Skill Relatedness and the Absorption of AI Across US Labor Markets*

Sidharth Rony

August 4, 2026

Abstract

Where does a new general-purpose technology take root, and why do some regional labor markets absorb it faster than others? I track firm AI adoption at biweekly frequency across US states and large metros in the Census Business Trends and Outlook Survey, and relate it to each region's pre-existing skill structure, measured from 433.6 million job postings. Three findings emerge. First, the geography of adoption is largely pre-determined: skill structure alone predicts a fully held-out region's adoption with a correlation of 0.63, and the 2019 skill base predicts 2023–25 adoption nearly as well as the 2024 base. Second, the absorption margin is the connective tissue of the local skill ecosystem (its relatedness density), not its sophistication: one standard deviation of density predicts about one percentage point of additional adoption growth, while skill complexity predicts none. Third, each frontier model release widens the adoption gap between exposed and unexposed regions by roughly 0.1 percentage points, a treadmill that compounds; this dynamic reading is the hardest to separate from a smooth differential trend. Because laggards converge only to the steady states their fundamentals imply, a readiness typology of regions, and the margin each should target, follows.

JEL: J24, O33, R11, R12.

*First draft: June 2026; this version: August 2026. Preliminary. Data: Census Business Trends and Outlook Survey (public tables); Howell, Feldman, Lanahan, Kalathil, and Johnson (2026) county skill panel (Zenodo 20330086, CC-BY); BLS OEWS; Felten, Raj, and Seamans (2021); Eloundou, Manning, Mishkin, and Rock (2023). The event calendar, intensity scores, exposure measures, lag range, and placebo dates were fixed in a specification memo *ex ante of estimation* (Appendix B).

1 Introduction

General-purpose technologies have never diffused evenly across space: hybrid corn crossed Iowa a decade before it reached Alabama (Griliches, 1957), adoption lags of major technologies persist for decades (Comin and Hobijn, 2010), and the information-technology wave sorted American cities into diverging tiers (Moretti, 2012). Artificial intelligence arrives faster and lumpier still: frontier capability lands in discrete, dateable jumps (a model release, a price collapse, a new modality) on regional labor markets that differ enormously in their capacity to absorb it. If absorption capacity is unevenly distributed and slow to build, the geography of the AI economy is being written now, by pre-existing labor-market structure rather than by policy. This paper measures that process directly, asks what kind of regional capability governs it, and finds that the answer is written years in advance.

Three ingredients make the measurement possible. The Census Bureau's Business Trends and Outlook Survey (BTOS) provides the only high-frequency, representative, geographically resolved AI adoption series in existence: 54 biweekly waves for 50 states, DC, Puerto Rico, and the 25 largest metros over September 2023 to November 2025 (Bonney, Breaux, Buffington, Dinlersoz, Foster, Goldschlag, Haltiwanger, Kroff, and Savage, 2024). The Howell et al. (2026) county skill panel provides a deep *pre-period* measure of every labor market's capability base (skill specialization, complexity, and relatedness) computed from 433.6 million job postings over 2010–2024. And the 2023–2025 release calendar provides nine repeated, third-party-dateable capability shocks. I combine them in three complementary designs, whose identifying choices were fixed ex ante of estimation (Appendix B): an out-of-sample prediction protocol, a skill-moderation design, and a stacked multi-shock event design.

The first and most secure finding is that *the geography of adoption is largely pre-determined.* A model trained with a geography entirely deleted from the data predicts that geography's late-2025 adoption from pre-period skill structure alone with a correlation of 0.63 across state units, and a model trained only on pre-2025 data beats naive persistence out of sample, so this is prediction, not inertia. The capability that matters is old: in a vintage horse race, the 2019 skill base predicts 2023–25 adoption nearly as well as the 2024 base. And there is no unconditional convergence across regions; convergence appears only conditional on skill structure, implying laggards drift toward permanently lower steady states. Absorption capacity is a slow-moving structural endowment, not something regions built during the AI era itself. (The claim is carried by the state layer; the 25-metro layer is noisier.)

The second finding says *what* that endowment is. Two natural candidates from economic geography are the sophistication of a region's skill portfolio, its complexity, and the relatedness of its existing skills to the capabilities around them, what this paper calls the skill ecosystem's connective tissue (Section 2.2). The data consistently select connective tissue. One standard deviation of relatedness density predicts about 1.1 percentage points of additional adoption growth over the sample and more than doubles the hazard of adoption breakout, while complexity predicts nothing conditional on density; the same ordering holds under every exposure measure, under

four alternative complexity definitions, and under a doubly-robust machine-learning estimator that shares none of the event design's assumptions. The moderation attenuates in the thinnest, noisiest geographies, a caveat I carry openly in Section 4.6.

The mechanism evidence points the same way. Dense regions convert the adjacent skills they do not yet possess into revealed strengths at a faster rate, and their skill demand reshuffles more year to year, which is exactly what a connective-tissue account requires. Where the ecosystem is thin, non-adopting firms overwhelmingly report that AI is simply not applicable to their business, an applicability channel that tracks density. And worker-side usage style is strikingly uniform across states (augmentation versus automation shares vary by barely a percentage point), so the regional margin of divergence is *whether* firms adopt at all, not how AI is used once adopted.

Third, the release calendar makes it possible to watch the endowment work in real time: a *treadmill* in which every frontier release widens the gap between exposed and unexposed regions. Each typical release widens the adoption gap between a one-standard-deviation-more-exposed region and the average region by roughly 0.1 percentage points, about 0.8 points cumulated over the nine regime-1 releases, against a national rise of 6.2 points; the response builds one to four months after release and is absent on impact, inconsistent with hype or anticipation. Because releases arrive on a roughly even cadence, this step-by-step reading is only weakly separable from a smooth differential trend in the pooled data; the domain-matched design, in which only coding-domain shocks interacted with coding-intensive workforces carry the response, survives that trend control, and Section 4.6 weighs the evidence in full.

The policy product is a readiness typology. Comparing each region's actual adoption to the level its own fundamentals predict separates regions underperforming their fundamentals (Mississippi, at 4.3 points below prediction, New Jersey, Los Angeles, Detroit), the natural targets of diffusion policy, from overperformers such as Washington DC (8.6 points above). Each gap is decomposed by the region's relatively weaker fundamental, occupational *composition* versus skill *depth*, which map to different instruments: industry attraction versus long-horizon ecosystem building. The vintage result is a warning about the second margin: skill depth moves on multi-year horizons, so short-run training policy should not be expected to move adoption much.

This paper connects three literatures. A rapidly growing body of work measures AI adoption and its correlates: McElheran, Li, Brynjolfsson, Kroff, Dinlersoz, Foster, and Zolas (2024) map the descriptive geography of firm adoption in Census microdata, Acemoglu, Autor, Hazell, and Restrepo (2022) measure establishment-level AI demand in vacancies, Babina, Fedyk, He, and Hodson (2024) show AI-investing firms grow through product innovation, and worker-side generative-AI take-up has been strikingly fast (Bick, Blandin, and Deming, 2024; Humlum and Vestergaard, 2025). I add the diffusion margin those cross-sections cannot see: high-frequency regional adoption responding to repeated frontier shocks, with the response governed by pre-existing structure. A second literature builds occupational exposure measures for AI and generative AI (Felten et al., 2021; Felten, Raj, and Seamans, 2023; Eloundou et al., 2023; Webb, 2020); these measures serve here as pre-determined treatment intensity in a regional-exposure design in the lineage of Autor,

Dorn, and Hanson (2013), and the results show that exposure is not destiny: the same exposure produces adoption only where the surrounding skill ecosystem can absorb it. Third, the paper gives an AI-era test of the task-based and regional-branching tradition, in which human capital transfers across related activities (Gathmann and Schönberg, 2010) and regions diversify into activities related to their existing portfolio (Neffke, Henning, and Boschma, 2011; Boschma, 2017; Hidalgo, Klinger, Barabási, and Hausmann, 2007): the connective-tissue measure is exactly that tradition's relatedness density (Balland, Boschma, Crespo, and Rigby, 2019), operationalizing the absorptive capacity of Cohen and Levinthal (1990) at the regional level, and its failure to be beaten by any sophistication measure extends the branching logic to the adoption of a new general-purpose technology at biweekly frequency. The convergence result puts a technology-adoption face on the breakdown of regional convergence under skill-biased technical change (Giannone, 2022; Barro and Sala-i Martin, 1992). Finally, Autor (2024) argues AI could rebuild middle-class work by letting adjacent workers perform expert tasks. The evidence here is compatible with that optimism but relocates its condition: the augmentation opportunity is realized where connective tissue is dense, and since worker-level usage style is uniform across space, the binding regional margin is adoption itself, precisely the margin pre-existing skill structure controls.

Section 2 describes the data and defines the skill measures. Section 3 lays out the three designs. Section 4 presents results in order of evidential strength: prediction and persistence (4.1), the absorption margin (4.2), mechanisms (4.3), treadmill dynamics (4.4), the typology and worker-side validation (4.5), and a consolidated robustness and threats assessment (4.6). Section 5 concludes.

2 Data

2.1 BTOS adoption series

The BTOS surveys roughly 200,000 firms per biweekly panel, representative of nonfarm employer businesses, and publishes response estimates for 50 states + DC + Puerto Rico and the 25 largest MSAs. The core AI question, "In the last two weeks, did this business use Artificial Intelligence (AI)…", exists in two wording regimes which I *never splice*: regime 1 ("…in producing goods or services", collection periods 2023-09-11 through 2025-11-16, 54 published biweeks) and regime 2 ("…in any of its business functions", 2025-11-17 onward). Measured adoption jumps discontinuously at the wording change (Figure 1), which is why all event analysis is confined to regime 1 and the Nov-2025 release cluster (Gemini 3, GPT-5.1, Opus 4.5) is excluded from the event calendar ex ante. The regime-1 history is published in the Census archive file *AI Core Questions (original wording)*; regime-2 data come from the current cumulative releases (Appendix A).

Coverage and noise. A pre-registered coverage audit (experiment log, exp001) found the regime-1 AI cell published for 86% of state×biweek cells and 84% of MSA cells after disclosure suppression; three geographies (Alaska, North Dakota, San Antonio) have no published regime-1 cells and drop, and five more are thin (St. Louis, Wyoming, Mississippi, Charlotte, Vermont publish 18–33% of

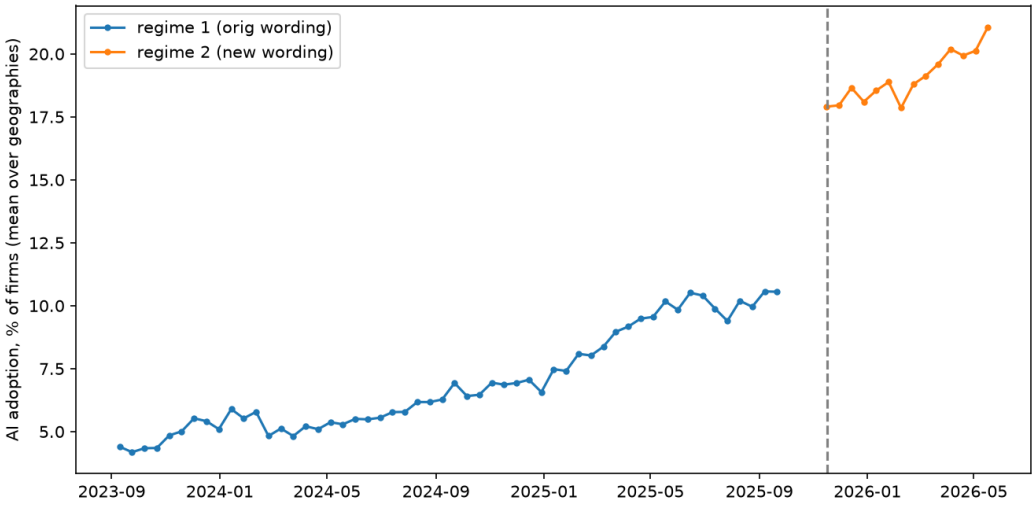


Figure 1: BTOS AI adoption, mean across geographies, by question regime. The dashed line marks the 2025-11-17 wording change; the two series are never spliced. Shares of firms reporting AI use (percent), unweighted mean over the 74 published geographies.

biweeks). Unit response averages 13.7%, drifting −5.5pp over the sample. The aggregate series is smooth, but the within-geography series is noisy (median within-geography first-difference s.d. 1.85pp), which motivates the sustained-threshold definition of breakout (Section 4.2) and the smoothing robustness reported below. Suppression is concentrated in small states, a fact revisited when interpreting the typology (Sections 4.5 and 4.6); cell-level detail is in Appendix A.

