that higher-wage postings carry larger task-AI-substitutability scores, consistent with the distributional heterogeneity documented in Section 7. The sign reversal relative to the keyword index (negative vs. positive coefficient) reflects the conceptual distinction between task substitutability and worker skill demand, as discussed in Section 8; the *shape* of the premium trajectory is robust across both measurement approaches.

10 Conclusion

This paper documents three stylized facts about digital skill markets in China: time-varying premia, sub-additive skill combinations, and skill upgrading without wage compression. A task-based production framework (Section 3) unifies the dynamics, complementarity, and wage-dispersion findings.

First, digital-skill wage premia are not static, but their level is subject to a measurement caveat. The AI/ML premium followed a hump-shaped trajectory over 2016–2025 in raw OLS, consistent with a demand shock followed by a lagged supply response; however, adding salary-interval-width fixed effects reduces the pooled estimate from 0.434 to 0.011 log points (significant at 5%, $p = 0.013$), indicating that advertised salary midpoints substantially overstate realized premia. We treat the 0.177–0.225 log-point range from aggressive trimming as a conservative lower bound and the raw OLS as an upper bound. The hump-shaped *trajectory* remains informative about relative dynamics even if the level is upward-biased. Temporal dynamics matter: cross-sectional studies that sample near a peak or trough will yield very different conclusions.

Second, skill combinations are sub-additive across all five tested pairs. Holding multiple digital skills yields a joint premium consistently below the sum of individual premia, with the largest gap at Programming $\times$ Data Engineering (-22.9 pp). The AI $\times$ Programming interaction is also negative and significant (-4.8 pp, $p < 0.001$), confirming sub-additivity across all five pairs. Importantly, all five significant interactions survive the addition of job-category fixed effects (top-200 categories), ruling out occupation-composition sorting as the primary driver. The finding challenges O-ring theories of complementarity and implies that the marginal return to a second digital skill is substantially below the return to the first, though the policy implication depends on job type, not a blanket case for depth over breadth.

Third, Chinese digital labor demand exhibits skill upgrading, not employment polarization, yet diffusion and inequality are decoupled. The middle-skill tier expanded from 10.9% to 18.0% of postings while the high-skill tier contracted, inconsistent with a U-shaped shift. Despite this upgrading, wage inequality widened at the upper tail. Quantile regression reveals heterogeneous patterns: AI/ML premia peak in the upper-middle wage quantiles, while programming premia fall monotonically from Q10 to Q90. Broad diffusion of digital skills does not guarantee compressed inequality when the returns to those skills are heterogeneously distributed across the wage distribution.

These findings point toward three policy priorities. (i) *Curriculum timing*: because premia are time-varying, higher-education systems that adapt degree programs with 3–5 year lags may inadvertently overproduce workers for premia that are already compressing. (ii) *Depth over breadth*: the sub-additivity result supports focused mastery over broad digital literacy in skill-formation policy; students maximize returns by investing deeply in one high-premium skill rather than spreading effort thinly across multiple digital domains.

(iii) *Redistribution alongside diffusion*: policymakers who assume that broader digital-skill demand will automatically narrow wages are likely to be disappointed. The wage-dispersion paradox, diffusion without compression, calls for complementary redistributive instruments, particularly for lower-wage workers whose programming premia are large but whose absolute wages remain low.

Two methodological contributions sharpen the substantive findings. First, we document a salary-interval-width confound that is likely endemic to job-posting wage data: AI/ML postings advertise salary ranges twice as wide as non-AI postings, and once interval-width decile fixed effects are added, the pooled AI/ML premium collapses from 0.434 to 0.011 log points ($p = 0.013$). This finding is a caution for the growing literature that uses advertised salary midpoints as a wage proxy: headline OLS estimates from such data should be treated as upper bounds, and researchers should report interval-FE and trim-sensitivity checks alongside baseline results. Second, the continuous AI-exposure index (Section 8) confirms that our binary keyword results are not threshold-sensitive, aligning our measurement approach with the task-exposure framework of Acemoglu et al. (2022). The post-2020 compression in AI/ML premia is a robust national average trend consistent with lagged supply expansion; Ministry of Education data on STEM graduates show a contemporaneous correlation of $r = -0.80$ with the annual AI/ML premium, and supplementary descriptive evidence on the post-2020 period is provided in Appendix D.

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4B, pages 1043–1171. Elsevier.

