Discussion of "Weak Bundle, Strong Bundle: How AI Redraws Job Boundaries"

By Luis Garicano, with Jin Li, Yanhui Wu, Leo Yang Yang and Zhenyu Zhu

Discussant: Benjamín Villena-Roldán

Facultad de Economía y Negocios - Universidad Andrés Bello

XXIX Annual Conference of the Central Bank of Chile, November 20, 2026

AI removes a task only where splitting the job is cheap

- Jobs bundle a codifiable and a contextual task; splitting them destroys output share c.
- Weak bundle (low c): AI takes the task as AI quality increases, the job narrows.
- Strong bundle (high c): AI assists inside the job.
- Model captures in c a reduced form of various reasons behind human attachment to jobs.
- Interesting departure from Acemoglu–Restrepo: the unit is the job; their task displacement is the weak-bundle case ($c \to 0$).
- Empirics based on 210 million Chinese postings, 2023–2026.

Outcome	β	t	t, no trend
Ads requiring GenAI tools (pp)	−0.37	−2.93	−3.32
Ads with no exposed task (pp)	−0.60	−3.79	−3.25
Share of exposed tasks (pp)	+0.29	3.21	0.35
Posted wage (log ×100)	+1.14	2.89	0.71

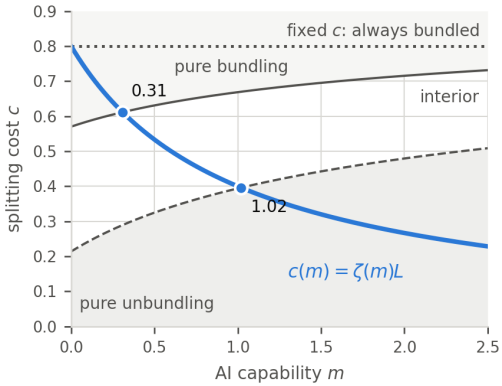
β: exposure (tasks an LLM can do) × bundle strength, per SD of each. "No trend": net of pre-trend. Paper, Table 4.

Comment 1: Liability raises c only through legal frictions

- In the model, c is share of lost output due the technology of the job..
- Liability decides who pays a harm D. It binds only if the worker can pay $L < D$, AI cannot be sued, or law requires a human signature.
- O*NET measures D in the US: nine of the ten strongest bundles are physicians and dentists.
- China: the hospital is sued and pays in case of malpractice; since 2022 AI may not replace the physician.
- **Suggestion:** take components separately; split licensed occupations.

Comment 2: with liability, c falls as AI improves

- Suppose task 2 is only review, $t_1 + t_2 = 1$. AI errs with probability $\zeta(m)$, $\zeta' < 0$; review finds the error with probability t_2; an error costs L (in units of revenue).
- Bundled: $P\,\mu_H(m)^{1-\eta}(t_1 s_1)^{\eta}\,[1 - \zeta(m)L\,t_1]$
  Split: $P\,\mu_A(m)\,[1 - \zeta(m)L]$, so $c(m) = \zeta(m)L$ falls with m.
- Review time $t_2^* = \max\{0,\, 1 - \frac{\eta}{(1+\eta)\,\zeta(m)L}\}$ is zero from $\hat{m} = 0.87$ on: the signature becomes a formality.
- Unless the law requires the signature, the bundle splits sooner than the paper's regimes imply.



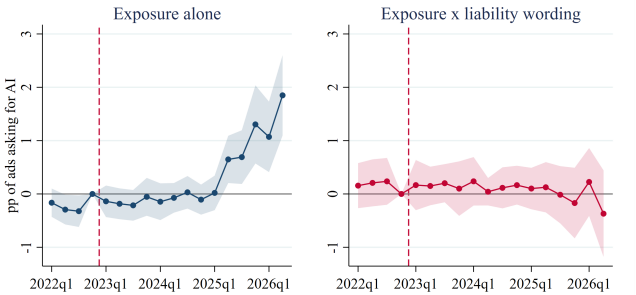
$\zeta(m) = 0.1/(1+m)$, $L = 8$. Regimes from the paper's Figure 1 benchmark ($\alpha = 0.5$, $\eta = 0.75$, $\mu_H = 1 + m$, $\mu_A = 1.5 + 1.5m$). With c fixed at 0.8 the job stays bundled; with $c(m)$ it enters the interior at $m = 0.31$ and fully splits at $m = 1.02$.