2.2 Measuring skill complexity and connective tissue

The Howell et al. (2026) panel provides 201 county-year measures, 2010–2024, for 3,194 counties, derived from 433.6M Lightcast postings. Two of its measures play the lead roles in this paper, and because the results turn on the distinction between them, I define both explicitly.

Both start from revealed comparative advantage in skills. Let $q_{r,s,t}$ be the share of region r's posting skill mentions in year t accounted for by skill s. Region r has revealed comparative advantage (RCA) in skill s when it mentions the skill more intensively than the nation does:

$$\mathrm{RCA}_{r,s,t} \;=\; \frac{q_{r,s,t}}{q_{\mathrm{US},s,t}} \;>\; 1. \tag{1}$$

Following Hidalgo et al. (2007), the relatedness of two skills is their tendency to be mastered together: the conditional probability that a region holds RCA in one skill given RCA in the other, taking the symmetric minimum,

$$\varphi_{s,s',t} \;=\; \min\Big\{ \Pr\big(\mathrm{RCA}_{s,t} \,\big|\, \mathrm{RCA}_{s',t}\big),\; \Pr\big(\mathrm{RCA}_{s',t} \,\big|\, \mathrm{RCA}_{s,t}\big) \Big\}, \tag{2}$$

with probabilities computed across regions. *Relatedness density*, the measure I call the skill ecosys-

tem's **connective tissue**, is the average closeness of the skills a region does *not* yet command to the skills it does (Balland et al., 2019): writing $R_{r,t}$ for the region's RCA skill set and $N_{r,t}$ for its complement,

$$\text{Density}_{r,t} = \frac{1}{|N_{r,t}|} \sum_{s \in N_{r,t}} \frac{\sum_{s' \in R_{r,t}} \varphi_{s,s',t}}{\sum_{s' \neq s} \varphi_{s,s',t}}. \tag{3}$$

A dense region is one whose next capabilities sit close to its current ones. In the economic-geography tradition this is the variable that predicts which activities a region acquires next (Neffke et al., 2011; Boschma, 2017), the regional counterpart of task-specific human capital that transfers across related activities (Gathmann and Schönberg, 2010), and a direct operationalization of absorptive capacity (Cohen and Levinthal, 1990); Section 4.3 verifies the skill-acquisition property in this panel. *Complexity*, by contrast, measures the sophistication of what a region already commands: the economic complexity index (ECI) of Hidalgo and Hausmann (2009) computed over the region×skill RCA matrix (the eigenvector associated with its second-largest eigenvalue), standardized within year. Complexity asks how rare and sophisticated a region's portfolio is; density asks how connected it is to what lies nearby. The two are correlated but conceptually and, as Section 4.2 shows, empirically distinct.

Counties aggregate to states and to the 25 BTOS MSAs (2023 CBSA delineations) via posting-weighted means (an employment-weighted alternative gives the same results, Section 4.6), and both measures are z-scored across the estimation sample, so every coefficient on them reads per one standard deviation. Skill structure enters every specification *lagged*: adoption in biweek b is explained by structure measured in calendar year $y(b) - 1$, capped at the panel's 2024 endpoint. I never use posting data from 2025 onward, post-ChatGPT postings are themselves shaped by AI-driven labor demand, which would put the outcome in the regressor; an assertion in the build pipeline enforces this.

2.3 Occupational exposure

$\text{Exposure}(g) = \sum_o \text{EmpShare}(g, o) \cdot \text{Intensity}(o)$, with May-2023 OEWS detailed-occupation employment shares (May-2022 as robustness) and three pre-registered intensity vectors: the Felten et al. (2021) AIOE score (z-scored; the primary, "general" domain), the Eloundou et al. (2023) human-β LLM exposure share, and the employment share of computer occupations (SOC 15-12xx; the "coding/agentic" domain). These occupational exposure measures descend from the task-based tradition (Autor, Levy, and Murnane, 2003; Webb, 2020; Felten et al., 2023); they enter not as impact forecasts but as pre-determined treatment intensity, a shift-share with fixed pre-period shares in the sense of Goldsmith-Pinkham, Sorkin, and Swift (2020). Cross-geography dispersion of the primary measure: s.d. 0.108, range −0.23 to 0.55 (DC). All inputs predate every outcome period.

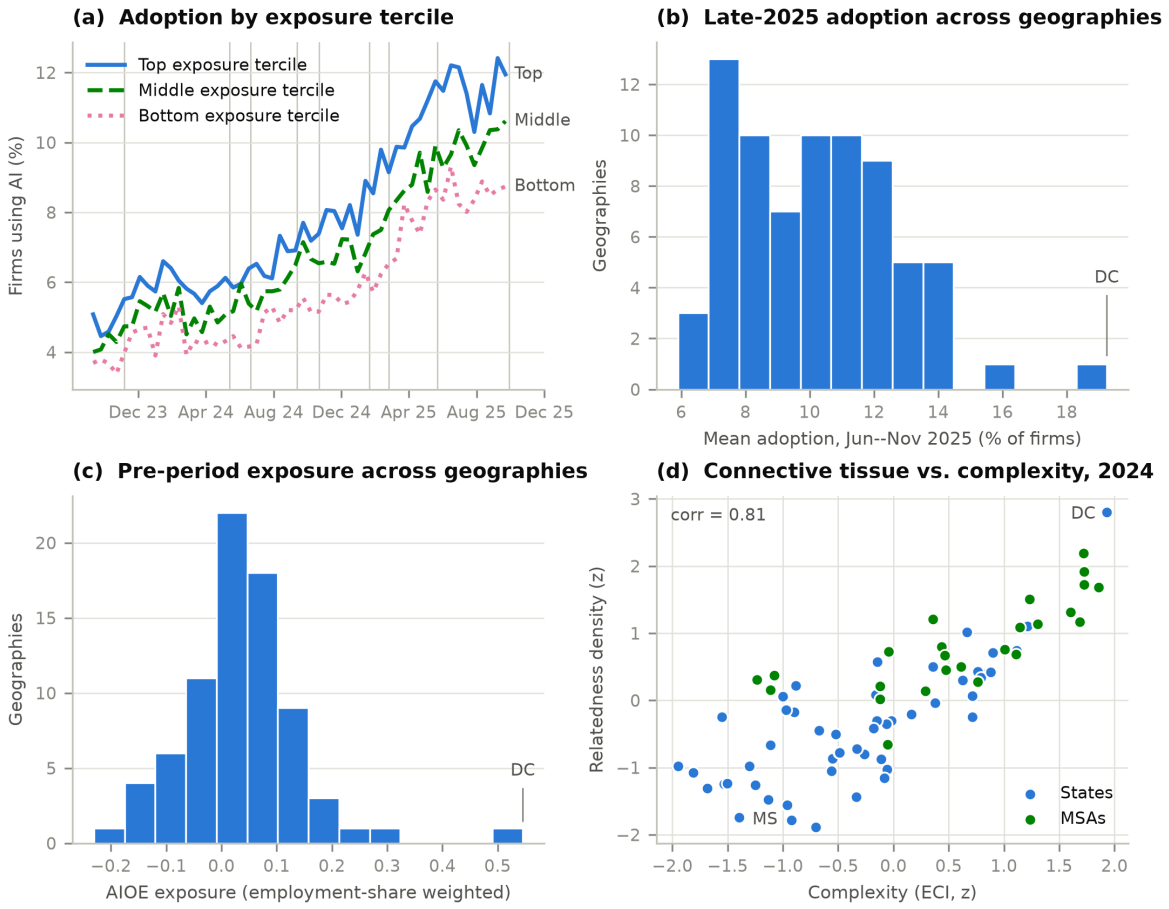


Figure 2: Descriptive patterns. (a) Mean BTOS AI adoption (percent of firms) by tercile of pre-period AIOE exposure, unweighted across published geographies, regime 1; vertical hairlines mark the nine frontier releases; the top-minus-bottom gap grows from 1.2pp (first six biweeks) to 2.9pp (last six). (b) Distribution of mean adoption, June–November 2025, across the 74 published geographies. (c) Distribution of pre-period AIOE exposure (employment-share weighted; s.d. 0.108). (d) 2024 relatedness density against complexity (both standardized), states and the 25 BTOS MSAs; correlation 0.81.

2.4 Estimation sample

Regime-1 panel: 54 biweeks × 77 geographies (50 states + DC + PR + 25 MSAs, net of full suppression), pooled with geography-type×time effects; the MSA layer adds precision and is never used as a standalone $N = 25$ cross-section. The 25 MSAs are geographic subsets of the pooled states; geography-type×biweek effects absorb level differences between the layers, and Section 4.6 shows the results are unchanged dropping the MSA layer entirely or clustering by parent state. Estimation N = 2,877 geography-biweeks in the distributed-lag specification (2,906 in the stock specification, which loses no lag periods). Appendix Table 14 reconciles every estimation sample in the paper.

The raw patterns. Figure 2 shows the raw material behind every estimate. Panel (a) plots mean adoption by tercile of pre-period exposure: the top-versus-bottom gap widens from 1.2 to 2.9 percentage points over regime 1, so the divergence the event design measures is visible in the raw series before any regression. Panel (b) shows the outcome's cross-section: late-2025 adoption spans roughly 6 to 19 percent of firms across geographies, with Washington DC the far right tail. Panel (c) shows the treatment's cross-section: exposure is tightly distributed with a long right tail (DC at 0.55 against a standard deviation of 0.108), which is why the leverage diagnostics of Section 4.6 matter. Panel (d) shows the two skill measures: relatedness density and complexity are strongly correlated across regions (0.81) yet empirically separable, and the horse race of Section 4.2 exploits exactly the regions where they disagree; DC is high on both margins, Mississippi low on both.

3 Empirical Design

3.1 Baseline specification and identification

The treatment is continuous: regions differ in pre-determined occupational exposure to frontier AI, and each release raises the stock of available frontier capability. Writing $K(b)$ for the cumulative capability stock at biweek b (the running sum of release intensities from the calendar in Section 3.2), the baseline specification is a generalized difference-in-differences in which the exposure gradient of adoption grows with the capability stock:

$$\text{adoption}_{g,b} \;=\; \beta\,\big[K(b)\times \text{Exposure}_g\big] \;+\; \alpha_g \;+\; \lambda_{type(g),b} \;+\; \varepsilon_{g,b}, \tag{4}$$

with geography fixed effects, geography-type×biweek fixed effects, and standard errors clustered by geography. The "treated" units are more-exposed regions, the "controls" are less-exposed regions in the same biweek and geography type, and β is the treadmill gradient: how much the adoption gap between them widens per unit of accumulated capability. The identifying assumption is that of any difference-in-differences with a continuous treatment: *absent frontier releases, high- and low-exposure regions would have evolved in parallel within geography-type×biweek cells.* Two testable implications discipline it: exposed regions must not diverge *before* releases (event-study leads), and no divergence should appear at dates when no release occurred (placebos); both hold (Sections 4.4 and 4.6).

The transparent event-study complement stacks windows around the three temporally isolated releases (GPT-4o, o1, DeepSeek-R1; selected ex ante), in the spirit of Cengiz, Dube, Lindner, and Zipperer (2019):

$$\text{adoption}^{(e)}_{g,b} \;=\; \sum_{\substack{\tau=-8\\ \tau\neq -1}}^{+12} \delta_\tau\,\mathbf{1}\big[b-b_e=\tau\big]\times \text{Exposure}_g \;+\; \alpha_{g,e} \;+\; \lambda_{b,e} \;+\; \varepsilon_{g,b,e}, \tag{5}$$

with geography×event and biweek×event effects, the biweek before the release omitted. The leads

$\{\delta_{-8}, \ldots, \delta_{-2}\}$ are the pre-trend test.

