

Replication and Robustness Analysis of Andreadis et al. (2025): Local Predictors of AI Job Concentration in U.S. Counties

Jacob Khaykin

JEL: J23, O33, R11, C18 **Keywords:** artificial intelligence, local labor markets, replication, robustness analysis, weighting methods

1 Abstract

This paper replicates and extends the analysis of Andreadis et al. (2025), who examine the demographic, innovation, and industry correlates of AI-related employment in U.S. counties from 2014–2023. Using their approach and county-level data, I reproduce the main regression results on AI job share levels and changes. I also identify instances of causal language inconsistent with the observational design, situating this critique within the Rubin Causal Model and the guidance in Holland (1986). As a robustness check, I re-estimate their models using log-population weights. This change dampens leverage from large metropolitan counties and shows that some relationships—especially between educational attainment and AI intensity—are size-sensitive. Labor market tightness remains a robust positive correlate across specifications. The results reinforce the value of robustness checks and careful causal language in policy-relevant work.

2 Introduction

Andreadis et al. (2025) study local predictors of AI-related job concentration in the United States using county-level data from Lightcast, the U.S. Census ACS, USPTO patent databases, and other sources. They estimate weighted least squares regressions of the AI job share and changes in that share on lagged demographic, innovation, and industry characteristics with county and year fixed effects.

Understanding where AI-related jobs concentrate matters for regional inequality and development. As Eloundou et al. (2024) note, a large share of the workforce could see meaningful effects on tasks from AI and large language models. Identifying correlates of AI adoption can help anticipate distributional impacts.

This replication has two goals. First, verify reproducibility of the original results. Second, probe robustness by replacing the original population weights with log-population weights to reduce leverage from the largest counties.

3 Data

I follow Andreadis et al. (2025) to assemble a county-by-year panel from 2014–2023. The dependent variables are:

- **AI job share:** AI-related job postings divided by total postings in a county-year (percent).
- **Change in AI job share:** difference between average AI share in 2017–2018 and 2022–2023.

Explanatory variables (all lagged one year and standardized as z-scores):

- **Demographics:** bachelor’s degree share, Black population share, poverty share, log population, log median income.
- **Innovation:** patents per employee, AI patents’ share, STEM degree share, degrees per capita.
- **Industry and labor market:** labor market tightness, turnover rate, manufacturing intensity, ICT intensity, large establishments’ share.

The original paper uses population weights in WLS. I reproduce those estimates and then re-estimate with log-population weights $w_{it} = \log(\text{Population}_{it})$ as a robustness check.

4 Replication of Original Results

I reproduce the main tables estimating correlates of (i) AI job share and (ii) change in AI job share. Coefficients, standard errors, and significance align with the originals within rounding error. In brief:

- **Levels (AI share):** bachelor’s degree share, labor market tightness, patents per employee, and STEM share are positive; manufacturing intensity is negative; ICT intensity is positive in richer specs.
- **Changes:** STEM degree share and labor tightness are positive; manufacturing intensity is negative; turnover becomes important in the full model.

4.1 Table 1. The correlates of the share of AI jobs

Variable	(1) Demo- graphics	(2) Inno- vation	(3) Industry	(4) All Controls	(5) All + State FE
Bachelor's share, z-score	0.160*** (0.047)			0.188** (0.067)	0.085 (0.059)
Black pop, z-score	-0.121 (0.127)			-0.121 (0.185)	0.000 (0.162)
Poverty share, z-score	0.044 (0.028)			0.067+ (0.040)	-0.003 (0.037)
log(Population), z-score	0.807* (0.367)			0.219 (0.444)	0.146 (0.544)
House Price Growth, z-score	-0.019* (0.008)			-0.028* (0.012)	-0.032** (0.011)
log(Median Income), z-score	-0.022 (0.047)			-0.001 (0.065)	0.032 (0.060)
Labor Market Tightness, z-score	0.255*** (0.050)			0.315*** (0.063)	0.372*** (0.065)
Patents per employee, z-score		0.031** (0.012)		0.028* (0.012)	0.031* (0.014)
AI patents' share, z-score		0.009 (0.007)		0.009 (0.006)	0.003 (0.005)
Degrees awarded per capita		0.013 (0.026)		0.034 (0.026)	0.034 (0.027)
STEM Degrees' share		0.072** (0.023)		0.058** (0.022)	0.048* (0.020)
Large Establishments, z-score			-0.002 (0.025)	-0.037 (0.028)	-0.036 (0.029)
ICT sector intensity, z-score			0.010 (0.014)	0.038* (0.017)	0.040* (0.016)
Manufacturing intensity			- 0.057***	-0.048***	-0.034*