Comment 3: the AI-tools outcome and bundle index admit other readings

- In the model, the bundled worker uses AI tools. Assume that ads for residual workers request only task 2 and no AI tools. Therefore, ads asking for AI tools should rise most in strong bundles.
- The paper finds a larger rise in weak bundles (-0.37). That needs the residual worker to run the AI. Test: do AI-tools ads also list exposed tasks?
- Bundle strength is one number per occupation, from US ratings (O*NET, August 2026) matched by an LLM. Other occupation traits can load on it. Rebuild it from a pre-2024 release.
- Ads can measure bundling within an occupation, where the model's sorting happens: some make the holder answer for a result (a unit, compliance, a sign-off), others only list tasks.
- Effects are 0.3 to 0.6 pp per SD of each index. Without baseline shares (how many ads ask for AI tools, how many are residual), their size is hard to judge.

Chile: liability wording pays more, but the interaction is flat

- 774,015 Trabajando.com ads, 2021–2026; 26.9% say who answers for people, a unit or compliance.
- Within occupation and quarter they ask +0.08 skill categories (mean 1.16) and pay +5.1 log points.
- Exposure alone: +1.8 pp of ads asking for AI per SD, by 2026.
- Exposure × liability wording: −0.24 (SE 0.25).



88 white-collar occupations; quarterly coefficients, 2022q4 = 0, 95% intervals. Exposure and liability wording: 2021 ad shares, standardized.

Summing up

- The job is the right unit of analysis, and the paper says plainly which results survive once pre-existing trends are removed.
- Comment 1: liability raises c only through legal frictions, and who bears it differs across countries (US: the physician; China: the hospital).
- Comment 2: the liability part of c falls as AI improves. A bundle held by a legally required signature survives, as a rent.
- Comment 3: if residual ads ask for no AI tools, the model predicts that ads asking for AI tools rise most in strong bundles; the paper finds the opposite. Rebuild bundle strength from a pre-2024 O*NET release, and report baseline shares.
- Enter the three components of c separately, split licensed occupations, and measure bundling in the ads. Chile: ads with liability wording post wages 5.1 log points higher; the exposure $\times$ wording interaction is imprecise (-0.24, SE 0.25).

Backup: three smaller points

- **Sign-off does two jobs in the paper.** It explains why AI's help inside the job grows more slowly than AI working alone (Assumption 1), and it is one source of the splitting cost c. One model of human review would link the two.
- **The pay cut after a split may be smaller.** In Lemma 2 the worker's share of revenue falls from κ to $1 - \alpha$ when the job splits. If the residual worker still signs for the AI's output, she is paid for that risk, and her share stays above $1 - \alpha$.
- **The two bundle measures move differently with exposure.** Across occupations, liability wording in Chilean ads rises with exposure (correlation $+0.24$), while the paper's O*NET index falls with exposure (-0.35). They may capture different things.

Backup: Table 4 of the paper; two main results survive the trend check

Outcome	Change, late − pre	SE	t	Pre-trend t	t net of trend
Ads asking for AI tools (pp)	−0.367	0.125	−2.93	1.18	**−3.32**
Share of task mentions that are exposed (pp)	0.294	0.092	3.21	1.43	0.35
Ads with no exposed task, "residual" (pp)	−0.599	0.158	−3.79	0.78	**−3.25**
Posted wage (log points)	1.144	0.396	2.89	1.13	0.71
Asks for a bachelor's degree or more (pp)	0.118	0.176	0.67	0.45	0.09
Asks for three or more years of experience (pp)	0.232	0.214	1.08	0.01	0.29
New postings (log points)	−1.979	1.859	−1.06	0.59	−0.98

Each row is the coefficient on exposure × bundle strength, both measured in standard deviations. Change: its average in January–June 2026 minus its average in August 2023–March 2024. A negative value means the outcome rose less, or fell more, in exposed occupations with stronger bundles. Pre-trend t: December 2023–March 2024 against August–November 2023. t net of trend: the late value tested against a straight line fitted to the pre months and extended to 2026; bold where it stays above 2 in absolute value. pp: percentage points. Standard errors clustered by O*NET code.