Because releases arrive every three to four months and adoption responds with a lag, discrete event windows overlap; the distributed-lag form traces the response *shape* directly and is the paper's evidence on timing:

$$\text{adoption}_{g,b} = \sum_{\ell=0}^{12} \beta_\ell \big[\Delta K_{b-\ell} \times \text{Exposure}_g\big] + \sum_{\ell=0}^{12} \theta_\ell \big[\Delta K_{b-\ell} \times \text{Exposure}_g \times \text{Skill}_{g,y(b)-1}\big] + \alpha_g + \lambda_b + \varepsilon_{g,b} \tag{6}$$

with the same fixed effects and clustering (shift-share-robust Adão, Kolesár, and Morales, 2019; Borusyak, Hull, and Jaravel, 2022 inference in Section 4.6). The moderation coefficients θ_ℓ ask whether, at the same workforce exposure, regions with deeper lagged skill bases absorb the same capability jump into more adoption; Skill stands for lagged relatedness density and complexity, entered together. One property of equation (6) matters for everything that follows: with overlapping events and collinear lags, the *sum* of the lag coefficients is a valid sign and timing statistic but a poor magnitude estimator, and Section 4.4 shows that it overstates the long-run gradient roughly fifteen-fold relative to equation (4) and the cross-section. Magnitudes in this paper therefore always come from the stock and cross-sectional estimators. The ex-ante falsifiability commitments: a response peaking on impact indicates anticipation or hype confounding; no response within 12 months casts doubt on the capability-adoption link itself.

3.2 Shock calendar (fixed ex ante of estimation)

Table 1 lists the nine events, fixed from third-party sources (release announcements, benchmark publications, API price sheets) before estimation, each scored 0–1 on three domains (coding/agentic, general/multimodal, price-access). The memo also fixed the capability index construction, the expected lag range (1–6 months), and five placebo dates (Appendix B). The domain intensities are coarse hand-set judgments, so Section 4.6 shows the headline gradient is invariant to replacing them with a binary release indicator, benchmark-anchored deltas, or 500 random perturbations; and because release dates are public facts, the design's credibility rests on the rule-based event list and the placebo battery rather than on the memo's timestamp.

3.3 Three designs, and what machine learning does here

The event machinery above is one of three designs, and the paper's findings do not all lean on it. The *prediction protocol* (Section 4.1) needs no events at all: models are trained with a geography deleted, or with all post-2024 data deleted, and judged purely out of sample. The *moderation design* asks who absorbs a given shock (the θ terms above, plus event-free hazard, adjustment, and convergence analogues). The *event design* (equations 4 and 5) carries the causal-dynamics reading and is, by construction, the most demanding; the findings are ordered accordingly.

Machine learning enters in exactly two places, and the core identification is ordinary panel regression in both headline equations. First, the out-of-sample predictions use a gradient-boosted

Table 1: Event calendar and domain intensities (fixed ex ante of estimation)

Date	Event	Coding/agentic	General/multimodal	Price-access
2023-11-06	GPT-4 Turbo / DevDay	0.50	0.75	0.75
2024-05-13	GPT-4o free multimodal	0.25	1.00	1.00
2024-06-20	Claude 3.5 Sonnet	0.75	0.50	0.25
2024-09-12	o1 reasoning models	0.75	0.50	0.00
2024-10-22	Claude 3.5 + computer use	1.00	0.25	0.00
2025-01-20	DeepSeek-R1 price collapse	0.50	0.25	1.00
2025-02-24	Claude 3.7/Code + GPT-4.5	1.00	0.50	0.25
2025-05-22	Claude 4	1.00	0.50	0.25
2025-09-29	Claude Sonnet 4.5	1.00	0.25	0.25
Excluded ex ante: Nov-2025 cluster (coincides with the BTOS wording break).				

Notes: Intensities are coarse hand-set ex-ante judgments from benchmark coverage and price sheets, fixed in the specification memo (Appendix B). $\Delta K(b)$ sums the mean-domain intensity of events whose release date + 1 day falls in collection window b; total K at the end of regime 1 is 5.0 (mean event size 0.56). Every release bundles capability, price, and product; events are interpreted as *composite frontier shocks* throughout.

learner (with an elastic net as a linear benchmark) because the estimand there is predictive accuracy, and holdout discipline, not functional form, is what protects it. Second, I re-estimate the skill-moderation result with a doubly-robust double machine-learning estimator on the geography cross-section (Section 4.2): it guards the moderation against functional-form and selection-on-observables concerns the fixed-effects design does not address, and it is valid there precisely because the long difference removes the panel structure that makes off-the-shelf causal ML inapplicable to shift-share event designs. Estimator details, nuisance models, and tuning are in Appendix C.

4 Results

4.1 The geography of adoption is pre-determined

Two holdouts fixed ex ante discipline the claim that adoption geography is encoded in pre-existing structure. *Temporal:* models trained only on pre-2025 geography-biweeks (features: 2024-vintage skill measures, 2019→2024 growth, three exposure measures, time) predict 2025 regime-1 cells with a root-mean-squared error of 3.01 (elastic net) and 3.44 (gradient boosting) against 3.63 for naive persistence (carrying forward each geography's last observed pre-2025 value; $n = 1{,}158$ test cells). The model beats inertia, the minimal bar any structural claim must clear. *Leave-one-geography-out:* deleting each geography entirely and predicting its mean June–November 2025 adoption from skill structure alone yields a correlation of 0.63 between actual and predicted (RMSE 2.15 against a cross-geography standard deviation of 2.49). Squaring the correlation, pre-period labor-market structure predicts on the order of forty percent of the out-of-sample variance in where adoption ended up, with the caution that this is estimated across roughly fifty state units; the claim is carried by the state layer (within-state correlation 0.63), while the 25-MSA layer is noisier (0.39) and serves to add precision to the pooled fit rather than to carry the prediction.

Table 2: Vintage horse race: which year's skill base predicts 2023–25 adoption?

Skill vintage	Density coef.	(s.e.)	ECI coef. (s.e.)	Within-R^2
2024	26.45	(8.23)	−0.49 (0.28)	0.202
2022	21.36	(7.52)	−0.49 (0.28)	0.179
2019	21.68	(7.93)	−0.60 (0.29)	0.174

Notes: adoption(g, b) (BTOS share of firms using AI, percentage points, regime 1) on vintage-V relatedness density and ECI plus AIOE exposure, biweek fixed effects, clustered by geography; 76 geographies, $N = 2{,}877$ geography-biweeks (density t: 3.21, 2.84, 2.73; ECI t: −1.76, −1.76, −2.07; exp005). Growth terms (2019→2024 changes), added to the 2024 specification, are individually insignificant (ECI growth $t = 1.51$; density growth $t = -0.47$).

The structure doing the predicting is old. Table 2 re-estimates the determinants regression (adoption on skill structure and exposure, biweek effects, clustered by geography; conditional associations, no causal claim) using the 2024, 2022, and 2019 vintages of the same skill measures. The 2019 skill base, measured before COVID, before remote work reshuffled posting geography, and three years before ChatGPT, predicts 2023–25 adoption nearly as well as the 2024 base: the within-R^2 falls only from 0.202 to 0.174, and the density coefficient is stable (26.5 versus 21.7, both $t > 2.7$). Adding 2019→2024 growth terms to the 2024 specification adds nothing (both $|t| < 1.6$). *Capability memory is long*: regional absorption capacity is a slow-moving structural endowment, not something regions assembled once ChatGPT arrived.

Persistence has a converse: laggards are not catching up on their own. Unconditional β-convergence across regions is absent (initial-level coefficient +0.04, $t = 0.19$); convergence appears only *conditional* on skill structure, the signature of regions converging to different steady states indexed by their fundamentals. The gaps documented next should therefore be expected to persist.

4.2 Connective tissue, not complexity, governs absorption

Regions convert capability into adoption in proportion to the connective tissue of their skill ecosystem, and not at all in proportion to its sophistication. The magnitude is easiest to see in the event-free cross-section: conditional on exposure and baseline adoption, one standard deviation of lagged relatedness density predicts about 1.1 percentage points of additional adoption growth over the regime-1 window (+1.17, $t = 2.34$ in the convergence regression; +1.11, $t = 3.52$ in the mechanism cross-section of Section 4.3), and it more than doubles the hazard of adoption breakout (Cox hazard ratio 2.40, $z = 2.5$). Complexity, entered alongside density, predicts nothing anywhere.

The event design says the same about *who absorbs the shocks themselves*. Table 3 reports the summed moderation coefficients from equation (6): holding workforce exposure fixed, regions one standard deviation deeper in lagged relatedness density convert the same capability jump into significantly more adoption under every exposure measure (t = 3.0, 4.5, 4.2), while complexity enters negative and insignificant in all three columns. The table reads as sign and significance evidence; the stock-form analogue of the moderation (density interacted with the capability stock)

Table 3: Skill moderation of the capability response: relatedness density, not complexity (distributed-lag sums; sign and significance)

	AIOE (primary)	LLM-β	Coding share
Treadmill $\sum_\ell \beta_\ell$	24.29	87.16	94.00
	(10.93)	(69.48)	(83.39)
	$[t = 2.22]$	$[t = 1.25]$	$[t = 1.13]$
$\sum_\ell \theta_\ell$: complexity (ECI, lagged, z)	-14.98	-6.09	-67.07
	(13.89)	(4.79)	(52.58)
	$[t = -1.08]$	$[t = -1.27]$	$[t = -1.28]$
$\sum_\ell \theta_\ell$: relatedness density (lagged, z)	34.58	23.75	201.01
	(11.69)	(5.26)	(47.70)
	$[t = 2.96]$	$[t = 4.51]$	$[t = 4.21]$
N (geography-biweeks)	2,877	2,877	2,877

Notes: Equation (6) with geography and geography-type×biweek fixed effects; standard errors (clustered by geography) of the summed linear combination in parentheses. Outcome: BTOS share of firms using AI in producing goods or services (percentage points), regime-1 panel, Sep 2023–Nov 2025. Skill moderators are z-scored and lagged one calendar year (capped at the 2024 panel endpoint). Cross-geography exposure dispersion: AIOE s.d. 0.108, LLM-β 0.022, coding 0.015; the alternative-measure coefficients are larger because their regressors vary less. The distributed-lag sums establish sign and significance but overstate magnitudes ≈15-fold through overlapping events and collinear lags (Table 6); magnitude statements in the text come from the cumulative-stock and event-free estimators. Full lag profiles in Appendix Table 9.

is sign-consistent and implies a modest 0.03 to 0.05 percentage points of extra absorption per release per standard deviation of density, though it is imprecise ($t = 1.6$ with complexity included, 2.4 without).

The ordering is not an artifact of how sophistication is operationalized. Substituting Tacchella–Pietronero fitness, skill entropy, average ubiquity, or a skill Herfindahl for the ECI leaves density's moderation positive and significant against every alternative ($t = 3.0$–4.3), no sophistication measure earns a positive coefficient conditional on density, and the result is unchanged under level, log, and logit growth specifications (Appendix Table 15). The lone positive complexity-family coefficient, average ubiquity, measures the *commonness* of a region's skills, so its sign reinforces rather than overturns the reading, and it reverses under the doubly-robust estimator.

A doubly-robust check corroborates the moderation with none of the event design's structure. Treating pre-period exposure as the treatment and 2023–25 adoption growth as the outcome on the geography cross-section, a double machine-learning estimator with random-forest nuisance models recovers a null *average* effect of exposure (-1.23, 95% CI $[-5.17, 2.71]$), unsurprising given the modest treadmill, but a conditional effect that rises with skill density: the slope of the treatment effect on density is +3.1 (95% CI [1.2, 5.0]), the doubly-robust analogue of the moderation, on $n = 62$ geographies. That an estimator with entirely different identifying assumptions recovers the same density-buys-absorption pattern is the strongest single corroboration; because a single-modifier specification cannot discriminate among correlated moderators, the horse race above, with density co-included, remains the discriminating test (estimator details in Appendix C).

Table 4: Domain-matched capability shocks: $\sum_\ell \beta_{d\ell}$ by domain

Domain shock ΔK_d	Matched exposure	$\sum_\ell \beta_{d\ell}$	(s.e.)	t
coding	`exp_coding`	237.62	(33.68)	7.05
general	`exp_aioe_z`	-26.90	(19.19)	-1.40
price	`exp_llm_beta_human`	137.18	(124.80)	1.10

Notes: all three domain interactions estimated jointly (39 regressors) on the regime-1 panel; geography and geography-type×biweek fixed effects, clustered by geography. Outcome: BTOS share of firms using AI in producing goods or services (percentage points), regime-1 panel, Sep 2023–Nov 2025. Domain intensities and exposure pairing fixed ex ante (Appendix B). Under a common AIOE-exposure×time trend the coding response falls only to 198.5 ($t = 4.80$); see Table 7.

Two further event-free lenses agree, with less power. In a discrete-time hazard of adoption breakout (three consecutive biweeks at ten percent or more; 33 of 74 geographies broke out within regime 1), density raises the hazard and complexity does not (Appendix Table 11); the Cox variant sharpens the density effect to the hazard ratio of 2.40 quoted above, invariant to breakout thresholds from 6 to 14 percent and carried by the state layer. And a per-geography partial-adjustment rate toward the frontier (0.71 per biweek raw, roughly 0.20 on a smoothed series that removes sampling mean-reversion) has density as its only positive structural correlate (0.161, $t = 1.69$; complexity $t = -0.44$). Each lens is individually underpowered; their agreement, with the moderation and the prediction results, is the point.