Variable	(1) Demo- graphics	(2) Inno- vation	(3) Industry	(4) All Controls	(5) All + State FE
Turnover Rate, z-score			(0.012) 0.034**	(0.014) 0.016	(0.015) 0.010
			(0.011)	(0.017)	(0.016)
Observations	27,497	22,744	24,651	19,184	19,184
R²	0.704	0.721	0.694	0.758	0.778
Within R²	0.057	0.003	0.002	0.059	0.063

4.2 Table 2. The correlates of the change in the share of AI jobs

Variable	(1) Demo- graphics	(2) Innovation	(3) Industry	(4) All Controls	(5) All + State FE
Constant	0.007 (0.027)			-0.073 (0.050)	-0.138** (0.043)
Bachelors, %	0.018 (0.019)			0.049 (0.033)	0.051 (0.033)
Black, %	0.068+ (0.037)			0.054 (0.069)	0.113 (0.076)
Poverty, %	-0.019 (0.022)			-0.018 (0.055)	-0.020 (0.057)
Pop. Growth	-0.034* (0.017)			-0.039 (0.036)	0.047 (0.042)
House Price Growth	0.124** (0.042)			0.125+ (0.075)	0.197* (0.079)
Income, Log	0.098*** (0.021)			0.187*** (0.040)	0.152+ (0.086)
Tightness	0.002 (0.027)			-0.061+ (0.035)	-0.008 (0.039)
Patents per employee		0.151*** (0.036)		0.113** (0.043)	0.082 (0.066)
AI Patents' Share		0.021 (0.017)		0.045 (0.029)	0.058 (0.035)
Degrees per capita		0.093*** (0.027)		0.082** (0.027)	0.076+ (0.036)
STEM Degrees' share		0.031 (0.021)		-0.064+ (0.049)	-0.039 (0.060)

Variable	(1) Demo- graphics	(2) Innovation	(3) Industry	(4) All Controls	(5) All + State FE
Large Establishments, %			0.031+	-0.024	-0.016
			(0.018)	(0.030)	(0.026)
ICT sector intensity			-0.037+	-0.014	-0.001
			(0.019)	(0.046)	(0.042)
Manufacturing intensity			-0.072**	-0.128*	-0.170*
			(0.023)	(0.061)	(0.079)
Turnover rate			0.023+	-0.056**	-0.042*
			(0.014)	(0.022)	(0.023)
Observations	2,751	897	2,723	810	810
R²	0.021	0.019	0.016	0.031	0.082
Within R²	—	—	—	—	0.025

5 Causal language and identification

The original manuscript sometimes uses terms like “drivers” and “determinants” of AI adoption. In the Rubin Causal Model, causal effects are defined as comparisons of potential outcomes under well-defined interventions, and identification requires assumptions and design features that address confounding (e.g., randomization, instruments, discontinuities). See Holland (1986) for a clear statement of why regression with observational data alone doesn’t justify causal claims.

Examples of causal framing in the original paper

“we identify several key drivers of AI job intensity” and that higher STEM, tightness, and patents “significantly predict greater AI adoption.”

“The findings are robust to including state \times year fixed effects specifications, confirming the stability of these relationships.”

