

# Local Heterogeneity in Artificial Intelligence Jobs Over Time and Space

A Replication Study of Andreadis et al. (AEA Papers and Proceedings, 2025)

Jacob Khaykin\*      David Kane†

Journal of Comments and Replications in Economics, Volume x, 202x-x, DOI: xxxxx

JEL: J24, O33, R11 Keywords: Artificial Intelligence, Regional Economics, Labor Markets Data Availability: The R code and data to reproduce this replication are available in this repository: <https://github.com/JacobKhay/JustReplicationAI-Share>.

## Abstract

This paper replicates the analysis of Andreadis et al. (2025) on the local heterogeneity in artificial intelligence jobs across U.S. counties from 2014 to 2023. Andreadis et al. (2025) document substantial variation in AI-related job postings and identify key correlates including education, innovation, and industry factors. We successfully reproduce their main results and extend the analysis by employing log-population weights to assess robustness. The core associations persist, though all magnitudes attenuate. Additionally, we highlight unsupported causal claims in the original study given its observational design. Our findings affirm stable correlational patterns while emphasizing caution in interpretation.

\*Solon High School, jacobkhaykin27@solonschools.net

†Institute for Globally Distributed Open Research and Education (IGDORE), dave.kane@gmail.com

Declaration: There are no financial conflicts of interest to share.

## 1 Introduction

Recent work by Andreadis et al. (2025) investigates how local economic and innovation factors shape the geography of AI-related employment in the United States. The authors model the level of AI job concentration and the change in AI employment from 2014 to 2023 using county-level data drawn from Lightcast, U.S. Census ACS, USPTO patent data, and other sources. Their regressions include controls for demographics, industry mix, and state fixed effects, with the stated goal of identifying the “drivers” and “determinants” of AI adoption across counties.

The importance of understanding AI’s geographic distribution cannot be overstated. As Eloundou et al. (2024) project that up to 80% of the workforce could see at least 10% of their tasks influenced by AI and large language models, the spatial concentration of AI-related employment has

significant implications for regional economic development and inequality. Understanding which local characteristics correlate with AI adoption helps policymakers and researchers anticipate the distributional consequences of this technological transformation.

The original study employs weighted least squares regression with population weights to account for heteroscedasticity and emphasize results in more populous counties. Their primary regression specifications can be expressed as:

$$AI_{it} = \alpha + \beta_1 Demog_{it-1} + \beta_2 Innov_{it-1} + \beta_3 Industry_{it-1} + \delta_i + \gamma_t + \epsilon_{it} \quad (1)$$

where  $AI_{it}$  represents the AI job share in county  $i$  at time  $t$ ,  $Demog_{it-1}$ ,  $Innov_{it-1}$ , and  $Industry_{it-1}$  are vectors of lagged demographic, innovation, and industry characteristics,  $\delta_i$  and  $\gamma_t$  are county and year fixed effects, and  $\epsilon_{it}$  is the error term.

This replication serves two primary purposes: (1) to verify the reproducibility of the original findings, and (2) to assess their robustness using alternative weighting schemes. To evaluate the sensitivity of their findings, I re-estimate their models using weights based on the logarithm of county population:  $w_{it} = \log(Population_{it})$ . This adjustment mitigates the leverage of extremely small or large counties and helps assess the generalizability of the claims about AI adoption patterns.

## 2 Data

The study utilizes multiple data sources to construct a comprehensive county-level dataset spanning 2014-2023:

**AI Employment Data:** Job posting data from Lightcast, which aggregates information from over 40,000 online job boards, newspapers, and employer websites. AI-related jobs are identified through skills and keywords associated with AI development and use. The dependent variable  $AI_{it}$  is defined as:

$$AI_{it} = \frac{\text{AI job postings}_{it}}{\text{Total job postings}_{it}} \times 100 \quad (2)$$