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature*, 40(1):7–72.

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1):S293–S340.

Felten, E., Raj, M., and Seamans, R. (2019). The occupational impact of artificial intelligence: Labor, skills, and polarization. *SSRN Working Paper* 3368605.

Autor, D., Levy, F., and Murnane, R. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Deming, D. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics*, 132(4):1593–1640.

Ge, S. and Yang, D. (2011). Labor market developments in China: A neoclassical view. *China Economic Review*, 22(4):611–625.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, 104(8):2509–26.

Katz, L. and Murphy, K. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics*, 107(1):35–78.

Kremer, M. (1993). The O-ring theory of economic development. *Quarterly Journal of Economics*, 108(3):551–575.

Feng, S., Liu, J., and Lu, J. (2026). The decline of age-friendly jobs in China: Evidence from online job vacancies. *Economic Modelling*, 155:107399.

Wang, J., Cai, Z., Zeng, Z., and Liu, C. (2025). How do digital skills affect rural households' incomes in China? An explanation derived from factor allocation. *Sustainability*, 17(20):8967.

Wang, H. (2025). Research on the impact of artificial intelligence development on the wage gap of the labor force: From the perspectives of labor intensity and task cognition. *Advances in Economics, Management and Political Sciences*, 195:210–222.

Wu, Y., Lin, Z., Zhang, Q., and Wang, W. (2024). Artificial intelligence, wage dynamics, and inequality: Empirical evidence from Chinese listed firms. *International Review of Economics & Finance*, 96:103739.

Yang, G., Yao, S., and Dong, X. (2023). Digital economy and wage gap between high- and low-skilled workers. *Digital Economy and Sustainable Development*, 1:7.

A Data Construction and Descriptive Statistics

A.1 Data Source and Sample Coverage

The raw data are drawn from a commercial recruitment database covering job postings published by firms listed on the Shanghai Stock Exchange, the Shenzhen Stock Exchange, and the Beijing Stock Exchange, together with their subsidiaries and affiliated enterprises. As of end-2024, approximately 5,300 A-share listed companies are active on these exchanges. The database expands coverage to roughly 68,000 unique employer entities by including wholly-owned subsidiaries, majority-controlled subsidiaries, and equity-method associates (联营企业) of each listed parent. The mapping from listed parent to affiliated entity is maintained by the data vendor using the CSRC ownership-disclosure filings and the National Enterprise Credit Information Publicity System (国家企业信用信息公示系统). Each posting record carries both the recruiting entity's name and the associated listed-company stock code, allowing us to link postings to parent-firm industry classifications (CSRC first-level industry, 22 categories) and to verify that the employer is a genuine affiliate rather than a name-similar unrelated firm.

The raw database spans 2014–2026 and contains approximately 6.9 million posting records after deduplication within each calendar year. We restrict the analysis sample to 2016–2025 for two reasons: (i) coverage of Tier-2 and Tier-3 cities is sparse before 2016, creating a composition bias toward Tier-1 cities in early years; (ii) the AI/ML skill category contains fewer than 0.3% of postings before 2016, making annual estimates unreliable. The wage-analysis sample further requires a non-missing, non-zero advertised salary range, yielding 820,807 observations.

A.2 Sample Construction Pipeline

Step 1: Annual CSV ingestion and field standardization. Each year's raw CSV is read with UTF-8 encoding (GBK fallback for pre-2018 files). The following fields are retained and renamed: employer name (`company`), stock code (`stock_code`), parent–affiliate relationship (`relation`), CSRC industry (`industry`), job title (`job_title`), work city (`city`), advertised salary range (`salary_min`, `salary_max`), job description text (`job_description`), education requirement (`edu_requirement`), experience requirement (`exp_requirement`), and posting date (`pub_date`).

Step 2: Variable construction. *Salary.* Monthly salary midpoint is computed as $salary\_mid = (salary\_min + salary\_max)/2$, with one-sided imputation when only one bound is reported. Observations with $salary\_mid < 500$ yuan or $> 500{,}000$ yuan are set to missing and excluded from the wage analysis (these thresholds remove data-entry errors while retaining legitimate high-pay postings in finance and technology).

Education rank. The ordinal variable `edu_rank` maps the categorical education requirement to $\{0, 1, 2, 3, 4, 5, 6\}$ corresponding to {no requirement, junior high, senior high / vocational, associate, bachelor's, master's, doctoral}.