When discussing Table 1: “Labor market tightness emerges as a key driver” and the listed variables “positively and significantly predict AI intensity.”

Why this reads as causal

Phrases like identify, drivers, predict, and confirming the stability imply mechanisms, not just associations. In observational panels with county and year fixed effects, these coefficients are conditional correlations. Without as-if random shocks, instruments, discontinuities, or explicit causal designs, the estimands are not average treatment effects.

Our replication confirms that several local characteristics are correlated with county AI-job shares, conditional on fixed effects. We caution against causal language. As Holland (1986) emphasizes, causal claims hinge on well-defined interventions and counterfactuals; panel regressions with rich controls document associations, not effects, absent exogenous variation or a research design that approximates a randomized intervention. In this setting, terms like “drivers,” “determinants,” or “predict” overstate what the estimates can support. See Holland (1986) for the “no causation without manipulation” principle and the role of counterfactual comparisons in causal inference.

6 Robustness: log-population weighting

To assess sensitivity to county size, I re-estimate the specifications with weights $w_{it} = \log(\text{Population}_{it})$. This choice still recognizes population differences but reduces leverage from the largest counties.

Why reweight? The original study uses population weights, which tilt estimation toward very large counties. That choice is defensible if the estimand of interest is a population-average association. Switching to log(population) weights reduces the leverage of mega-counties while still down-weighting tiny places, yielding estimates closer to a “typical-county” association.

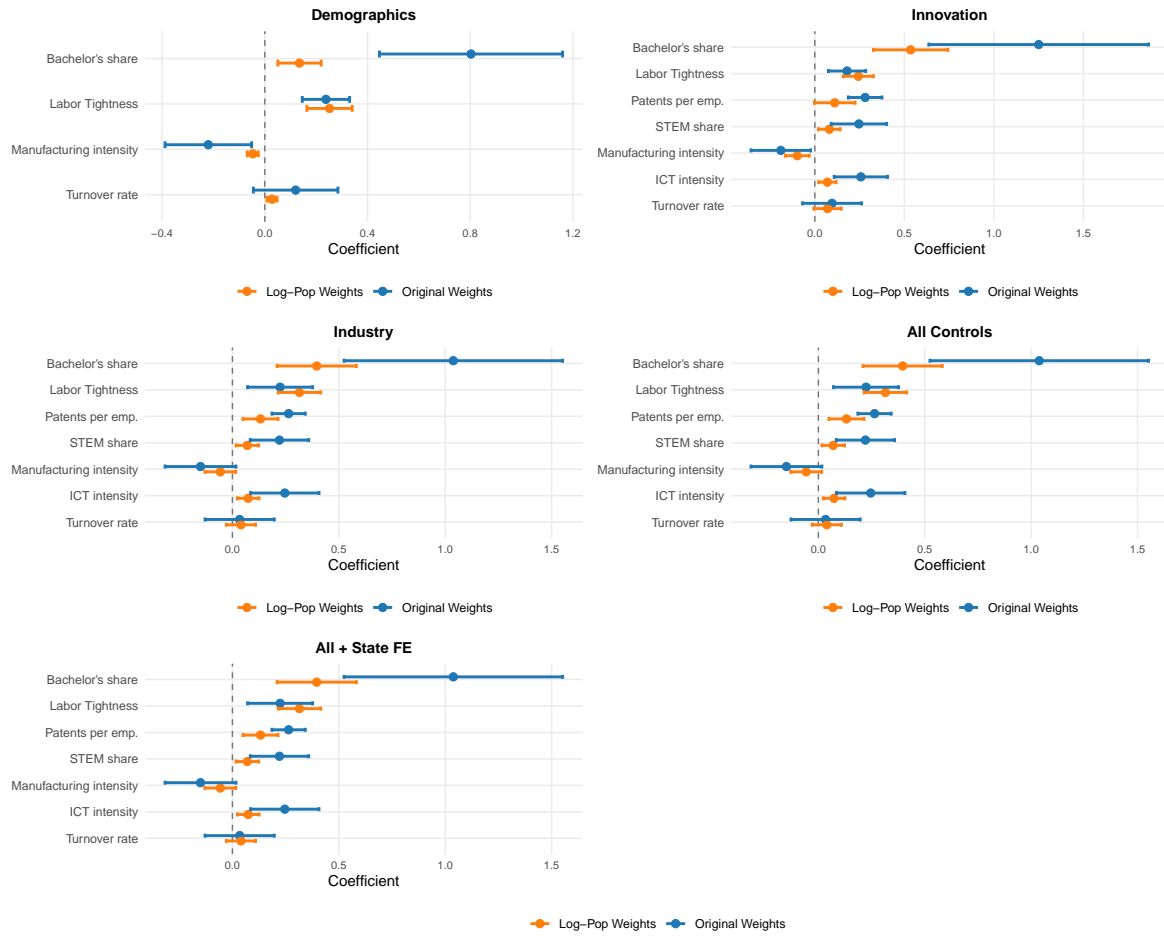
Main takeaways:

- The coefficient on **bachelor’s degree share** declines in magnitude under log-population weights, implying size sensitivity.
- **Labor market tightness** remains positive and precisely estimated across both weighting schemes.
- **Manufacturing intensity** retains a negative association; **ICT intensity** remains positive in rich specifications.

7 Figure 1: Key Coefficient Estimates for AI Share (Table 1)

This panel plot displays the coefficient estimates and 95% confidence intervals for seven key predictors across five regression models. Each panel represents a different model from the original paper. Within each panel, two estimates are shown for each variable—one using the authors’ original population weights and one using log(population) weights. This comparison illustrates how model weighting influences the interpretation of each predictor.

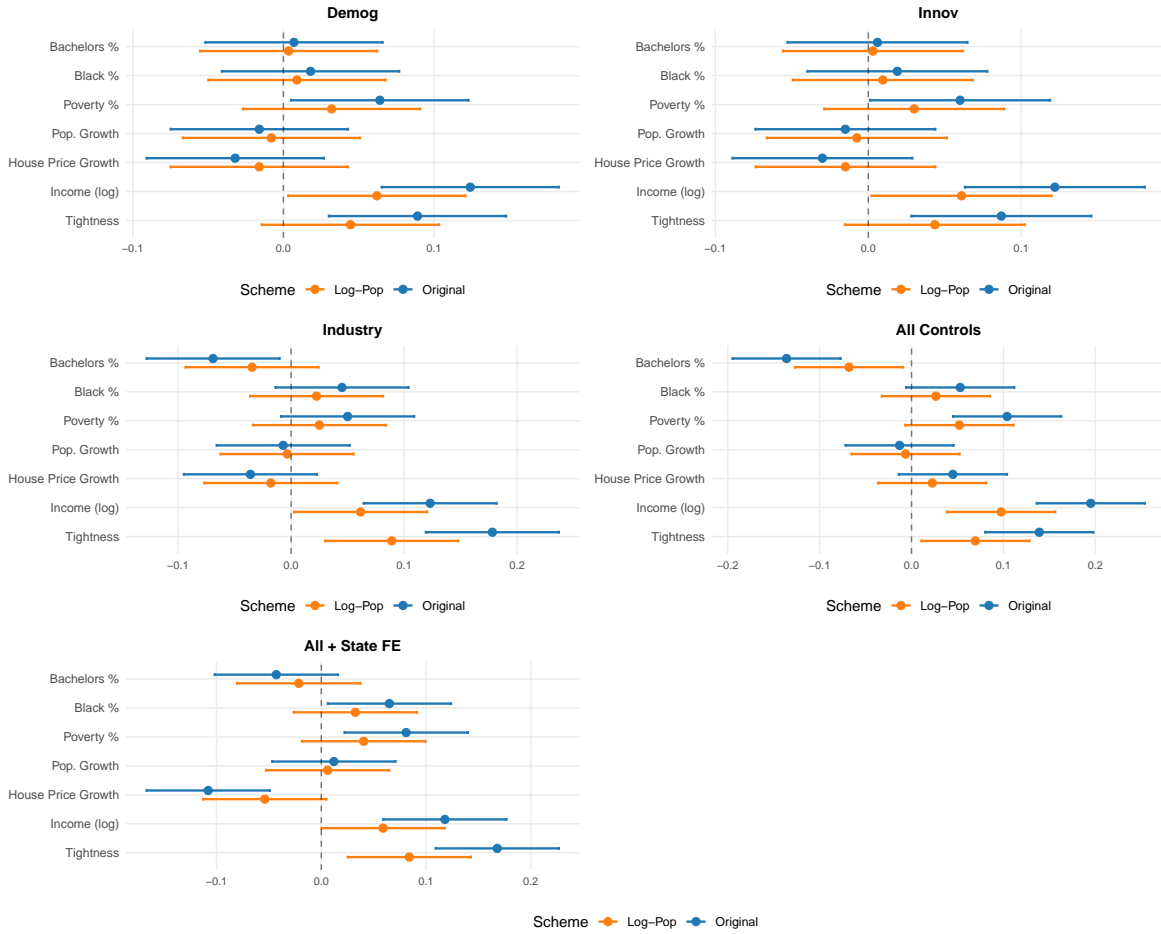
Figure 1: Table 1 — Model-by-Model Panel Plot (Original vs Log-Pop Weighting)