Backup: Chile, liability ads ask for more skills and post higher wages

Outcome	Liability ads minus other ads	SE
Number of skill types asked, 0 to 5 (mean 1.16)	0.08	0.02
comparing ads of the same firm	0.04	0.01
Posted wage (log points, about percent)	5.1	0.8
comparing ads of the same firm	4.9	0.6

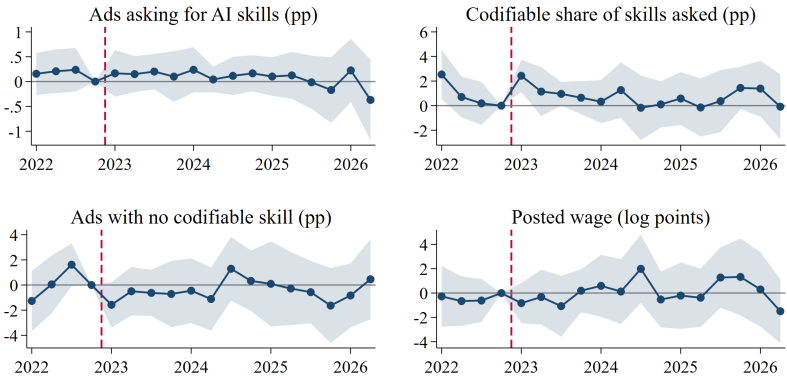
773,979 ads in 140 four-digit occupations, 2021–2026. Every comparison is between ads of the same occupation and quarter, holding ad length fixed; the wage rows also hold the education $\times$ experience cell fixed. Liability wording: one of the five phrases ("responsable de", "a su cargo", "velar por", "garantizar el cumplimiento", leading a team). Skill types: writing, customer service, financial, general computer, specific software. Standard errors clustered by occupation.

Backup: Chile, the interaction shows no clear pattern in seven outcomes

	Bundle strength measured by					
	liability wording				people management	
Outcome	β	SE	t	Pre-trend t	β	t
Ads asking for AI skills (pp)	−0.24	0.25	−0.96	−0.59	−0.09	−0.44
Codifiable share of skills asked (pp)	−0.04	0.91	−0.04	−2.26	1.08	1.34
Ads with no codifiable skill (pp)	−0.80	1.05	−0.76	1.58	−2.23	−1.95
Posted wage (log points)	0.79	1.00	0.79	0.19	1.82	1.89
Asks for three or more years of experience (pp)	−1.45	0.77	−1.89	0.16	−2.28	−2.95
Asks for college or more (pp)	−3.12	1.14	−2.73	−1.07	−1.90	−1.65
Number of ads (log points)	−1.32	4.06	−0.32	0.99	−5.15	−1.07

β is the paper's object: how much the effect of one standard deviation of exposure changes, from 2022 to 2025q3–2026q2, where bundle strength is one standard deviation higher. Exposure: the 2021 share of an occupation's ads asking for codifiable skills (writing, financial, office, computer, software or programming). Liability wording: the 2021 share of ads with one of the five phrases. People management: the 2021 share of ads asking the holder to manage people. Ads of 2021 are not in the regression. Pre-trend t: second half of 2022 against the first. 88 white-collar occupations, 1,475 occupation-quarter cells; occupation and quarter fixed effects; standard errors clustered by occupation. Two of the 14 estimates have $|t| > 2$; about one would by chance.

Backup: Chile, the interaction by quarter shows no break after ChatGPT



Each dot is the exposure × liability-wording coefficient for that quarter, relative to 2022q4, with its 95% interval (clustered by occupation). The dashed line is ChatGPT (30 November 2022); a rise after it would mean that exposed occupations where liability wording is common changed more than other exposed occupations. Exposure and liability wording: 2021 shares of an occupation's ads, standardized. Wages are net of education × experience cell means.