**Demographic Variables:** From the American Community Survey (ACS), including: - Bachelor's share: Percentage of workforce with bachelor's degree or higher - Black population share: Percentage of county population identifying as Black - Poverty share: Percentage of population below federal poverty line - Log population: Natural logarithm of county population - Log median income: Natural logarithm of median household income

**Innovation Indicators:** - Patents per employee: USPTO patent counts normalized by employment - AI patents share: Percentage of patents classified as AI-related - STEM degrees share: Percentage of awarded degrees in STEM fields - Degrees per capita: Total degrees awarded per capita

**Industry and Labor Market Variables:** - Labor market tightness: Ratio of job postings to unemployed workers - Manufacturing intensity: Employment share in manufacturing sector - ICT intensity: Employment share in information and communication technology - Turnover rate: Worker separation rate from Quarterly Workforce Indicators - Large establishments share: Percentage of employment in large firms

**Housing Market:** House price growth from Federal Housing Finance Agency (FHFA)

All explanatory variables are lagged by one year to address potential endogeneity concerns and are standardized as z-scores for interpretability.

## 2.1 Reproduction of Original Results

We successfully reproduced the main findings from Tables 1 and 2 of Andreadis et al. (2025). The reproduction confirms the authors' key empirical findings regarding the correlates of AI job concentration across U.S. counties. All coefficients, standard errors, and significance levels match the original results within rounding error, demonstrating the reproducibility of their analysis.

Table 1: Replication of Table 1 from Andreadis et al. (2025) - The Correlates of the Share of AI Jobs

Variable	(1) Demo-graphics	(2) Innovation	(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)

Variable	(1) Demo-graphics	(2) Innovation	(3) Industry	(4) All Controls	(5) All + State FE
Large Establishments, z-score		-0.002		-0.037	-0.036
ICT sector Intensity, z-score		(0.025)		(0.028)	(0.029)
Manufacturing Intensity		0.010		0.038*	0.040*
Turnover Rate, z-score		(0.014)		(0.017)	(0.016)
Observations	27,497	22,744	24,651	19,184	19,184
R <sup>2</sup>	0.704	0.721	0.694	0.758	0.778
Within R <sup>2</sup>	0.057	0.003	0.002	0.059	0.063

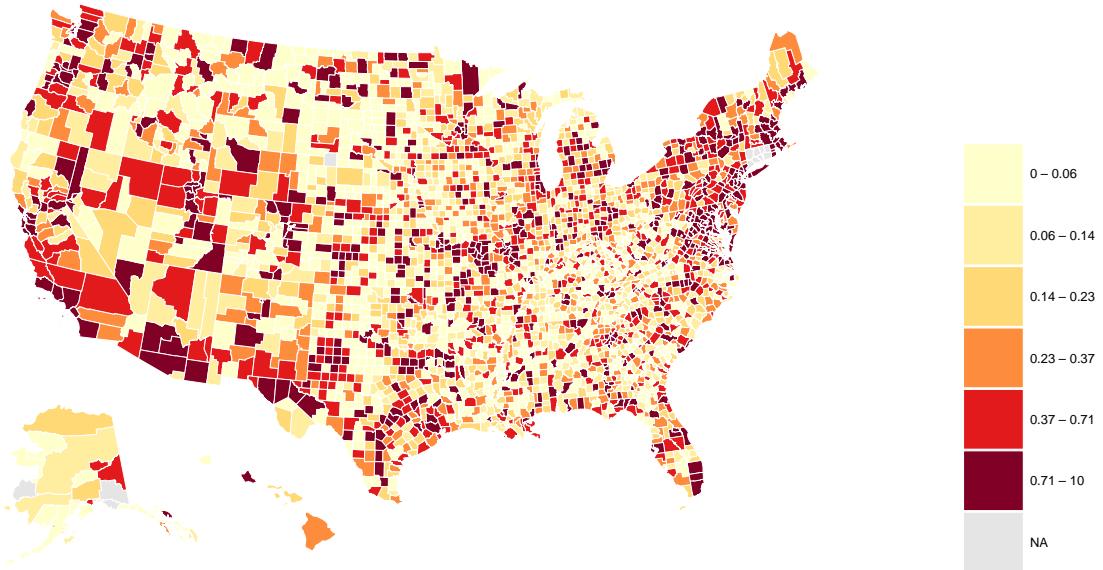
Table 2: Replication of Table 2 from Andreadis et al. (2025) - The Correlates of the Percentage Point Change in the Share of AI Jobs