Experience years. The string field is parsed by exact-match lookup against a standardized map (e.g., "1 年以内" $\rightarrow$ 0.5, "3 年" $\rightarrow$ 3, "10 年以上" $\rightarrow$ 10); unmatched strings are set to -1 (missing). A regular expression fallback extracts numeric years from free-text entries such as "工作经验 5 年以上".

City standardization. City names are normalized by stripping trailing administrative suffixes (市, 省, 区, 县) and parenthetical district qualifiers (e.g., "北京（海淀）" $\rightarrow$ "北京") using the regex `re.sub(r"[市省区县]$", "", s)`. Cities are then classified into three tiers: Tier 1 (Beijing, Shanghai, Guangzhou, Shenzhen), Tier 2 (20 major provincial capitals and sub-provincial cities), and Tier 3 (all remaining cities).

Step 3: Deduplication. Within each year, duplicate records are dropped on the composite key (`company`, `job_title`, `city`, `pub_year`). Across years, a string key is constructed and a running set tracks previously seen records, so postings that remain open across calendar years are counted only in their first appearance year.

Step 4: Skill extraction. Skill indicators are constructed by dictionary-based exact matching against the keyword lexicon described in Appendix B. For each posting, the

Python function `extract_skills_dict_match(text)` applies a compiled case-insensitive regular expression for each keyword (`re.compile(re.escape(word), re.IGNORECASE)`) and returns the set of matched categories. A binary indicator $S_{it}^{k} = 1$ if any keyword in category k is found in the job description or job title. All keywords are pre-registered in the `jieba` custom dictionary to ensure correct tokenization of compound technical terms (e.g., "深度学习", "数据仓库"). A TF-IDF scan of the full corpus (top-50 terms by mean document frequency) is run after each annual update to flag emerging terms not yet in the dictionary; terms appearing in $\geq 0.5\%$ of postings and semantically related to an existing category are added in the next dictionary revision.

A.3 Descriptive Statistics

Table 13: Descriptive Statistics

Variable	Mean	SD	P10	P90
Panel A. Wage and Posting Characteristics				
Monthly salary (yuan)	11,240	6,830	5,000	20,000
Log monthly salary	9.18	0.57	8.52	9.90
Education rank	3.62	0.81	3	5
Experience (years)	3.14	2.41	0.5	7
City tier	1.84	0.79	1	3
Panel B. Skill Prevalence				
Any digital skill	0.798			
AI / Machine Learning	0.013			
Programming	0.106			
Data Engineering	0.048			
Cloud / DevOps	0.035			
Data Analysis	0.216			
Panel C. Sample Composition				
Employers	68,241			
Listed parent firms	5,284			
Subsidiaries / affiliates	62,957			
Industries	22			
Cities	287			
Years	2016–2025			
Observations		820,807		

Notes: The wage sample includes postings with non-missing advertised salaries between 500 and 500,000 yuan per month during 2016–2025.
Education rank ranges from no requirement to doctoral degree. Experience observations without explicit requirements are excluded from summary calculations.
Skill prevalence measures the share of postings containing at least one keyword associated with each skill category.

A.4 Complete Skill Keyword Lexicon

Table 14 reports the full keyword dictionary used for skill extraction. The dictionary is organized into five digital-skill categories (the core research variables) and six supplementary categories (used as controls or for robustness checks). Keywords are matched case-insensitively; Chinese and English variants of the same term are listed together.