Backup: the liability keywords were checked by reading 380 snippets

Pattern	Strict	Broad	Decision
responsable de/del/por	29/40	36/40	kept, singular only
a (su) cargo	33/40	38/40	kept
velar por	31/40	40/40	kept
garantizar/asegurar el cumplimiento	30/40	40/40	kept
liderar/supervisar/dirigir (al) equipo	40/40	40/40	added
gerunds: *velando por, asegurando el...*	23/40	40/40	left out
toma de decisiones	2/40	7/40	dropped
rendición de cuentas	8/20	14/20	dropped
representante legal, responsabilidad civil...	4/20	5/20	dropped

For each pattern, random ads that match it; about 30 words around the match are read. Strict: the snippet makes the job holder answer for people, a unit, assets, a function, a target or compliance. Broad: also counts routine duties and competencies. Decision: whether the pattern enters the liability flag. Coded by an AI assistant (Claude); no person double-coded them. Missed cases: of 60 ads that use a related word but match no pattern, 6 contained a strict statement.

Backup: Chile, the pay gap of liability ads grew in exposed occupations

- **Question.** Did the wage gap between ads with and without liability wording grow more in occupations more exposed to AI? Ads are compared within the same occupation and quarter.
- **Answer.** Relative to 2022, the gap is wider by 3.2 log points per standard deviation of exposure in 2023–24 (SE 1.2), and by 2.5 in 2025–26 (SE 1.3).
- **Earlier trend.** 2021 against 2022: -0.4 (SE 1.2), so the gap was not already widening.
- **Same firm.** Comparing ads of the same firm, the 2025–26 estimate falls to 1.2 (SE 1.5), so the result is suggestive.
- This fits Comment 2: signing for AI output may carry a premium.

Backup: in China the hospital bears malpractice liability

	United States	China
Who is sued and pays	the physician, through insurance	the hospital, in full
The physician's position	self-employed or employed	employee of a public hospital
Lawsuits a year, per 100 physicians	1.8 (physicians sued)	0.34 (hospitals sued; Chengdu)
Average payment	US$348,065 (2018)	US$31,430 (2008–17)
What a lawsuit costs the physician	premiums, time, reputation	pay, career, stress, violence
Rules on AI replacing the physician	none at national level; licensing applies	barred in internet care since 2022

O*NET's liability items measure how much harm an error can do. Who pays for it differs: in China the physician bears it only through her employer.

China: Liu, Yi, Yuan and Zhao (2023), *J. Comp. Econ.* 51: 1142–1161; Chengdu hospitals 2009–2014, lawsuits per hospital-year divided by physicians per hospital. They find lawsuits raise spending on tests, drugs and diagnostic devices, with no change in physician employment. US: share of physicians sued in the past year, AMA survey. AI rule: 2022 national rules on internet diagnosis and treatment, which also forbid AI-generated prescriptions.

Backup: the Chinese evidence fits bundles but does not yet rule out tasks

Prediction about	Acemoglu–Restrepo	Garicano et al.	Chinese data (t net of trend)
What changes ads in exposed occupations	exposure alone	exposure, more where bundles are weak	residual ads -3.25: fits; ads asking for AI tools -3.32: the literal model predicts the opposite sign
Effect of exposure on tasks, at average bundle strength	exposed tasks leave	they leave only where bundles are weak	none ($t = -0.20$): fits, or exposure is measured with noise
Tasks listed in strong-bundle jobs	exposed tasks leave	unchanged	share of exposed tasks, $t = 0.35$: no clear difference
Wages	fall with automation	fall only in weak bundles	posted wage, $t = 0.71$: no clear difference
Education and experience asked	no prediction	fall when the job splits	$t = 0.09$ and 0.29: no change
Shape of the response	gradual	a jump once c falls below a threshold	the paper fits straight lines; a high/low split is not significant
Employment	falls unless new tasks appear	falls in weak bundles if demand is inelastic	new postings, $t = -0.98$: imprecise

t is net of the pre-trend (paper, Table 4), except the second row (late minus pre, Cambridge slides, p. 32) and the high/low split (Table 9). A task model in which automation costs differ by occupation, and correlate with shared context or liability, also produces the interaction; telling the two apart needs the jump, unchanged tasks in strong bundles, and easier entry after a split. The authors chose the sample of platforms among 16,416 combinations; corrected for that search, $p = 0.136$ (five-platform search) and 0.106 (full search).