4.3 Mechanisms

Three questions unpack the absorption margin: through which workforces does frontier capability propagate, what does a dense ecosystem actually do, and which margins do *not* vary. Table 5 collects the new evidence.

What propagates. Pairing each domain's capability shocks with its matched exposure (fixed ex ante) shows the 2023–25 episode was *coding-tool-led*: the response concentrates entirely in coding/agentic shocks interacted with coding-intensive workforces ($t = 7.05$), while general/multimodal and price-access shocks add nothing conditional on it (Table 4). This domain concentration also does the heavy lifting for identification, because the coding response survives the differential-trend control that absorbs the pooled gradient (Section 4.6).

What a dense ecosystem does. Relatedness density is defined by the promise that a region's next capabilities sit close to its current ones. The county panel bears the promise out on the conversion margin: in county-year regressions with county and year fixed effects, a standard deviation of lagged density predicts a significantly higher rate at which the skills a county does *not* yet command convert into revealed strengths the following year ($t = 5.1$), and a larger year-over-year reshuffling of the county's skill-demand profile ($t = 3.9$), while complexity predicts neither (Table 5, Panel A). The volume of acquisition, by contrast, does not rise with density (dense, diversified counties have proportionally fewer gaps left to fill, and count models with base-size controls load negative), so the

Table 5: Mechanism evidence: what relatedness density does, and does not, predict

	Density coef.	(s.e.)	[t]	ECI t	N
Panel A: county-year skill acquisition on lagged density (z), Howell panel 2011–2024					
Entries per at-risk skill (not yet RCA)	0.201	(0.040)	[5.06]	-6.66	41,447
Skill-profile change (cosine distance)	0.006	(0.002)	[3.91]	-4.90	41,447
Entries per existing RCA skill[a]	-0.125	(0.015)	[-8.08]	-0.63	41,447
Panel B: 2023–25 adoption growth and pre-period acquisition, 77 BTOS geographies					
Growth on acquisition rate 2020–23 (z)	-0.603	(0.288)	[-2.09]	–	62
Growth on acquisition rate 2015–19 (z, placebo)	-0.325	(0.284)	[-1.15]	–	62
Density (z), controlling acquisition	1.085	–	[2.49]	–	62
Acquisition (z), controlling density	-0.023	(0.401)	[-0.06]	–	62
Panel C: worker-side usage style (AEI augmentation share), US states					
Augmentation share on density (z)	0.532	(1.324)	[0.40]	-0.63	49
Raw correlation with density	0.169	–	–	–	49

Notes: Panel A: county-year regressions of skill-acquisition outcomes on lagged relatedness density (z) and lagged complexity (ECI, z), with log postings, county and year fixed effects, posting weights, standard errors clustered by state; county-years with fewer than 50 postings and pseudo-county rows excluded (3,110 counties). At-risk skills = skills present in the county's postings without revealed comparative advantage. [a]The per-RCA-skill rate is saturation-confounded (diversified counties have fewer skills left to enter relative to their base) and is shown for completeness; the volume of acquisition does not rise with density (PPML with base-size controls: $t = -16.0$). Panel B: cross-sectional regressions of 2023–25 adoption growth (late minus early regime-1 mean, percentage points) on posting-weighted geography-level acquisition rates, controlling AIOE exposure, baseline adoption, and an MSA indicator; HC1 standard errors. Panel C: state augmentation share of Claude conversations (Anthropic Economic Index, Aug 2025, percent) on 2024 skill density (z), complexity (z), and exposure; HC1. The augmentation share is compressed across states (class means range 50.7–51.9%): regional variation in AI is in whether firms adopt, not in the automation-versus-augmentation style of use.

tissue works by making nearby skills reachable, not by making regions churn more. Firm self-reports corroborate the applicability side of this channel: among non-adopting firms in the 2026 BTOS AI supplement (U.S. Census Bureau, 2026), the dominant stated reason, "AI is not applicable to this business," falls with state skill density (correlation −0.34), while explicitly skill-based reasons are flat across states (≈11% everywhere), so the depth margin is identified from outcomes rather than from firms' self-diagnoses.

What does not vary. Two informative nulls bound the mechanism. First, pre-period acquisition *rates* do not predict 2023–25 adoption growth (Table 5, Panel B), and density's gradient is untouched by controlling for them: what absorbs AI is the accumulated relatedness stock, not recent skill churn. Second, worker-side usage *style* is geographically uniform: state augmentation-versus-automation shares of Claude conversations sit in a band of barely one percentage point and are uncorrelated with density (Panel C). The regional variation that matters is in whether firms adopt at all, the extensive margin the adoption series measures, not in how AI is used once adopted.

What remains unseen are worker flows and firm entry: no worker-level transition data or establishment-entry margin enters this paper, and the public skill panel is aggregate, so I cannot observe *which* skills enter (only their count and the profile shift). Whether connective tissue operates through workers moving into adjacent occupations, firms recombining incumbent workers, or entry of AI-complementary establishments is the natural next question for administrative

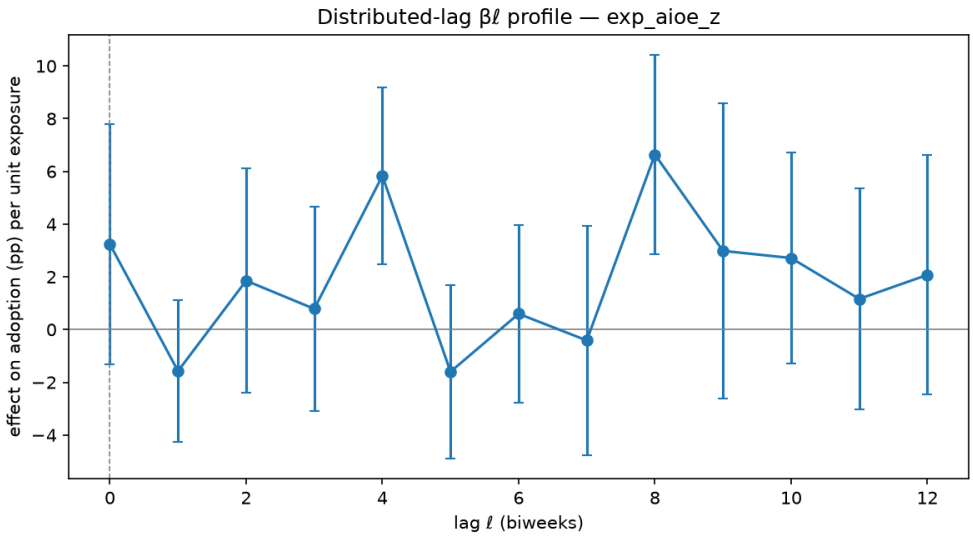


Figure 3: Distributed-lag profile β_ℓ (AIOE exposure), 95% CIs, from equation (6) on the regime-1 panel (N=2,877; clustered by geography). Lags 5–6 dropped by collinearity with event spacing. The response is absent on impact and builds one to four months after release.

worker-flow data.

4.4 Dynamics: the adoption treadmill

Frontier progress operates on this pre-determined geography as a treadmill: each release widens the gap between exposed and unexposed regions. The magnitude is modest and precisely signed. From the cumulative-stock specification (equation 4), the gradient is 1.56 per unit of capability ($t = 3.2$), which the cross-sectional long difference corroborates at 1.67 ($t = 5.1$; Table 6). Scaled by the cross-regional dispersion of exposure (s.d. 0.108) and the size of a typical release ($\Delta K = 0.56$): *each typical release widens the adoption gap between a one-standard-deviation-more-exposed region and the average region by roughly 0.1 percentage points*, about 0.3 points per year at the observed release cadence, and about 0.8 points cumulated over the nine releases of regime 1 (the widening is visible in the raw tercile series of Figure 2a). Against the aggregates, that cumulative wedge is roughly an eighth of the 6.2-point national adoption rise and a fifth of the late-2025 interquartile adoption gap across regions (3.8 points), a meaningful but not dominant re-sorting of the adoption distribution; consistent with McElheran et al. (2024), most adoption variation remains cross-sectional rather than shock-driven.

Timing behaves exactly as a diffusion process should. The response is near zero on impact (the ex-ante expectation, and evidence against hype or anticipation effects) and builds one to four months after release, inside the pre-registered one-to-six month window (Figure 3; Appendix Table 9). The stacked event study around the three isolated releases (Figure 4; Appendix Table 10) shows the same build-up, and its eight leads are individually insignificant (max $|t| = 1.20$): exposed regions were not already diverging before releases.

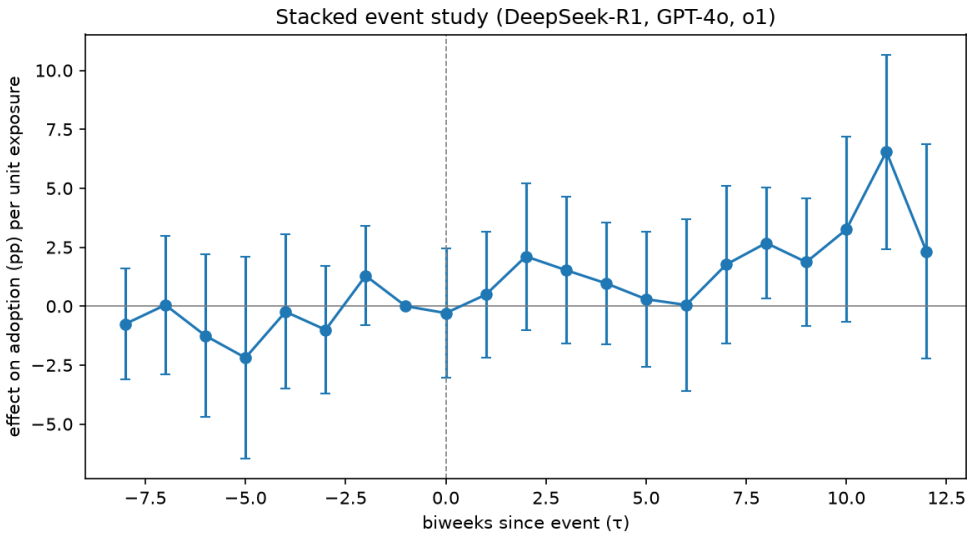


Figure 4: Stacked event study around the isolated events (GPT-4o, o1, DeepSeek-R1): δ_τ on AIOE exposure from equation (5), $\tau = -1$ omitted, 95% CIs, clustered by geography (N=3,343 stacked geography-biweeks). All eight leads are individually insignificant.

One reconciliation underlies every magnitude in this paper. The same long-run gradient can be estimated three ways, and Table 6 shows the distributed-lag sum (24.3) is roughly fifteen times the cumulative-stock (1.56) and cross-sectional (1.67) estimates, which agree closely with each other. The inflation is a collinearity artifact of overlapping events and oscillating individual lags: grouping the lags into windows collapses the sum monotonically (24.3 to 8.0 to 4.8) toward the stock estimate. I therefore quote the stock and cross-section estimators for magnitude throughout and keep the distributed lag only as evidence on sign and latency. The alternative exposure measures give the same qualitative pattern with noisier sums, reflecting their much smaller cross-regional dispersion.

The treadmill is the paper's most demanding claim, and two flags on it are argued in full in Section 4.6: with near-evenly-spaced releases the capability-step reading is only weakly separable from a smooth differential trend in the pooled data (the domain-matched coding response is what survives that control), and the level of the divergence, while precisely signed, is modest.