Table 14: Complete Skill Keyword Lexicon

Category	Keywords
Digital / Technology Skills (core research variables)	
AI / Machine Learning	机器学习, 深度学习, 人工智能, 神经网络, 自然语言处理, NLP, 计算机视觉, CV, 强化学习, 大模型, LLM, GPT, BERT, Transformer, TensorFlow, PyTorch, Keras, scikit-learn, 特征工程, 模型调优, AutoML
Programming	Python, Java, C++, C#, JavaScript, TypeScript, Go, Rust, Scala, Kotlin, Swift, Linux, Shell, Git, Docker, Kubernetes, 微服务, API 开发, 后端开发, 前端开发, 全栈
Data Engineering	大数据, Hadoop, Spark, Hive, Flink, Kafka, 数据仓库, 数据湖, 数据中台, 实时计算, 离线计算, MySQL, PostgreSQL, MongoDB, Redis, Elasticsearch
Cloud / DevOps	云计算, AWS, Azure, 阿里云, 腾讯云, 华为云, DevOps, CI/CD, 容器, Docker, Kubernetes, 运维
Data Analysis	数据分析, 数据挖掘, 统计分析, 商业分析, 业务分析, 数据可视化, BI, 报表, A/B 测试, 用户画像, Python, R 语言, SPSS, SAS, Tableau, PowerBI, Excel, SQL, 数据库, ETL
General Skills (control variables)	
Communication	沟通, 表达, 汇报, 演讲, 写作, 文案, 方案撰写, 跨部门协作, 客户沟通, 谈判
Management	项目管理, 团队管理, 人员管理, 绩效管理, 资源管理, PMP, 敏捷, Scrum, OKR, KPI
Problem Solving	逻辑思维, 结构化思维, 问题分析, 方案设计, 决策, 创新, 学习能力
Domain / Professional Skills (control variables)	
Finance	财务分析, 会计, 审计, 税务, 成本控制, 预算, 财务建模, 估值, DCF, 三大报表, CPA, CFA
Marketing	市场营销, 品牌, 广告, SEO, SEM, 内容营销, 社交媒体, 私域流量, 用户增长, 增长黑客, GMV
Supply Chain	供应链, 采购, 物流, 仓储, 库存管理, ERP, SAP, 供应商管理, 谈判, 降本增效
HR	招聘, 培训, 绩效, 薪酬, HRBP, 人力资源规划, 劳动法, 员工关系
Product Management	产品经理, 需求分析, 原型设计, Axure, Figma, 用户研究, 产品规划, 竞品分析, PRD

Matching is case-insensitive and uses exact substring search (`re.search` with `re.IGNORECASE`). All keywords are pre-registered in the `jieba` custom dictionary prior to tokenization. Docker and Kubernetes appear in both Programming and Cloud/DevOps; a posting is counted in each category independently. Python appears in both Programming and Data Analysis; the two categories are not mutually exclusive by design, as the paper estimates their premia separately in a single regression. The negation filter described in Section 4 removes matches preceded within 10 characters by negation phrases (不需要, 无需, 不要求, 非必须).

B Task-Content Index: Mapping Details

B.1 AI Progress Index Scores

Felten et al. (2019) measure AI progress across nine capability domains via a crowd-sourced benchmark (AI Progress Index, AIPI). Table 15 reproduces the nine domain scores used in this paper.

Table 15: AI Capability Domain Scores (AIPI)

AI capability domain	Representative tasks	AIPI score
Image recognition	Object detection, image classification	0.95
Speech recognition	Transcription, voice commands	0.93
Translation	Machine translation	0.85
Language modeling	Text generation, completion	0.88
Reading comprehension	Question answering over text	0.82
Visual QA	Image + text joint reasoning	0.80
Predictive modeling	Forecasting, anomaly detection	0.78
Recommendation	Collaborative filtering	0.72
Planning / games	Sequential decision-making	0.65
Management	Coordination, strategy (low-exposure reference)	0.12
Creative judgment	Design, artistic creation	0.18
Physical operation	On-site inspection, manual assembly	0.15

Rows 1–9: AIPI scores from Felten et al. (2019) Table 1, standardized to $[0, 1]$. Rows 10–12: low-exposure reference levels derived from the mean AI-exposure score of managerial, creative, and physical occupations in Acemoglu et al. (2022) Table A1; these are not Felten scores. Postings with no keyword match receive the median AIPI score (0.80).

B.2 Job-Title Keyword to Felten Domain Mapping

Table 16: Job-Title Keyword to Felten AI Capability Domain

Job-title keyword	Felten domain(s)	AIPI score
视觉, 图像, CV	Image recognition, Visual QA	0.95, 0.80
语音, ASR	Speech recognition	0.93
NLP, 自然语言, 大模型, LLM	Language modeling, Reading comprehension	0.88, 0.82
翻译	Translation	0.85
算法, 机器学习, 预测, 风控	Predictive modeling	0.78
深度学习	Predictive modeling, Image recognition	0.78, 0.95
推荐	Recommendation	0.72
强化学习	Planning / games	0.65
数据标注	Image recognition, Language modeling	0.95, 0.88
总监, 经理, 总裁, HR, 销售	Management	0.12
设计	Creative judgment	0.18
生产, 质检, 运维	Physical operation	0.15
(no match)	—	0.80 (median)

When a keyword matches multiple domains, the posting's score is the mean AIPI across matched domains. Domain scores are from Felten et al. (2019) Table 1 (rows 1–9 of Table 15) or the low-exposure reference levels (rows 10–12); no score is author-assigned.