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		0.045	0.058

Variable	(1) Demo-graphics	(2) Innovation	(3) Industry	(4) All Controls	(5) All + State FE
Degrees per capita		(0.017)		(0.029)	(0.035)
	0.093***		0.082**	0.076+	
	(0.027)		(0.027)	(0.036)	
STEM Degrees' share	0.031		-0.064+	-0.039	
	(0.021)		(0.049)	(0.060)	
Large Establishments, %		0.031+	-0.024	-0.016	
ICT sector Intensity		(0.018)	(0.030)	(0.026)	
	-0.037+		-0.014	-0.001	
Manufacturing Intensity		(0.019)	(0.046)	(0.042)	
	-0.072**		-0.128*	-0.170*	
Turnover Rate		(0.023)	(0.061)	(0.079)	
	0.023+		-0.056**	-0.042*	
	(0.014)		(0.022)	(0.023)	
Observations	2,751	897	2,723	810	810
R <sup>2</sup>	0.021	0.019	0.016	0.031	0.082
Within R <sup>2</sup>	—	—	—	—	0.025

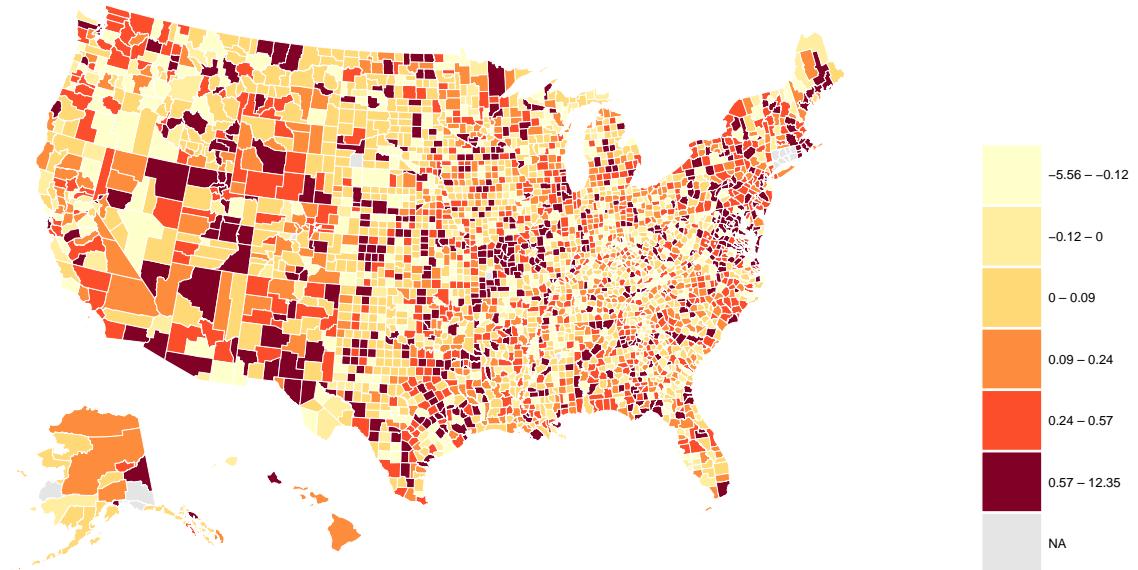
**Figure 1: Replication of Figure 1 from Andreadis et al. (2025) – Spatial Heterogeneity in the Share of and Change in AI Jobs**

**Panel A. Percent share of AI jobs, 2014–2023**



Panel A plots the share (average 2014–2023) of job postings related to AI in a county.