4.5 Readiness typology, policy margins, and worker-side validation

Figure 5 plots actual against predicted adoption; median splits define four types (29 frontier, 28 structurally lagging, 8 untapped readiness, 8 overperformers; full listing with coverage annotations in Appendix Table 16). The residual is the diagnosis: a region *below* its prediction is underperforming its own fundamentals, fixable in principle by diffusion policy, while a region at a low prediction is constrained by the fundamentals themselves. The largest underperformers are Mississippi (−4.3pp below prediction), New Jersey (−4.0), Los Angeles (−4.0), and Detroit (−3.7); the largest overperformers are Washington DC (+8.6; actual 23.1% against predicted 14.5%), Delaware (+5.6), Denver (+4.0), Wyoming and Vermont (+3.1, +3.0). These labels are read with two grains of salt

Table 6: Three estimators of the long-run treadmill gradient (per unit of cumulative capability K, AIOE exposure)

Estimator	coef	s.e.	t	N
Distributed-lag $\sum_\ell \beta_\ell$ (lag-by-lag)	24.29	10.93	2.22	2,877
lags windowed 0–12 (five windows)	8.04	4.37	1.84	2,877
lags windowed 0–26 (seven windows)	4.81	4.86	0.99	2,877
Cumulative-stock $K(b)\times$Exposure	1.56	0.49	3.21	2,906
Cross-section long-difference / total K	1.67	0.33	5.10	62

Notes: all rows target the per-unit-K widening of the exposed/unexposed adoption gap. Outcome: BTOS share of firms using AI in producing goods or services (percentage points), regime-1 panel, Sep 2023–Nov 2025. Distributed-lag and windowed rows estimate equation (6) (windowed variants group the lags as indicated) with geography and geography-type×biweek fixed effects, clustered by geography. The cumulative-stock row regresses adoption on $K(b)\times$Exposure with the same fixed effects and clustering. The cross-section row is the long difference of adoption (late minus early regime-1 mean) on exposure, HC1, divided by total $K = 5.0$. The lag-by-lag sum is inflated by overlapping events and collinear lags; grouping the lags collapses it monotonically toward the stock estimate, which is the magnitude quoted in the text (≈0.1pp per release per s.d. of exposure).

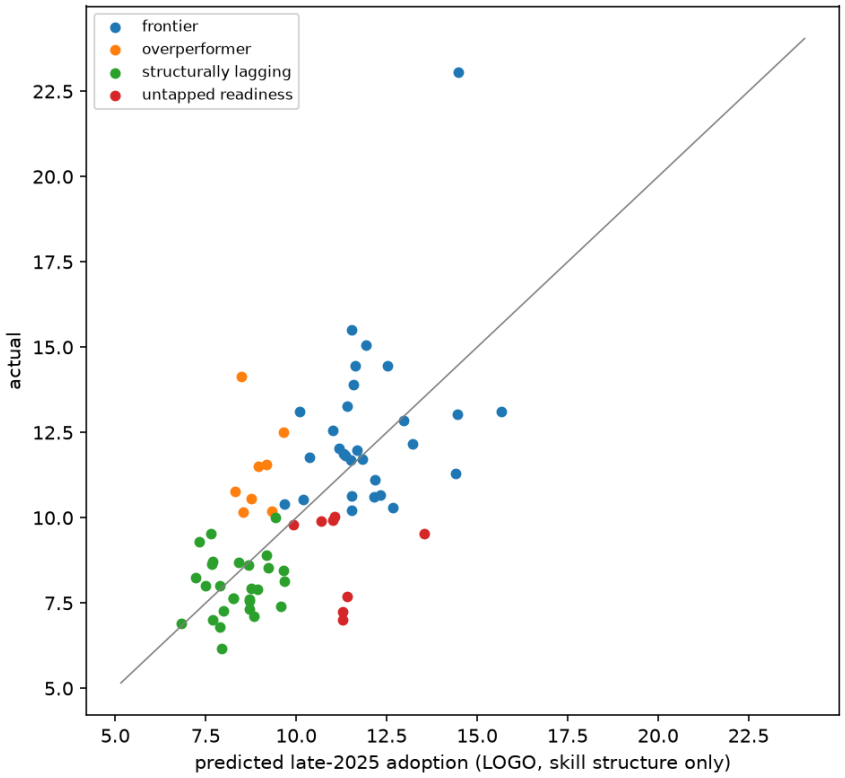


Figure 5: Actual vs. LOGO-predicted late-2025 adoption (percent of firms using AI, June–November 2025 mean). The 45-degree line is prediction = outcome; the vertical distance is the residual that defines the typology. Predictions from leave-one-geography-out gradient boosting on pre-period skill structure and exposure.

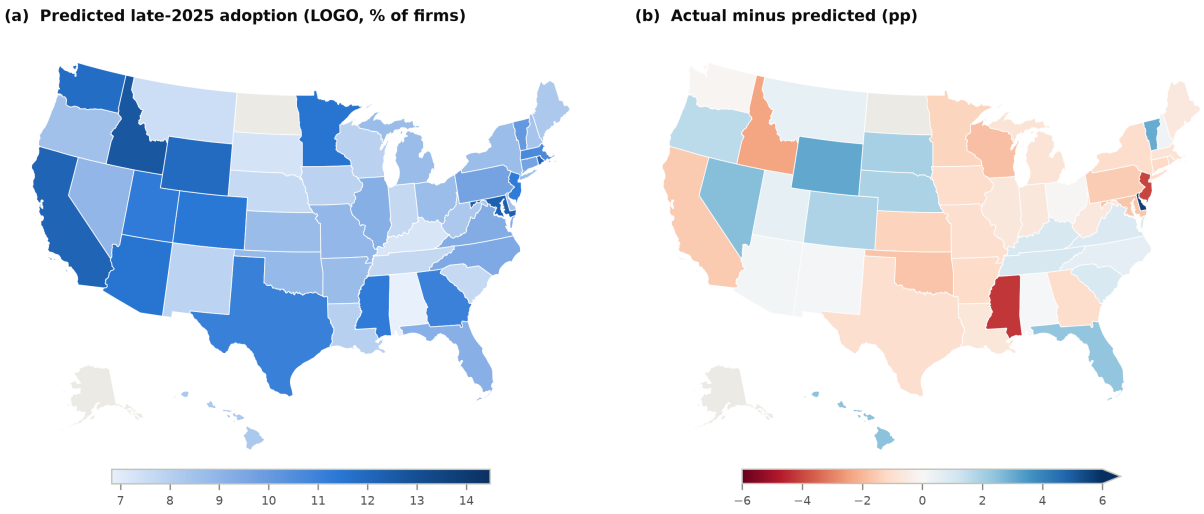


Figure 6: The predicted geography of adoption, and the residual that defines the typology. (a) LOGO-predicted late-2025 adoption (percent of firms) from pre-period skill structure and exposure alone; each state is predicted by a model that never saw it. (b) Actual minus predicted adoption: red states adopt below their own fundamentals (Mississippi −4.3pp, New Jersey −4.0), blue states above (Delaware +5.6, Wyoming +3.1). The diverging scale is capped at ±6pp; Washington DC (+8.6) exceeds the cap but is too small to render at map scale. Alaska and North Dakota (gray) publish no regime-1 AI cells; Puerto Rico and the 25 MSAs are not shown. States only, the layer that carries the prediction claim (Section 4.1).

developed below: residuals on roughly fifty units carry sampling noise, and several headline units are exactly the thin-coverage states where the BTOS cell is often suppressed. Figure 6 draws the same two objects on the map: the predicted geography of adoption, and the residual that defines the typology.

Each gap is decomposed by the region's *relatively* weaker fundamental, whichever percentile is lower: occupational *composition* (exposure percentile below the region's own relatedness-density percentile) versus *depth* (the converse: the occupations present but the surrounding ecosystem shallow). This is a within-region ranking, not an absolute deficit. Los Angeles, for instance, is tagged composition despite sitting at the 76th exposure percentile overall; it ranks composition only because its relatedness density ranks higher still (97th percentile), so its shortfall is best read as a near-pure residual on strong fundamentals. The genuinely depth-constrained cases, below-median on relatedness density, are Mississippi, New Jersey, and Detroit. The two margins map to different instruments, industry attraction versus long-horizon ecosystem building, and the vintage result warns that the depth margin moves on multi-year horizons.

Worker-side validation. The BTOS measures firm-side adoption. External validation comes from a fully independent measurement of state AI intensity: the Anthropic Economic Index (Handa, Tamkin, McCain, Huang, Durmus, et al., 2025), which records Claude usage per working-age capita

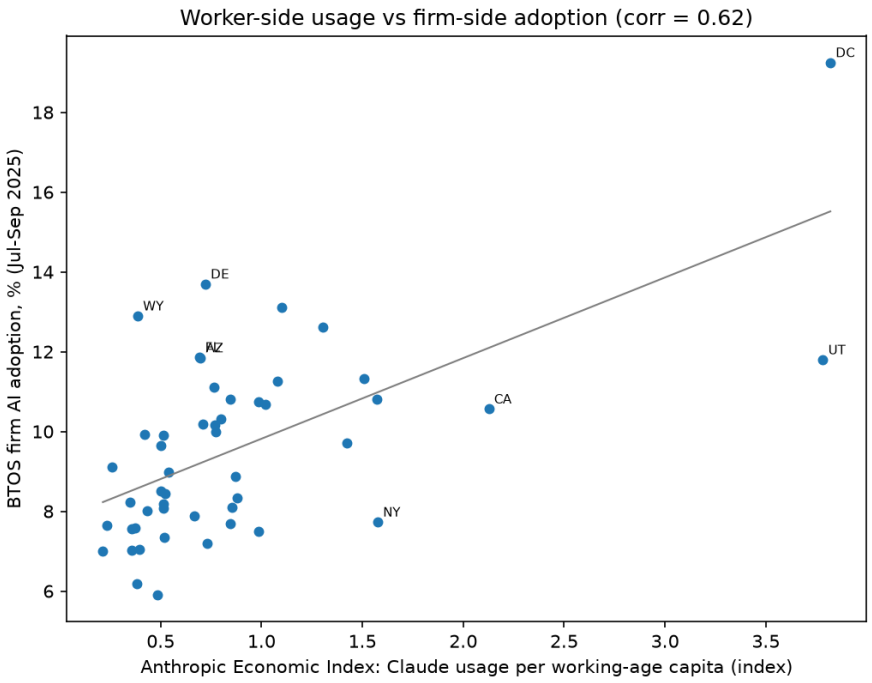


Figure 7: Worker-side Claude usage per working-age capita (Anthropic Economic Index, Aug 2025) against BTOS firm-side AI adoption (Jul–Sep 2025, percent of firms), US states. Correlation 0.62 ($n = 49$). Labeled states have wedge $|z(\text{usage}) - z(\text{adoption})| > 1.2$.

from millions of conversations (week of 2025-08-04).[1] Against BTOS state adoption averaged over July–September 2025 (bracketing the AEI window), the correlation is 0.62 (rank correlation 0.55, $n = 49$ matched states; Figure 7): two instruments with entirely different respondents, collection modes, and failure modes agree on the geography of AI intensity.

The *wedge* between the two measures is itself informative. States where worker usage runs furthest ahead of firm adoption (Utah, New York, California, Missouri, Wisconsin) are candidates for bottom-up diffusion: individual workers adopting before their employers formalize use. States where firm adoption runs ahead of worker usage (Wyoming, Delaware, Florida, Arizona, Vermont) include exactly the small-state *overperformers* of the typology: their firm-reported adoption exceeds both their skill fundamentals and their own workers' personal usage, suggesting their overperformance partly reflects business composition (e.g., registration-domicile effects in Delaware and Wyoming) rather than ecosystem strength, a caveat I attach to the typology rather than a contradiction of it.

4.6 Robustness and threats to validity

The three findings differ in evidential strength, and I state the ranking plainly: the prediction result is the most secure (it needs only out-of-sample validity, which holds), the density moderation next (robust across designs and measures, with two soft spots documented below), and the treadmill's

[1] The AEI is produced by Anthropic from its own platform data; I use it here strictly as an external cross-check on the BTOS geography, and the mechanism test of Section 4.3 uses only its augmentation/automation classification.

Table 7: Step versus trend: the pooled gradient weakens under an Exposure×time control, but the coding-domain response survives it

Specification	gradient / $\sum_\ell \beta_{d\ell}$	t
Pooled cumulative-stock gradient (per unit K)		
$K(b)$×Exposure	1.56	3.22
+ Exposure×time (linear)	2.08	1.30
+ Exposure×time (quadratic)	1.88	1.18
Domain-matched coding/agentic response ($\sum_\ell \beta_{d\ell}$)		
no time control	237.6	7.05
+ common Exposure×time	198.5	4.80
+ separate Exposure×time per domain	47.4	0.80

Notes: outcome is BTOS adoption (pp), regime-1 panel (N=2,906 stock rows; N=2,877 domain rows); geography and geography-type×biweek effects, clustered by geography. The domain rows are the coding/agentic cumulative distributed-lag sum interacted with coding exposure, estimated jointly with the general and price domains (both null throughout). The common-trend row adds a single AIOE-exposure×time term, the one that reduces the pooled gradient to $t = 1.30$; the per-domain row adds a separate exposure×time trend for *each* of the three domains (near-collinear over 54 biweeks).

causal-dynamics reading the most demanding. This section collects every threat and what survives it.