B.3 Relationship to Acemoglu et al. (2022)

Acemoglu et al. (2022) construct an occupation-level AI-exposure index via the chain: SOC occupation → O*NET task descriptions → Felten et al. (2019) AIPI scores → weighted average. Our index follows the same terminal step, anchoring scores in the Felten et al. (2019) benchmark, but replaces the O*NET intermediate step with direct job-title keyword matching, because O*NET crosswalks for Chinese occupations are unavailable.

The key methodological property is preserved: all domain scores are externally anchored in the Felten et al. crowdsourced benchmark and are not author-assigned. The main limitation relative to Acemoglu et al. (2022) is that our mapping is coarser (keyword-level rather than task-description-level) and does not exploit within-occupation task variation.

C Future Research Directions

Two open questions are left for future work:

1. **Worker-level skill–wage matching.** This paper measures employer-stated skill *requirements*; whether workers actually possess those skills, and how the match quality affects realized wages, cannot be assessed from job-posting data alone. Linking posting data to worker-level survey data (e.g., CFPS, CHNS) or platform-level résumé data would allow direct estimation of skill–wage matching premia.

2. **Generative AI heterogeneity.** The 2023 rebound in AI/ML premia coincides with the diffusion of large language models. Whether LLM-related skills (prompt engineering, fine-tuning, RAG pipelines) command premia distinct from traditional ML skills, and whether they substitute for or complement programming skills, is an open empirical question that the current keyword dictionary cannot fully resolve.

D COVID-19 Shock: Supplementary Descriptive Analysis

This appendix provides supplementary descriptive evidence on the 2020–2024 pandemic episode and its association with digital-skill wage premia. Three caveats apply to all results below. First, high-severity cities (predominantly Tier-1 cities) already exhibited faster-declining skill premia before 2020; the parallel-trends assumption is not satisfied, and no result in this appendix should be interpreted as a causal effect. Second, city-level cumulative confirmed cases are a coarse proxy for pandemic severity: they cannot capture lockdown stringency, the distinct dynamics of multiple waves (initial 2020 outbreak, Delta, Omicron 2022), or within-city heterogeneity. Third, these findings do not affect the paper's core conclusions, which rest on the full 2016–2025 long-run trend; the pandemic is a short-lived disruption within that trend.

D.1 Annual Descriptive Evidence

We use city-level confirmed COVID-19 cases (cumulative peak through end-2020, from the DXY COVID-19 dataset) as a proxy for city-level pandemic severity, $Severity_c$, covering 473 prefecture-level cities across 32 provinces. The estimating equation is:

$$\begin{aligned}\ln w_{icft} = \alpha + \beta_1 Post_t + \beta_2 Severity_c + \beta_3 Post_t \times Severity_c \\ + \beta_4 S^k_{icft} \times Post_t + \beta_5 S^k_{icft} \times Post_t \times Severity_c + \gamma X_{it} + \mu_j + \varepsilon_{icft},\end{aligned}$$

where $Post_t = \mathbf{1}[t \geq 2020]$ and $Severity_c$ is the standardized log of province-level confirmed cases. Standard errors are clustered at the province level.

Table 17 reports the triple-interaction estimates. The $S^k \times Post$ coefficients confirm the average post-2020 compression documented in Section 5, with AI/ML, data engineering, and Cloud/DevOps premia falling by 13–16 log points. The triple interaction β_5 is insignificant for all four skills (AI/ML: -0.015, $p = 0.330$; Programming: 0.002, $p = 0.947$; Data Engineering: 0.004, $p = 0.863$; Cloud/DevOps: 0.030, $p = 0.140$), indicating no detectable differential effect by city-level COVID severity.

Pre-trend coefficients are large and significant for all skills (AI/ML: -0.067, $p < 0.001$; Programming: -0.084, $p < 0.001$; Data Engineering: -0.100, $p < 0.001$; Cloud/DevOps: -0.049, $p < 0.001$), confirming that high-severity provinces already had faster-declining premia before 2020. **The parallel-trends assumption is not satisfied.** All estimates in this appendix reflect correlational patterns, not causal effects.