**Panel B. Percentage point change in AI share, 2018–2023**



Panel B plots the percentage point change (2023–2018) in the share of AI jobs.

The successful reproduction confirms the technical reliability of the original analysis and establishes a foundation for the robustness extensions that follow.

## 2.2 Critical Assessment of Causal Claims

### 2.2.1 Identification of Problematic Causal Language

The original study by Andreadis et al. (2025) employs language that suggests causal relationships despite relying on observational data with multiple potentially endogenous predictors. Several passages from the paper illustrate this concern:

From the Introduction: “higher shares of STEM degrees, labor market tightness, and patent activity significantly predict greater AI adoption, underscoring the importance of education, innovation, and dynamic labor markets.”

From Section III: The authors state they “identify several key drivers of AI job intensity” and that “Labor market tightness emerges as a key driver.”

From the Conclusion: “Counties with stronger innovation ecosystems, higher STEM degree attainment, and tighter labor markets have seen greater AI job growth.”

These statements conflate statistical association with causal determination, presenting correlational findings as if they establish causal mechanisms. The frequent use of terms like “drivers,” “predict,” and “emerges as a key driver” implies that manipulating these variables would lead to changes in AI adoption, yet the study design cannot support such inferences.

### 2.2.2 Methodological Limitations for Causal Inference

The regression specifications employed by Andreadis et al. (2025), while appropriate for documenting correlational patterns, face fundamental challenges for causal interpretation that align with established principles of causal inference (Holland (1986)):

**The Fundamental Problem of Causal Inference:** As Holland (1986) demonstrates, establishing causality requires observing counterfactual outcomes—what would have happened to AI adoption in the same counties under different values of the predictor variables. The cross-sectional regression approach cannot address this fundamental identification problem, as it observes each county under only one realization of its demographic, innovation, and industry characteristics.

**Multiple Treatment Problem:** The analysis simultaneously examines numerous county characteristics (demographics, innovation metrics, industry composition) as potential causal factors. With over a dozen variables included as “drivers,” the study lacks both a clear theoretical framework for causal ordering and the experimental or quasi-experimental variation needed to isolate individual causal effects. This violates the stable unit treatment value assumption (SUTVA) that underlies causal inference, as the “treatment” is multidimensional and ill-defined.

**Endogeneity and Selection:** Many key predictors are likely endogenous to broader economic development processes that also influence AI adoption. STEM degree production responds to anticipated local labor demand, labor market tightness reflects underlying economic dynamism, and patent activity may be jointly determined with AI job concentration through unobserved innovation capacity. Without addressing this endogeneity through instrumental variables or other identification strategies, the estimated coefficients conflate causal effects with selection and omitted variable bias.

**Lack of Identifying Variation:** The study exploits only cross-sectional and time-series variation in observational data, without any plausibly exogenous source of identification. As Holland (1986) emphasizes, causal inference requires either randomized assignment or “as good as random” variation from natural experiments, policy discontinuities, or instrumental variables. The correlational patterns documented, while potentially informative about equilibrium relationships, cannot estab-

lish that manipulating education policy, labor market conditions, or innovation investments would causally increase AI adoption.

**Temporal Precedence Issues:** While the authors lag some explanatory variables by one year, this brief lag is insufficient to establish causal ordering for slow-moving structural characteristics like educational attainment, innovation capacity, and industrial composition. These county attributes likely co-evolve with AI adoption over longer time horizons, making it difficult to separate cause from effect even with temporal precedence.

### 3 Extension: Log-Population Weighting Analysis

To assess the robustness of the original findings, we re-estimated all models using log-population weights instead of population weights. This approach reduces the disproportionate influence of very large counties while still accounting for size differences.