Step versus trend. With roughly evenly spaced events the capability stock is nearly linear in time, so the capability-*step* reading competes with "exposed regions simply trend up." The divergence is not merely a smooth exposure trend, but the pooled specification alone cannot prove it: adding an exposure×time term leaves the pooled gradient positive (2.08) while cutting its t to 1.30 (Table 7). What separates step from trend is the domain-matched design: the coding-domain response survives the very same common trend control that defeats the pooled gradient (its cumulative effect falls only from 237.6, $t = 7.05$, to 198.5, $t = 4.80$), because coding releases cluster late in the window and trace a capability path a single linear trend cannot absorb. Only a near-saturated specification with a separate trend per domain (three collinear trends over 54 biweeks) extinguishes it ($t = 0.80$), the limit of what the sample resolves. A broader covariate-trend battery sharpens this division of labor. On the state layer the pooled gradient is absorbed not only by the exposure trend but by income-, software-share-, or density-by-time trends (all $t \leq 1.4$), as it must be if exposed regions differ on slow-moving characteristics. The coding-domain response survives that entire battery: with density, complexity, and software-share trends added its sum is 166 ($t = 4.0$); adding the exposure trend as well, 167 ($t = 4.2$); and on the state layer with an income trend added too, 126 ($t = 2.5$), while the general-domain sums stay null throughout. The treadmill claim therefore rests on the surviving domain response, the clean leads, the placebo battery, and the event-free cross-section and hazard results; the pooled gradient is a magnitude summary, not the identification.

Table 8: Placebo battery: exposure gradient at pre-registered non-release dates

Placebo date	$\sum_{\ell=0}^{6} \beta_{\ell}^{placebo}$	(s.e.)	t
2024-03-11	-14.38	(7.61)	-1.89
2024-08-05	-4.13	(6.92)	-0.60
2024-12-09	2.59	(6.19)	0.42
2025-04-14	24.48	(11.67)	2.10
2025-07-28	8.42	(7.47)	1.13

Notes: Each row re-estimates the distributed-lag specification with a unit "capability jump" assigned to a pre-registered non-release biweek (exp000 §5; all ≥ 6 weeks from any true event), lags 0–6, AIOE exposure; geography and geography-type×biweek fixed effects, clustered by geography, regime-1 panel. Outcome: BTOS share of firms using AI in producing goods or services (percentage points), regime-1 panel, Sep 2023–Nov 2025. Under the identifying assumption the cumulative placebo effect is zero.

Placebos. At four of five pre-registered non-release dates the cumulative "effect" is statistically zero, and one placebo is *negative*, the opposite of what a secular tech-region drift would produce (Table 8). The one borderline placebo (April 2025, $t = 2.10$) sits seven weeks after the February-2025 release cluster, within the anticipated response window of a true event; re-running the battery controlling for the full true-event lag structure drops it to $t = 1.54$ while another placebo turns significantly negative (Appendix Table 13). Across the ten placebo estimates in both variants, two reject at the five percent level, one in each direction; under the null the chance that at least one of five placebos rejects is about 23 percent.

Negative controls and event dependence. Two non-AI outcomes measured identically in the BTOS (the share of firms reporting longer supplier delivery times, and the share reporting increased prices charged) provide negative controls, and they are informative about what the pooled design can and cannot do. Both load positively on the pooled capability-stock gradient (delivery delays $t = 2.8$), and both gradients vanish entirely under the same exposure-by-time control ($t = 0.4$ and $t = -1.2$): a pooled stock gradient can be manufactured by differential trends alone. What the negative controls lack, and adoption has, is everything else (the surviving domain response, the post-release timing, the clean placebos). Dropping any single release leaves the adoption gradient between 1.55 and 1.90 ($t \geq 3.1$ in all nine cases), so no one event drives it.

Nested geographies. The panel pools states with 25 MSAs that are subsets of those states, raising double-counting and cross-layer dependence concerns beyond what geography-type×biweek effects absorb. Neither drives the results: dropping the MSA layer entirely leaves the stock gradient at 1.68 ($t = 2.72$, states only), and clustering by parent state (50 clusters, each MSA assigned to its primary state) leaves inference intact ($t = 2.93$). The density result is likewise state-carried: re-estimating the determinants regression within layers gives $t = 3.2$ states-only, with the MSA

layer adding noise rather than signal ($t = 1.4$), which is why the prediction claim is stated at the state level.

Leverage. Washington DC is the highest-exposure unit. Dropping it lowers the stock gradient modestly to 1.20 ($t = 2.88$; DC and Puerto Rico together: 1.20, $t = 2.87$), and a binned scatter of the long-difference cross-section shows exposure-quintile means rising monotonically with DC sitting *on* the fitted line, not bending it (slope 8.5 with DC, 8.3 without; quadratic term $t = 1.1$; Appendix Figure 9).

Inference. Shift-share designs can over-reject (Adão et al., 2019). Four checks say this one does not. Randomization inference: reassigning exposure across geographies 500 times, no draw attains the observed gradient ($p < 0.002$; the no-moderator statistic sits 4.7 permutation standard deviations from zero). Wild-cluster bootstrap p-values agree ($p = 0.0003$ pooled, $p = 0.020$ states only; 9,999 draws). Two-way clustering by geography and biweek *tightens* the distributed-lag inference ($t = 2.62$). Analytical AKM occupation-level standard errors are an order of magnitude *smaller* than geography-clustered ones (0.03 versus 0.49 for the stock gradient), because occupation shares are highly diffuse (≈95 effective occupations, mean pairwise cosine similarity 0.91), so the share-induced correlation the correction targets is negligible; a Borusyak–Hull–Jaravel shock-level computation agrees (shock-level and region-level gradients coincide exactly at 8.50; shock-level s.e. 0.99 versus 1.65 conventional). Geography clustering is therefore the binding, conservative choice throughout. Computation details are in Appendix C, and the full specification grid (two-way clustering, smoothing, thin-geography exclusion, May-2022 exposure weights) is Appendix Table 12.

Shock-intensity invariance. Because the domain intensities are coarse ex-ante judgments, the robustness battery re-estimates the stock gradient under a binary release indicator ($t = 3.19$), benchmark-anchored intensities ($t = 3.26$), and 500 random perturbations of the domain scores (gradient positive in every draw); the inference and the cumulated magnitude do not move (Appendix C).

Soft spots of the moderation. The density moderation attenuates under outcome smoothing ($t = 1.6$) and thin-geography exclusion ($t = 1.1$): part of it is identified off small, noisy geographies, and its *magnitude* is the least settled object in the paper even though its sign is robust in every lens. Aggregation weighting is not the reason: posting-weighted and QCEW-employment-weighted county-to-region aggregation give the same result ($t = 3.2$ both), and only equal-county weighting, which hands a 5,000-person county the weight of Los Angeles County, collapses it ($t \approx 1$). The density relationship also reappears under the regime-2 wording ($t = 3.4$; Appendix E), so it is not an artifact of the regime-1 question.

Data quality and design-level caveats. BTOS is an experimental product: unit response averages 13.7 percent and drifts −5.5pp over the sample (Bonney et al., 2024); published rates are firm-weighted (employment-weighted checks await Census publication); and disclosure suppression is concentrated in small states, several of which (Mississippi, Wyoming, Vermont; 25–33 percent coverage, Appendix A) are headline typology cases, which is why Table 16 carries coverage annotations and I treat individual small-state residuals as noisy. Every release bundles capability, price, and product changes: I make no decomposition claim and interpret events as composite frontier shocks. All event analysis is confined to the regime-1 wording; the regime-2 series enters only the persistence appendix as it accrues. The raw partial-adjustment rate is noise-contaminated (0.71 raw versus roughly 0.20 smoothed), and the hazard rests on 33 events; both sharpen as biweeks accrue.

5 Conclusion

The geography of the AI economy is, to a first approximation, already written in pre-AI labor-market structure. Where adoption ended up in late 2025 was predictable from skill structure measured before the outcomes began (out-of-sample correlation 0.63 across states), the 2019 skill base predicts it nearly as well as the 2024 base, and laggards show no unconditional convergence. What converts frontier capability into adoption is the connective tissue of a region's skill ecosystem, the relatedness of its existing skills to adjacent capabilities, worth about one percentage point of adoption growth per standard deviation over the sample window, while the sophistication of the skill base adds nothing. Frontier releases then compound the pattern in real time: each typical release widens the exposed-unexposed gap by roughly 0.1 percentage points, arriving one to four months after release, though I am candid that this step-by-step reading is the paper's most weakly identified claim. The actionable residual is the set of regions adopting below their own fundamentals, and the composition-versus-depth decomposition says which instrument each needs; the vintage result cautions that the depth margin moves on multi-year horizons, so the returns to ecosystem building are real but slow. Whether connective tissue works through worker transitions, incumbent retraining, or firm entry is the question the aggregate data cannot answer, and the natural next step for administrative worker-flow data.

References

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

Adão, R., M. Kolesár, and E. Morales (2019). Shift-share designs: Theory and inference. *Quarterly Journal of Economics 134*(4), 1949–2010.

Autor, D. (2024). Applying AI to rebuild middle class jobs. Working Paper 32140, National Bureau of Economic Research.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review 103*(6), 2121–2168.

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

Babina, T., A. Fedyk, A. He, and J. Hodson (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics 151*, 103745.

Balland, P.-A., R. Boschma, J. Crespo, and D. L. Rigby (2019). Smart specialization policy in the European Union: Relatedness, knowledge complexity and regional diversification. *Regional Studies 53*(9), 1252–1268.

Barro, R. J. and X. Sala-i Martin (1992). Convergence. *Journal of Political Economy 100*(2), 223–251.

Bick, A., A. Blandin, and D. J. Deming (2024). The rapid adoption of generative AI. Working Paper 32966, National Bureau of Economic Research.

Bonney, K., C. Breaux, C. Buffington, E. Dinlersoz, L. Foster, N. Goldschlag, J. Haltiwanger, Z. Kroff, and K. Savage (2024). Tracking firm use of AI in real time: A snapshot from the business trends and outlook survey. Working Paper CES-WP-24-16, U.S. Census Bureau, Center for Economic Studies.

Borusyak, K., P. Hull, and X. Jaravel (2022). Quasi-experimental shift-share research designs. *Review of Economic Studies 89*(1), 181–213.

Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda. *Regional Studies 51*(3), 351–364.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *Quarterly Journal of Economics 134*(3), 1405–1454.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal 21*(1), C1–C68.

Cohen, W. M. and D. A. Levinthal (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly 35*(1), 128–152.

Comin, D. and B. Hobijn (2010). An exploration of technology diffusion. *American Economic Review 100*(5), 2031–2059.

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models. arXiv:2303.10130.

Felten, E., M. Raj, and R. Seamans (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal 42*(12), 2195–2217.

Felten, E. W., M. Raj, and R. Seamans (2023). Occupational heterogeneity in exposure to generative AI. SSRN Working Paper 4414065.

Gathmann, C. and U. Schönberg (2010). How general is human capital? A task-based approach. *Journal of Labor Economics 28*(1), 1–49.

Giannone, E. (2022). Skill-biased technical change and regional convergence. Working paper, CREI.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica 25*(4), 501–522.

Handa, K., A. Tamkin, M. McCain, S. Huang, E. Durmus, et al. (2025). Which economic tasks are performed with AI? evidence from millions of Claude conversations. arXiv:2503.04761.

Hidalgo, C. A. and R. Hausmann (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences 106*(26), 10570–10575.

Hidalgo, C. A., B. Klinger, A.-L. Barabási, and R. Hausmann (2007). The product space conditions the development of nations. *Science 317*(5837), 482–487.

Howell, A., M. Feldman, L. Lanahan, N. Kalathil, and E. Johnson (2026). An economic geography dataset of U.S. skill specialization, relatedness, and complexity. arXiv:2606.09918; data: Zenodo record 20330086.

Humlum, A. and E. Vestergaard (2025). The unequal adoption of ChatGPT exacerbates existing inequalities among workers. *Proceedings of the National Academy of Sciences 122*(1), e2414972121.

McElheran, K., J. F. Li, E. Brynjolfsson, Z. Kroff, E. Dinlersoz, L. Foster, and N. Zolas (2024). AI adoption in america: Who, what, and where. *Journal of Economics & Management Strategy 33*(2), 375–415.

Microsoft Research (2019). EconML: A python package for ML-based heterogeneous treatment effects estimation. Version 0.16; https://github.com/py-why/EconML.