Figure 7 plots the grouped event study: $S^k \times Year$ coefficients (relative to 2017 baseline) estimated separately for high- and low-severity provinces. The divergence between the two groups *before* 2020 confirms the pre-trend failure visually.

Table 17: COVID-Period Heterogeneity Estimates

	AI/ML (1)	Programming (2)	Data Engineering (3)	Cloud/DevOps (4)
Panel A. Triple-Interaction Estimates				
$S^k \times Post$	-0.157***	-0.083***	-0.151***	-0.133***
	(0.024)	(0.014)	(0.017)	(0.024)
$S^k \times Post \times Severity$	-0.015	0.002	0.004	0.030
	(0.015)	(0.026)	(0.021)	(0.020)
Panel B. Pre-trend Test				
$S^k \times Trend$	-0.067***	-0.084***	-0.100***	-0.049***
	(0.018)	(0.012)	(0.013)	(0.013)
Controls	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Observations	500,214	500,214	500,214	500,214
Sample Period	2016–2023	2016–2023	2016–2023	2016–2023

Notes: All specifications include education and experience controls.
Standard errors clustered at the province level are reported in parentheses.
Panel B reports pre-trend tests. Significant estimates indicate deviations from parallel trends, suggesting that results should be interpreted as conditional associations rather than causal effects.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

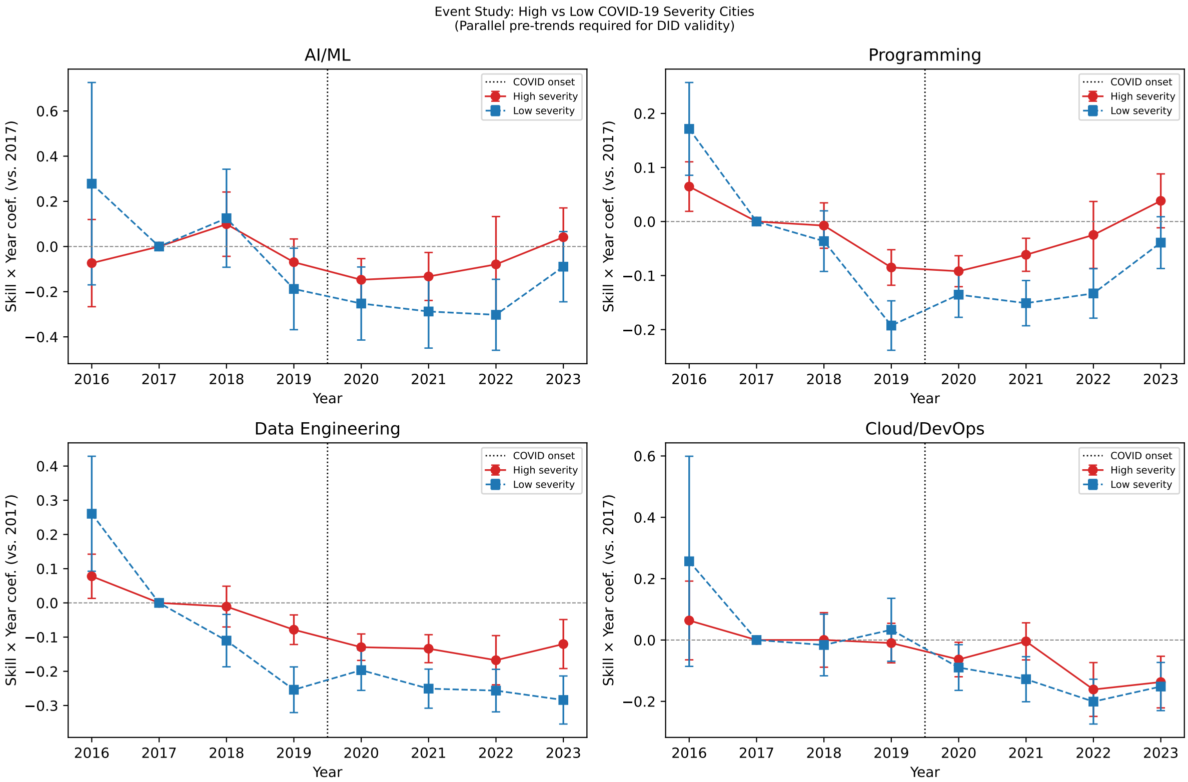

Figure 7: Grouped event study.