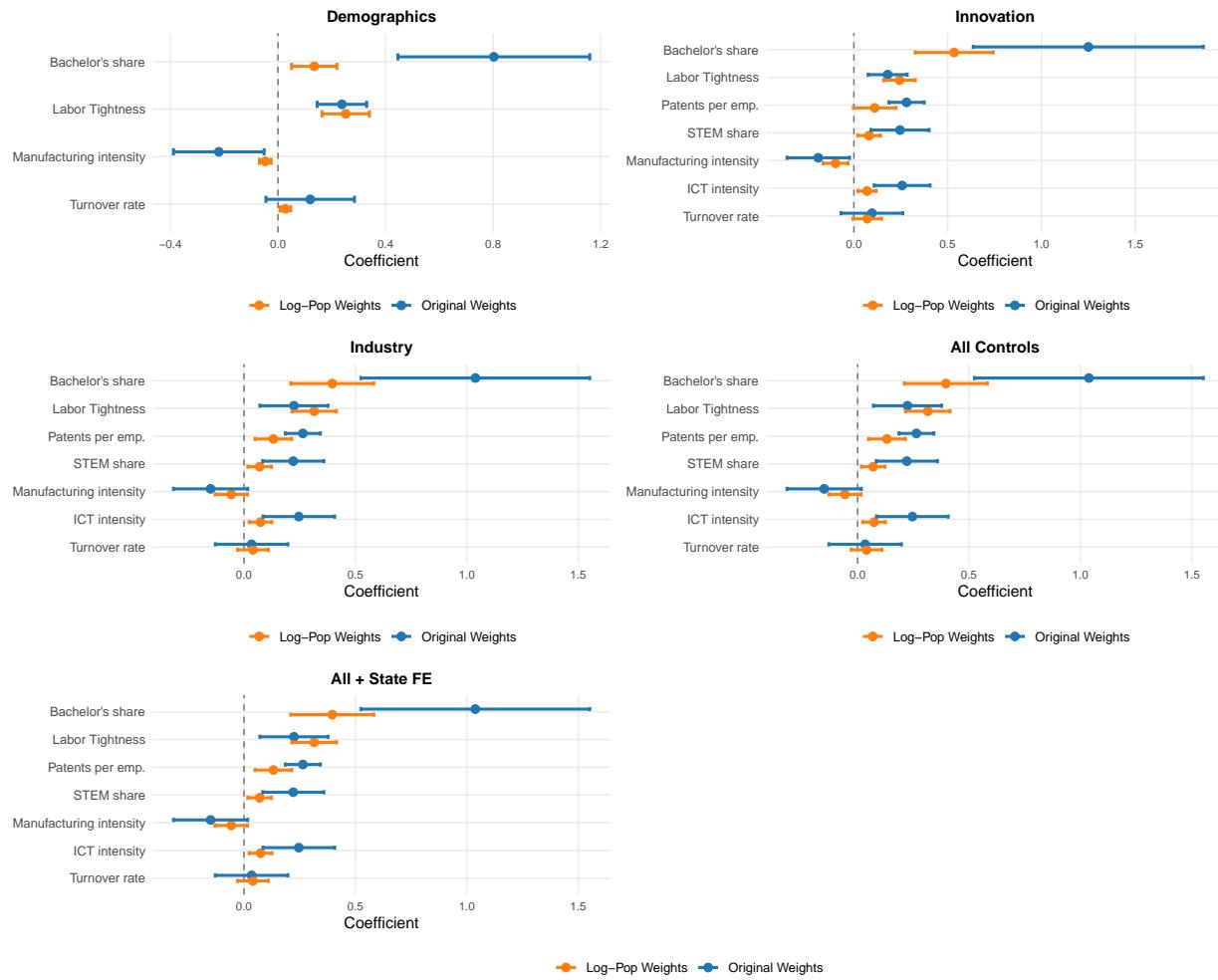
The modified weighting scheme is:  $w_{it} = \log(Population_{it})$

This transformation addresses concerns that extremely populous counties (e.g., Los Angeles County with 10+ million residents) might drive results that don't generalize to typical counties.