Moretti, E. (2012). *The New Geography of Jobs*. Boston: Houghton Mifflin Harcourt.

Neffke, F., M. Henning, and R. Boschma (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography 87*(3), 237–265.

U.S. Census Bureau (2026). The microstructure of AI diffusion: Evidence from firms, business functions, and worker tasks. Working Paper CES-WP-26-25, U.S. Census Bureau, Center for Economic Studies.

Webb, M. (2020). The impact of artificial intelligence on the labor market. SSRN Working Paper 3482150.

A Data Appendix

BTOS sources. Current cumulative releases (`census.gov/hfp/btos`): *State.xlsx*, *Top 25 MSA.xlsx* (response estimates by question×answer×period), *URR.xlsx* (unit response rates), AI supplements. The regime-1 AI series (original wording, periods 202319–202520) is published separately in the archive file *AI Core Questions.xlsx* (state and Top-25-MSA sheets); the current cumulative files carry only the regime-2 series. Values are parsed from published percentages; `S` cells (disclosure suppression) become missing. Period dates and the regime flag come from the published collection/reference-date table; regime 2 begins with the collection period starting 2025-11-17. Coverage detail behind Section 2.1: after suppression the regime-1 AI cell is published for 86% of state×biweek cells and 84% of MSA cells (regime 2: 97%, 99%); Alaska, North Dakota, and San Antonio have no published regime-1 cells; St. Louis (18%), Wyoming (25%), Mississippi (27%), Charlotte (32%), and Vermont (33%) are thin. The aggregate series is smooth (first-difference s.d. 0.40pp on a series rising 4.4%→10.6%; lag-1 autocorrelation 0.98) but noisy within geography (median within-geography first-difference s.d. 1.85pp, p90 2.64pp).

Geography. BTOS MSA display names are matched to 2023 CBSA delineations on (first principal city, first state); counties aggregate to CBSAs via the OMB delineation file. One override: OEWS retains the legacy NECTA code 71650 for Boston-Cambridge (CBSA 14460). Howell counties aggregate to states by FIPS prefix. Posting-weighted means aggregate intensive skill measures; counts are summed.

Exposure inputs. OEWS May-2023 state and MSA all-data files (detailed SOC × area employment); AIOE scores from the published data appendix of Felten et al. (2021); LLM exposure from the Eloundou et al. (2023) replication files (O*NET-SOC truncated to 6-digit SOC, averaged); shares renormalized over occupations with non-missing intensity. The full download manifest (URLs, dates, SHA-256) ships with the replication package.

B Pre-registration Memo (exp000)

The memo, reproduced in the replication package and version-controlled before any outcome regression was estimated, froze: (i) the nine-event calendar and intensity vectors of Table 1, with the Nov-2025 cluster excluded because it coincides with the BTOS wording break; (ii) the capability index $K_d(b) = \sum_{b' \le b} \Delta K_d(b')$ and its event-to-biweek mapping; (iii) the three exposure measures and OEWS vintages; (iv) the skill moderators (lagged complexity and relatedness density) and the 2024/2022/2019 vintage horse race; (v) the expected 1–6 month lag range with its two falsifiability commitments; (vi) the five placebo dates (2024-03-11, 2024-08-05, 2024-12-09, 2025-04-14, 2025-07-28), each $\geq$ 6 weeks from any event; (vii) the breakout definition (10%, 3 consecutive biweeks) and robustness grid; (viii) inference (geography clustering; AKM/BHJ in queue) and the sample rule (drop fully suppressed geographies). *Caveat, stated openly:* the memo is internally

version-controlled but was not externally time-stamped before estimation; the memo is unchanged since freezing. Because the memo cannot be externally verified, the paper's language reserves "pre-registered" for design elements whose ex-ante status is corroborated by the rule-based construction itself, and the identification never rests on the timestamp.

C Estimation and Machine-Learning Details

Doubly-robust heterogeneity (Section 4.2). The cross-sectional check uses the linear double machine-learning estimator of Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey, and Robins (2018) as implemented in Microsoft Research (2019): outcome Y_g = 2023–25 adoption growth (mean of the last six regime-1 biweeks minus the first six), treatment T_g = pre-period AIOE exposure, effect modifier X_g = lagged relatedness density (z), controls W_g = baseline adoption, an MSA indicator, and the other skill moderator; nuisance functions are random forests (200 trees, minimum leaf size 5), cross-fitted; $n = 62$ geographies with sufficient published biweeks. The reported statistic is the slope of the estimated conditional treatment effect on the modifier, the doubly-robust analogue of the moderation θ. Off-the-shelf DML is *invalid* for the panel event-study core, where the treatment is a constructed shift-share interacted with two-way fixed effects; the long difference removes that structure, which is what licenses the estimator here. Single-modifier specifications cannot discriminate among correlated moderators (each rides the common signal), so the discriminating test remains the fixed-effects horse race with density co-included.

Prediction protocol (Section 4.1). Gradient boosting with hyperparameters pinned by cross-validation inside the training folds only (elastic net as the linear benchmark); features are the 2024-vintage skill measures, 2019→2024 skill growth, the three exposure measures, and time. Temporal holdout: train on pre-2025 geography-biweeks, test on 2025 regime-1 cells ($n = 1{,}158$). Leave-one-geography-out: delete each geography entirely, predict its June–November 2025 mean. The LOGO correlation is insensitive to re-tuning (0.63 baseline, 0.65 CV-pinned).

Distributed-lag windowing (Table 6). The windowed rows group the lag terms of equation (6) into bins (0–2, 3–4, 5–6, 7–9, 10–12, and additionally 13–18, 19–26 in the long variant) and re-estimate, summing the window coefficients; grouping removes the offsetting oscillation of collinear adjacent lags that inflates the lag-by-lag sum.

Placebo construction (Table 8). Each placebo assigns a fictitious unit capability jump to one of the five pre-registered non-release biweeks and re-estimates the distributed-lag specification with lags 0–6; the conditional variant adds the full true-event lag structure as controls, so the placebo is identified only from variation beyond genuine lagged responses.

AKM and BHJ inference (Section 4.6). Exposure is a shift-share in 672 detailed occupations, so I compute Adão et al. (2019) (AKM) occupation-level standard errors from occupation score



Figure 8: Breakout timing (biweek index of first sustained 10% adoption) against lagged relatedness density, geographies that broke out by end of regime 1 (Section 4.2).

sums $\psi_o = \sum_g s_{go}\,\tilde{x}_g\hat{\varepsilon}_g$, using the exact renormalized OEWS×AIOE shares that enter Exposure. For the cumulative-stock specification the AKM standard error is 0.03 against 0.49 clustered by geography (cross-section: 0.07 versus 1.65). The correction is an order of magnitude smaller than geography clustering because shares are diffuse (share Herfindahl 0.011, ≈95 effective occupations) and nearly identical across regions (mean pairwise cosine similarity 0.91), so the share-induced cross-region error correlation the correction targets is negligible; I verified the implementation by replacing the shares matrix with geography indicators, which reproduces the clustered standard error. A Borusyak et al. (2022) shock-level check agrees: the region-level long-difference gradient equals the exposure-weighted occupation-level computation exactly (8.50 either way), with shock-level standard error 0.99 against the conventional 1.65. The binding, conservative inference is therefore the geography-clustered (and two-way clustered) standard errors reported throughout.

Shock-intensity schemes (Section 4.6). A bare binary release indicator gives the stock gradient with $t = 3.19$; a benchmark-anchored intensity vector (approximate SWE-bench, MMLU, and price-per-token deltas) gives $t = 3.26$; and 500 random perturbations of the three domain scores (each ± 0.25, clipped to $[0, 1]$) leave the gradient positive in every draw (s.d. 0.08, 5th–95th percentile $[1.47, 1.73]$). Raw coefficients rescale across schemes only because each changes the units of total K; the inference and the cumulated magnitude do not move.

D Appendix Tables and Figures

Table 9: Distributed-lag profile β_ℓ: capability jumps × regional exposure

Lag ℓ (biweeks)	AIOE exposure β_ℓ	(s.e.)	LLM-β exposure β_ℓ	(s.e.)	Coding exposure β_ℓ	(s.e.)
0	3.24	(2.32)	8.95	(14.79)	14.93	(18.84)
1	-1.57	(1.37)	-7.54	(7.85)	-17.07	(11.88)
2	1.85	(2.18)	18.73	(13.15)	19.53	(18.10)
3	0.79	(1.98)	-14.15	(16.85)	-31.88	(17.29)
4	5.83	(1.72)	18.84	(10.80)	41.00	(12.05)
5	-1.60	(1.68)	-2.15	(11.60)	-4.86	(14.75)
6	0.60	(1.72)	-1.69	(9.20)	1.40	(14.38)
7	-0.41	(2.22)	-20.06	(14.46)	-23.98	(17.93)
8	6.64	(1.93)	55.12	(14.57)	55.27	(21.08)
9	2.98	(2.86)	11.75	(17.31)	-1.69	(22.69)
10	2.71	(2.05)	8.65	(13.19)	-11.53	(14.74)
11	1.16	(2.14)	-3.48	(11.90)	-10.95	(18.03)
12	2.07	(2.32)	14.18	(14.69)	63.83	(17.26)

Notes: Coefficients on $\Delta K(b-\ell) \times \text{Exposure}(g)$ from equation (6), estimated on the regime-1 panel (N=2,877 geography-biweeks, 74 geographies) with geography and geography-type×biweek fixed effects; standard errors clustered by geography. Outcome: BTOS share of firms using AI in producing goods or services (percentage points), regime-1 panel, Sep 2023–Nov 2025. Lags 5–6 of the AIOE specification are dropped by collinearity with the event spacing. Triple-interaction terms with lagged complexity and relatedness density are included and reported in Table 3.

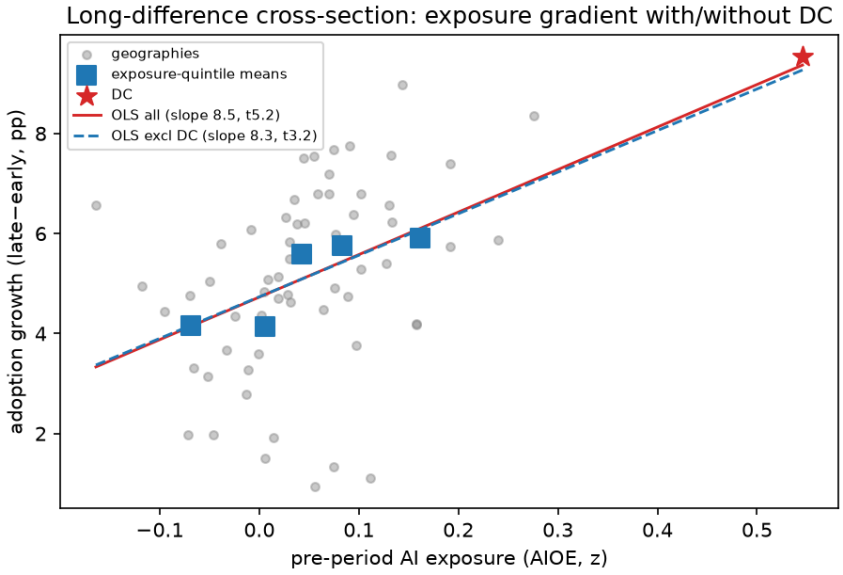


Figure 9: Long-difference cross-section: adoption growth (late minus early regime-1 mean, pp) against pre-period AIOE exposure. Grey points are geographies; blue squares are exposure-quintile means (DC excluded from the binning); the red star is DC. The OLS fit with DC (solid) and without DC (dashed) are nearly identical (slopes 8.5 and 8.3), and the quintile means are linear in exposure.

Table 10: Stacked event study: $\delta_\tau \times$ AIOE exposure around isolated releases

τ (biweeks since event)	δ_τ	(s.e.)
-8	-0.75	(1.20)
-7	0.05	(1.49)
-6	-1.25	(1.76)
-5	-2.19	(2.18)
-4	-0.24	(1.67)
-3	-1.00	(1.38)
-2	1.29	(1.07)
-1	0.00	(0.00)
0	-0.29	(1.39)
1	0.49	(1.36)
2	2.11	(1.59)
3	1.53	(1.58)
4	0.97	(1.31)
5	0.30	(1.46)
6	0.05	(1.85)
7	1.76	(1.70)
8	2.68	(1.20)
9	1.87	(1.38)
10	3.25	(2.00)
11	6.55	(2.10)
12	2.32	(2.32)

Notes: Cengiz et al. (2019)-style stacked design over the three pre-registered temporally isolated events (GPT-4o, o1, DeepSeek-R1); event windows $\tau \in [-8, +12]$ biweeks, $\tau = -1$ omitted; geography×event and biweek×event fixed effects; standard errors clustered by geography; N=3,343 stacked geography-biweeks. Outcome: BTOS share of firms using AI in producing goods or services (percentage points), regime-1 panel, Sep 2023–Nov 2025.