D.2 Quarterly Dynamic Event Study

To exploit quarterly variation in COVID-19 exposure, we re-estimate the event study at the city–quarter level using a continuous treatment design. The estimating equation is:

$$\ln w_{icqt} = \sum_{\tau \neq 0} \beta_\tau \Big(Severity_c \times \mathbf{1}[q = \tau] \Big) + \gamma X_{it} + \delta_c + \delta_q + \varepsilon_{icqt}, \qquad (10)$$

where q indexes calendar quarters, τ is measured relative to the base period 2019Q4, and δ_c, δ_q are city and quarter fixed effects. Standard errors are clustered at the city level. We include interactions of three 2019 city-level characteristics (log GDP per capita, digital infrastructure index, tertiary-sector share) with a linear time trend to control for differential pre-existing trends.

Pre-period coefficients ($\tau = -3, -2, -1$) remain positive and significant, reflecting the composition of high-severity cities (Tier-1 metropolitan areas with high concentrations of digital-economy employers). These estimates should be interpreted as *conditional correlations* rather than causal effects.

Table 18 reports phase-specific estimates. Programming skills show the most consistent pattern: positive and significant across all four phases, with the largest coefficient in Phase 3 (0.067, $p < 0.01$). Data Engineering premia are significant in Phases 1–3 but turn insignificant in Phase 4, consistent with a normalization of data-infrastructure demand after reopening. Cloud/DevOps premia are positive and significant in Phases 1 and 2 (0.036 and 0.024, both $p < 0.05$), consistent with sustained remote-infrastructure demand during the endemic period, but insignificant thereafter. AI/ML premia show no significant association with COVID severity in any phase.

Table 18: Skill Premium Associations Across COVID Phases

	Phase 1 (1)	Phase 2 (2)	Phase 3 (3)	Phase 4 (4)
AI/ML	0.003	-0.005	0.023	0.018
	(0.017)	(0.012)	(0.026)	(0.012)
Programming	0.027***	0.022**	0.067***	0.040***
	(0.009)	(0.009)	(0.018)	(0.015)
Data Engineering	0.029***	0.016**	0.036**	-0.010
	(0.009)	(0.007)	(0.016)	(0.015)
Cloud/DevOps	0.036**	0.024***	-0.024	-0.016
	(0.016)	(0.006)	(0.019)	(0.012)
Controls	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	158,214	158,214	158,214	158,214

Notes: Phase definitions are: Phase 1 (2020Q1–2020Q2), Phase 2 (2020Q3–2022Q2), Phase 3 (2022Q3–2022Q4), and Phase 4 (2023Q1–2024Q4).
All specifications include city-level controls interacted with linear time trends, including log GDP per capita, digital infrastructure, and tertiary-sector share.
Standard errors clustered at the city level are reported in parentheses.
Estimates should be interpreted as conditional associations rather than causal effects.
*, **, and *** denote significance at the 10%, 5%, and 1% levels.