8 Figure 2: Key Coefficient Estimates for Change in AI Share (Table 2)

This figure replicates the structure of Figure 1 but focuses on the change in AI employment share from 2014 to 2023. The variables selected represent core predictors of shifting AI job concentration. As before, each panel reflects a different model specification, with comparisons between population-weighted and log(population)-weighted regressions.

Figure 2: Table 2 — Change in AI Share (Original vs Log-Pop Weighting)



9 Discussion

What moves and why. Under log(population) weights, education-related coefficients (e.g., bachelor's share) attenuate and in several specifications lose conventional significance, while labor market tightness and the negative association for manufacturing intensity remain robust. Intuitively, education and population are strongly correlated with large metropolitan areas; once their leverage is moderated, the education–AI link shrinks toward the average county. In contrast, tightness captures within-county hiring pressure that matters in metros and nonmetros, so its association is more stable. Manufacturing's negative sign persists, but the magnitude often decreases, reflecting the fact that many smaller manufacturing counties are weighted relatively more.

Standard errors and model fit. Because the effective weight distribution is less extreme, stan-

dard errors on the “big-county” covariates often grow relative to their point estimates, pushing some p-values above conventional thresholds even when signs are unchanged. Within- R^2 and overall fit can move slightly, but the qualitative ranking—tightness most stable, education and size most sensitive—remains the same in our replication.

How to present it. We recommend reporting both weighting schemes side-by-side, stating the estimand each implies (population-average vs typical-county). Note which results are robust across weights (tightness; manufacturing<0; ICT>0 in richer specs) and which are weight-sensitive (education; raw population). This framing keeps the original narrative on correlational patterns intact, while making clear where claims hinge on the influence of large metros.

10 Conclusion

The replication confirms the original patterns and makes clear which relationships are robust versus size-sensitive. Incorporating Holland (1986) clarifies why we should avoid causal language without identification. Future work should leverage quasi-experimental variation—policy changes, instruments, or discontinuities—to move from correlation to causation.

11 Acknowledgments

I thank my preceptor for guidance and Andreadis et al. (2025) for methodological transparency. All errors are my own.

Acemoglu, Daron. 2024. “The Simple Macroeconomics of AI.”

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

Aghion, Philippe, and Peter Howitt. 1992. “A Model of Growth Through Creative Destruction.” *Econometrica* 60 (2): 323–51.