Table 11: Discrete-time (cloglog) hazard of adoption breakout

	Coef.	(s.e.)	z
Constant	-22.67	(2.63)	-8.61
Complexity (ECI, lagged, z)	-0.22	(0.53)	-0.41
Relatedness density (lagged, z)	0.64	(0.54)	1.20
AIOE exposure	3.20	(3.69)	0.87
log duration	4.93	(0.71)	6.93
MSA indicator	0.93	(0.58)	1.60

Notes: Breakout = first of three consecutive published biweeks with adoption $\geq$ 10% (pre-registered). Complementary-log-log discrete-time hazard on the regime-1 risk set (a geography exits at breakout): coefficients on the log-hazard scale; log duration and an MSA indicator included; skill moderators z-scored and lagged one calendar year. Risk set 2,542 geography-biweeks, 33 breakout events; standard errors clustered by geography. A Cox proportional-hazards variant gives a density hazard ratio of 2.40 ($z = 2.5$), threshold-invariant from 6% to 14% (exp014).

Table 12: Robustness of the treadmill and density-absorption statistics

	Treadmill $\sum_\ell \beta_\ell$			$\sum_\ell \theta_\ell$ (density)			
Specification	Coef.	(s.e.)	t	Coef.	(s.e.)	t	N
baseline (geography cluster)	24.29	(10.93)	2.22	34.57	(11.69)	2.96	2,877
two-way cluster (geography & biweek)	24.29	(9.29)	2.62	34.57	(13.11)	2.64	2,877
3-biweek smoothed adoption	23.40	(10.98)	2.13	27.94	(17.14)	1.63	2,882
exclude <50%-coverage geographies	30.05	(12.22)	2.46	30.00	(26.56)	1.13	2,698
May-2022 OEWS exposure weights	27.33	(10.93)	2.50	34.09	(11.46)	2.97	2,877

Permutation inference: 500 random reassignments of exposure across geographies;
no draw attains $|\sum \beta| \geq$ observed ($p < 0.002$; the no-moderator statistic is 4.7 permutation s.d. from zero).

Notes: each row re-estimates equation (6) (AIOE exposure) under the stated variant; geography and geography-type×biweek fixed effects unless noted; standard errors clustered by geography (row 2: two-way by geography and biweek). Outcome: BTOS share of firms using AI in producing goods or services (percentage points), regime-1 panel, Sep 2023–Nov 2025. Permutation inference uses the no-moderator specification and compares like with like.

Table 13: Placebo battery conditional on true-event lag terms

Placebo date	$\sum_{\ell=0}^{6} \beta_\ell^{placebo}$	(s.e.)	t
2024-03-11	-4.40	(7.37)	-0.60
2024-08-05	-25.21	(11.12)	-2.27
2024-12-09	-2.39	(6.88)	-0.35
2025-04-14	16.18	(10.53)	1.54
2025-07-28	14.07	(7.97)	1.77

Notes: as Table 8, adding the full true-event distributed-lag terms as controls, so each placebo is identified only from variation beyond true lagged responses; same specification, fixed effects, and clustering.

Table 14: Sample accounting: every estimation sample in the paper

Sample	N	Used in
BTOS geographies (50 states + DC + PR + 25 MSAs)	77	universe; 3 units fully suppressed
Geographies with regime-1 AI data	74	all panel analyses
Distributed-lag panel (geo-biweeks)	2,877	eq. (6); 12 lags consumed
Cumulative-stock panel (geo-biweeks)	2,906	eq. (4); no lags lost
State-only stock panel (geo-biweeks)	1,938	nested-geography check
Vintage regressions (geo-biweeks)	2,877	76 geographies with skill data
Long-difference cross-section	62	needs 12+ biweeks; DML, convergence
LOGO prediction	73	one prediction per geography
Breakout risk set (geo-biweeks)	2,542	33 breakout events
AEI-matched states	49	validation; augmentation margin

Notes: all samples derive from the regime-1 BTOS panel (Sep 2023–Nov 2025) after disclosure suppression. Differences across rows reflect lag consumption (distributed lag), minimum-published-biweeks requirements (growth cross-section), skill-panel coverage (vintage regressions), breakout exits (hazard risk set), and AEI state matching.

Table 15: Complexity-measure horse race: relatedness density versus alternative complexity operationalizations (distributed-lag moderation, AIOE exposure, $N = 2{,}877$)

Complexity measure C (density co-included)	$\sum_\ell \theta_{\text{density}}$ (t)	$\sum_\ell \theta_C$ (t)
ECI (incumbent)	34.6 (2.96)	−15.0 (−1.08)
Tacchella–Pietronero fitness	27.2 (3.46)	−10.8 (−1.47)
Skill entropy	45.1 (3.58)	−27.7 (−1.65)
Average ubiquity	71.1 (3.45)	+60.9 (+2.48)
Skill HHI	37.0 (4.34)	+16.7 (+1.68)

Notes: each row re-estimates the distributed-lag moderation (equation (6)) with relatedness density and one complexity measure C co-included as moderators; entries are the cumulative $\sum_\ell \theta$ and its t (clustered by geography); geography and geography-type×biweek fixed effects; outcome is BTOS adoption (pp), regime-1 panel. Density's moderation is positive and significant against every alternative complexity measure. The sophistication measures (ECI, fitness, entropy) are null or negative conditional on density. Average ubiquity is the lone positive, but it measures the *commonness* of a region's skills (inverse sophistication), so its positive sign reinforces "connectivity, not sophistication," and it reverses under the doubly-robust DML estimator (exp009). Growth specifications (level/log/logit) leave the ordering unchanged.

Table 16: Readiness typology: all geographies (LOGO predictions, late regime 1)

Geography	Actual	Predicted	Residual	Type	Weaker margin
Boston, MA MSA	11.3	14.4	-3.1	frontier	depth
Washington, DC MSA	13.1	15.7	-2.6	frontier	depth
ID	10.3	12.7	-2.4	frontier	composition
MD	10.7	12.3	-1.7	frontier	depth
CA	10.6	12.2	-1.6	frontier	composition
Seattle, WA MSA	13.0	14.5	-1.4	frontier	composition
MN	10.2	11.5	-1.3	frontier	depth
RI	11.1	12.2	-1.1	frontier	depth
Dallas, TX MSA	12.2	13.2	-1.0	frontier	composition
Baltimore, MD MSA	10.6	11.5	-0.9	frontier	depth
WA	11.7	11.8	-0.1	frontier	composition
Tampa, FL MSA	12.9	13.0	-0.1	frontier	depth
AZ	11.7	11.5	+0.2	frontier	composition
Atlanta, GA MSA	12.0	11.7	+0.3	frontier	composition
Riverside, CA MSA	10.5	10.2	+0.3	frontier	composition
Charlotte, NC MSA†	11.8	11.4	+0.5	frontier	composition
UT	11.9	11.3	+0.6	frontier	composition
Houston, TX MSA	10.4	9.7	+0.7	frontier	composition
Minneapolis, MN MSA	12.0	11.2	+0.9	frontier	depth
Portland, OR MSA	11.8	10.4	+1.4	frontier	depth
Miami, FL MSA	12.6	11.0	+1.5	frontier	depth
CO	13.3	11.4	+1.9	frontier	depth
San Francisco, CA MSA	14.5	12.5	+2.0	frontier	composition
Phoenix, AZ MSA	13.9	11.6	+2.3	frontier	composition
San Diego, CA MSA	14.4	11.6	+2.8	frontier	composition
VT†	13.1	10.1	+3.0	frontier	depth
WY†	15.1	11.9	+3.1	frontier	composition
Denver, CO MSA	15.5	11.5	+4.0	frontier	depth
DC	23.1	14.5	+8.6	frontier	depth
VA	10.2	9.3	+0.9	overperformer	depth
OR	10.2	8.5	+1.6	overperformer	depth
St. Louis, MO MSA†	10.6	8.7	+1.8	overperformer	depth
FL	11.6	9.2	+2.4	overperformer	depth
HI	10.8	8.3	+2.5	overperformer	composition
NV	11.5	9.0	+2.5	overperformer	composition
Orlando, FL MSA	12.5	9.6	+2.9	overperformer	composition
DE	14.1	8.5	+5.6	overperformer	depth
New York, NY MSA	7.4	9.6	-2.2	structurally lagging	depth
WI	6.2	7.9	-1.8	structurally lagging	composition
OK	7.1	8.8	-1.7	structurally lagging	composition
PA	8.1	9.7	-1.5	structurally lagging	depth
KS	7.3	8.7	-1.4	structurally lagging	composition
CT	8.4	9.6	-1.2	structurally lagging	depth
AR	7.6	8.7	-1.2	structurally lagging	composition

Geography	Actual	Predicted	Residual	Type	Weaker margin
IA	6.8	7.9	-1.1	structurally lagging	depth
NY	7.6	8.7	-1.1	structurally lagging	depth
MO	7.9	8.9	-1.0	structurally lagging	depth
MI	7.9	8.8	-0.8	structurally lagging	depth
LA	7.3	8.0	-0.7	structurally lagging	composition
IL	8.5	9.2	-0.7	structurally lagging	composition
IN	7.0	7.7	-0.7	structurally lagging	composition
ME	7.6	8.3	-0.6	structurally lagging	depth
WV	7.7	8.3	-0.6	structurally lagging	depth
Chicago, IL MSA	8.9	9.2	-0.3	structurally lagging	composition
OH	8.6	8.7	-0.1	structurally lagging	composition
AL	6.9	6.8	+0.1	structurally lagging	composition
NM	8.0	7.9	+0.1	structurally lagging	depth
NH	8.7	8.4	+0.3	structurally lagging	depth
MT	8.0	7.5	+0.5	structurally lagging	depth
NC	10.0	9.4	+0.6	structurally lagging	composition
SC	8.6	7.7	+1.0	structurally lagging	depth
TN	8.7	7.7	+1.0	structurally lagging	composition
KY	8.2	7.2	+1.0	structurally lagging	depth
NE	9.5	7.6	+1.9	structurally lagging	composition
SD	9.3	7.3	+2.0	structurally lagging	depth
MS†	7.0	11.3	-4.3	untapped readiness	depth
NJ	7.3	11.3	-4.0	untapped readiness	depth
Los Angeles, CA MSA	9.5	13.5	-4.0	untapped readiness	composition
Detroit, MI MSA	7.7	11.4	-3.7	untapped readiness	depth
GA	9.9	11.0	-1.1	untapped readiness	depth
TX	10.0	11.1	-1.0	untapped readiness	composition
MA	9.9	10.7	-0.8	untapped readiness	depth
Philadelphia, PA MSA	9.8	9.9	-0.1	untapped readiness	depth

Notes: Actual = mean published adoption, June–November 2025 (share of firms using AI, percent). Predicted = leave-one-geography-out gradient-boosted prediction from pre-period skill structure and exposure alone (Section 4.1); Residual = actual − predicted (percentage points). Type: median splits of actual and predicted. Weaker margin = the region's *relatively* weaker fundamental (occupational composition vs. skill depth), a within-region ranking rather than an absolute deficit. †Regime-1 publication coverage below 35% of biweeks (disclosure suppression: MS 27%, WY 25%, VT 33%, Charlotte 32%, St. Louis 18%); residuals for these units carry extra sampling noise.

E Persistence under the post-November-2025 instrument

The regime-2 series (consistent "any business function" wording) accumulates biweekly; by submission a year-plus of post-break data will allow a full persistence re-estimate of the determinants and typology under the new instrument. The first 18 regime-2 biweeks (collection periods 202524–202615, 76 geographies) already let me check the central skill relationship under the new wording, *without* splicing the two regimes: regressing each geography's mean regime-2 adoption on its

2024 skill structure, relatedness density predicts adoption with a standardized coefficient of 2.32 ($t = 3.38$; raw density–adoption correlation 0.74), while complexity (ECI, $t = -0.4$) and exposure ($t = 0.6$) are null, the same density-over-complexity ordering as in regime 1, if anything sharper. The skill geography of adoption is therefore not an artifact of the regime-1 question wording. These figures are regenerated from the pipeline as further regime-2 biweeks accrue.