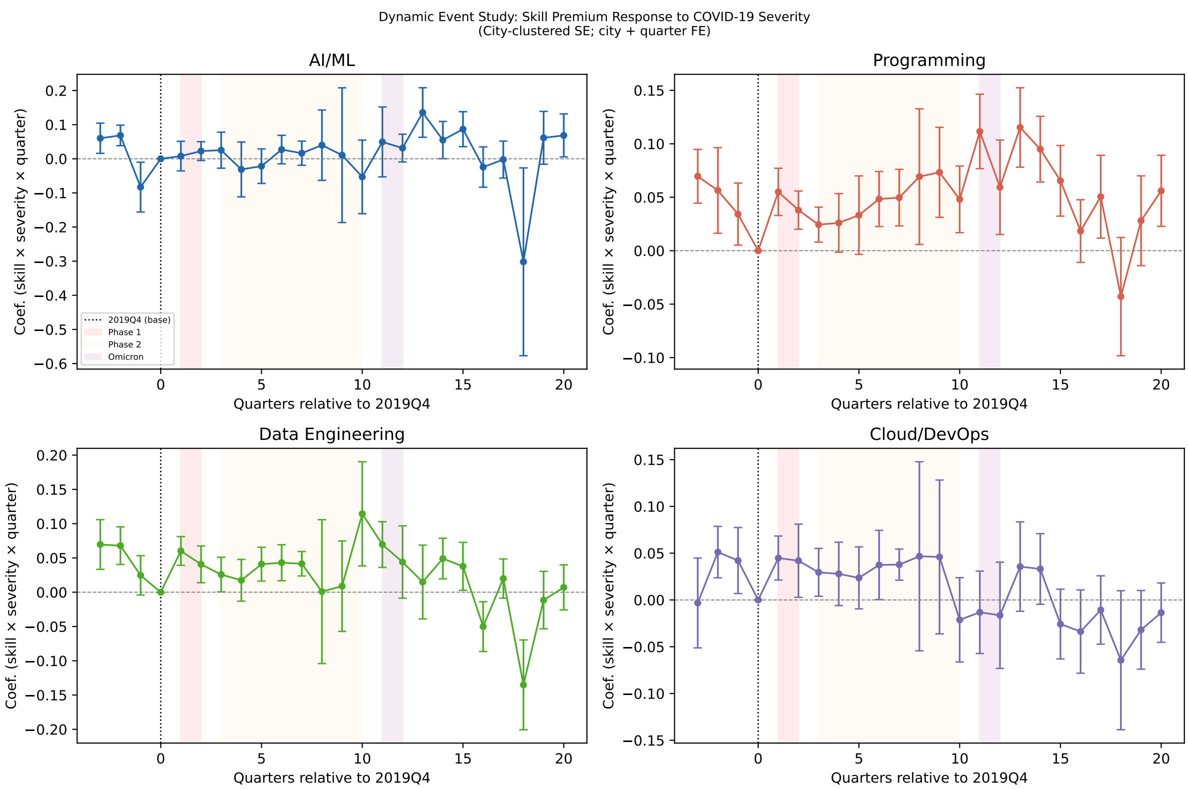

Figure 8: Dynamic event study: $Severity_c \times \mathbf{1}[q = \tau]$ coefficients relative to 2019Q4 base period.

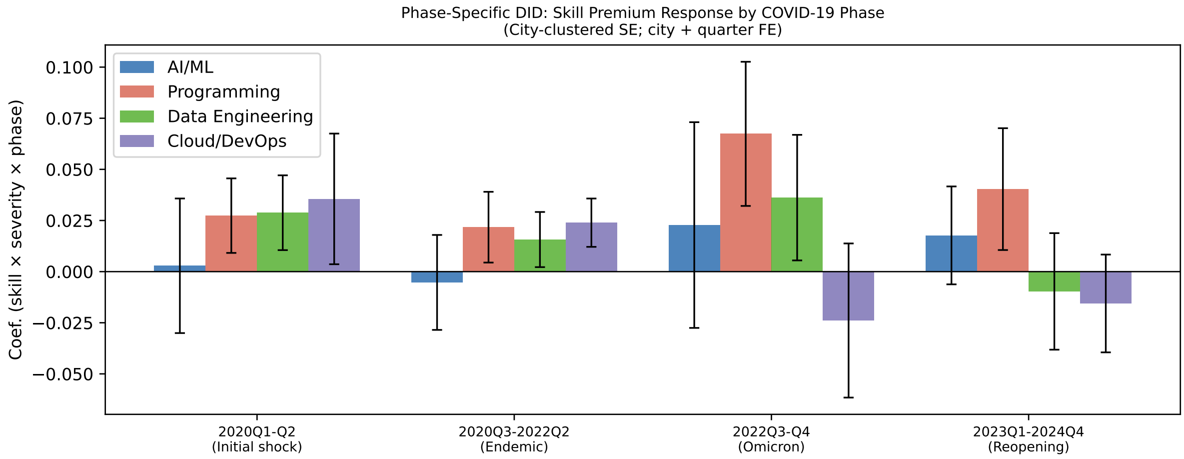

Figure 9: Phase-specific estimates ($Severity_c \times Phase_p \times S^k$).

D.3 Limitations and Future Identification

The descriptive patterns in this appendix are consistent with the long-run supply-expansion narrative in Section 5: the pandemic did not permanently alter the trajectory of digital-skill premia, which recovered to trend within two years. Causal identification of the pandemic's net effect would require city-level lockdown duration data (e.g., subway ridership or municipal health commission records) that can separate pandemic severity from the pre-existing concentration of digital-economy employers in Tier-1 cities. Quarterly data with such a proxy would also allow distinguishing the heterogeneous shocks of the initial 2020 outbreak, the Delta wave, and the Omicron surge, a natural direction for future work.