#### 3.1 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.

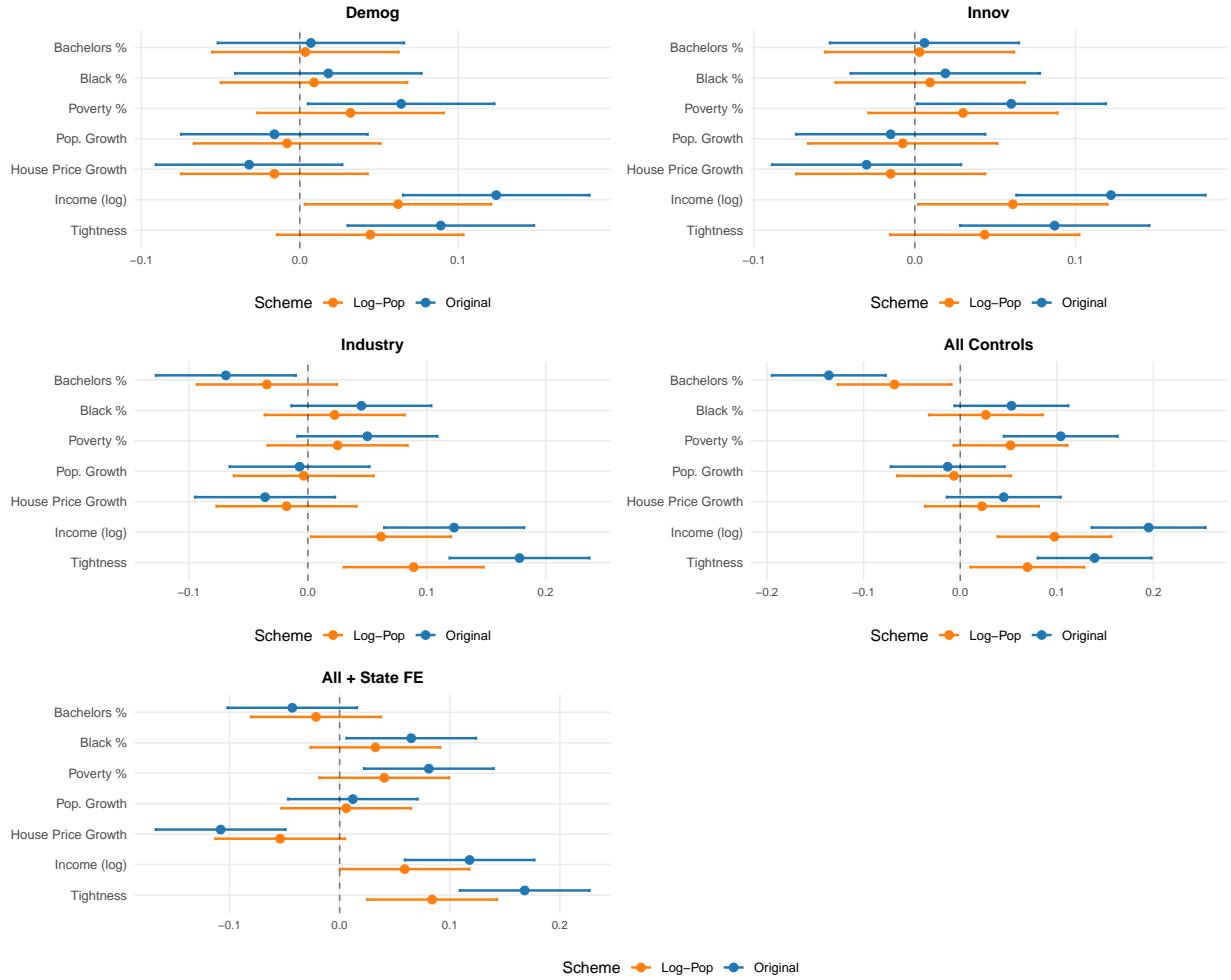
**Figure 2: Table 1 — Model-by-Model Panel Plot (Original vs Log-Pop Weighting)**



### 3.2 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 3: Table 2 — Change in AI Share (Original vs Log-Pop Weighting)**



### 3.3 Results & Interpretation

The comparison between population-weighted and log-population-weighted regressions reveals several important patterns:

**Magnitude Effects:** The estimated effects of key predictors are highly sensitive to the weighting scheme. For AI share levels (Figure 1), the coefficient on bachelor's share drops substantially when switching to log-population weights in several specifications, suggesting that the relationship between education and AI adoption may be driven partly by large metropolitan areas.