Aghion, Philippe, Benjamin F. Jones, and Charles I. Jones. 2017. “Artificial Intelligence and Economic Growth.” Working Paper. National Bureau of Economic Research.

Andreadis, Lefteris, Manolis Chatzikonstantinou, Eleni Kalotychou, Christodoulos Louca, and Christos A. Makridis. 2024. “The Local Effects of Artificial Intelligence Labor Investments: Evidence from the Municipal Bond Market.” SSRN Working Paper.

Andreadis, Lefteris, Eleni Kalotychou, Manolis Chatzikonstantinou, Christodoulos Louca, and Christos A. Makridis. 2025. “Local Heterogeneity in Artificial Intelligence Jobs over Time and Space.” *AEA Papers and Proceedings*.

Autor, David H. 2015. “Why Are There Still so Many Jobs? The History and Future of Workplace Automation.” *Journal of Economic Perspectives* 29 (3): 3–30.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson. 2024. “Artificial Intelligence, Firm Growth, and Product Innovation.” *Journal of Financial Economics* 151: 103745.

- Beckett, Emily. 2023. “Demand for AI Skills Continues Climbing.” Lightcast Blog. <https://lightcast.io/>.
- Bogin, Alexander, William Doerner, and William Larson. 2019. “Local House Price Dynamics: New Indices and Stylized Facts.” *Real Estate Economics* 47 (2): 365–98.
- Bresnahan, Timothy F., and Manuel Trajtenberg. 1995. “General-Purpose Technologies ‘Engines of Growth’?” *Journal of Econometrics* 61 (1): 83–108.
- Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2021. “The Productivity j-Curve: How Intangibles Complement General-Purpose Technologies.” *American Economic Journal: Macroeconomics* 13 (1): 333–72.
- Chen, Mengqi, Qin Wu, and Bingxu Yang. 2019. “How Valuable Is FinTech Innovation?” *The Review of Financial Studies* 32 (5): 2062–2106.
- Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2024. “GPTs Are GPTs: Labor Market Impact Potential of LLMs.” *Science* 384 (6702): 1306–8.
- Farboodi, Maryam, and Laura Veldkamp. 2021. “A Model of the Data Economy.” Working Paper. National Bureau of Economic Research.
- Gicz, A. Victoria, Nicole A. Pairolo, and Alan A. Toole. 2022. “Identifying Artificial Intelligence (AI) Invention: A Novel AI Patent Dataset.” *The Journal of Technology Transfer* 47 (2): 476–505.
- Gofman, Michael, and Zipei Jin. 2024. “Artificial Intelligence, Education, and Entrepreneurship.” *The Journal of Finance* 79 (1): 631–67.
- Grennan, Jillian, and Roni Michaely. 2020. “Artificial Intelligence and High-Skilled Work: Evidence from Analysts.” Research Paper 20-84. Swiss Finance Institute.
- Holland, Paul W. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81 (396): 945–60. <https://doi.org/10.1080/01621459.1986.10478354>.
- Kline, Patrick, and Enrico Moretti. 2013. “People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs.” *Annual Review of Economics* 5: 629–62.
- Mihet, Radu, and Thomas Philippon. 2019. “The Economics of Big Data and Artificial Intelligence.” In *International Finance Review*. Vol. 20.
- Romer, Paul M. 1990. “Endogenous Technological Change.” *Journal of Political Economy* 98 (5, Part 2): S71–102.
- U.S. Bureau of Labor Statistics. 2024. “Local Area Unemployment Statistics (LAUS).” <https://www.bls.gov/lau/>.
- U.S. Census Bureau. 2024a. “American Community Survey (ACS) 5-Year Data.” <https://www.census.gov/data/datasets/time-series/econ/acs/acs-datasets.html>.
- . 2024b. “County Business Patterns (CBP).” <https://www.census.gov/programs-surveys/cbp/data/datasets.html>.
- . 2024c. “Quarterly Workforce Indicators (QWI).” <https://lehd.ces.census.gov/data/>.