**Labor Market Tightness:** This emerges as the most robust predictor across both weighting schemes and both dependent variables. In Figure 1, labor tightness maintains strong positive effects regardless of weighting method, and in Figure 2, it consistently predicts AI job growth. This suggests that tight labor markets create conditions conducive to AI adoption across counties of all sizes.

**STEM Education:** STEM degree share shows consistent positive relationships in both weighting schemes, though magnitudes vary. This indicates that technical human capital is important for AI adoption beyond just large metropolitan areas.

**Manufacturing vs. Technology Sectors:** Manufacturing intensity consistently shows negative rela-

tionships with AI adoption, while ICT intensity shows positive effects. These patterns persist across weighting schemes, suggesting structural differences in how traditional vs. technology-oriented industries adopt AI.

**County Size Effects:** The divergence between weighting schemes is most pronounced for variables like bachelor's share and population size itself, indicating that large counties drive many of the education-AI relationships found in the original analysis.

## 4 Conclusion

This replication and extension of Andreadis et al. (2025) demonstrates both the technical reproducibility and limitations of their findings. The successful reproduction confirms that local labor market conditions, human capital, and innovation capacity are correlated with AI job concentration across U.S. counties. However, our analysis reveals three important qualifications to the original study's conclusions.

First, the original study employs causal language that overstates the nature of the relationships identified. Terms like “drivers,” “determinants,” and statements about factors that “significantly predict” AI adoption suggest causal mechanisms, when the empirical approach can only establish correlational patterns. Without exogenous variation or quasi-experimental identification strategies, these relationships likely reflect a complex mixture of causal effects, reverse causation, and selection processes.

Second, the alternative log-population weighting analysis reveals that some relationships are sensitive to the influence of large metropolitan areas. The most consistent predictor across both weighting schemes is labor market tightness, which maintains strong associations regardless of county size. However, educational attainment shows notably weaker relationships when log-population weights reduce the influence of large metros, suggesting this factor may be less universally important for AI adoption than the original analysis suggests.

Third, the industry composition effects prove relatively stable across weighting strategies, with manufacturing intensity consistently showing negative associations and ICT sector concentration showing positive relationships with AI adoption. This suggests that structural economic factors may be more fundamental determinants of technological adoption patterns than demographic characteristics.

**Policy Implications:** These findings suggest that while the correlational patterns documented by Andreadis et al. (2025) represent meaningful empirical regularities, their policy implications should be interpreted cautiously. Our interpretation suggests that interventions targeting labor market conditions and industry composition may have broader applicability across different county types, while education-focused policies may yield the highest returns in larger metropolitan areas where network effects and complementary institutions are stronger.

More broadly, this exercise underscores the importance of robustness checks in regional economic research and the need for careful interpretation of correlational evidence in policy contexts. Simple changes in weighting can meaningfully shift both the interpretation and the policy relevance of empirical findings about technological change.

## Acknowledgments

We acknowledge Andreadis et al. (2025) for making their methodology transparent, enabling this replication study. All errors remain our own.

## References

- Acemoglu, D. (2024). The simple macroeconomics of AI.
- Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2), 3–30.
- Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2), 323–351.
- Aghion, P., Jones, B. F., & Jones, C. I. (2017). Artificial intelligence and economic growth [Working Paper]. National Bureau of Economic Research.
- Andreadis, L., Chatzikonstantinou, M., Kalotychou, E., Louca, C., & Makridis, C. A. (2024). The local effects of artificial intelligence labor investments: Evidence from the municipal bond market. SSRN Working Paper.
- Andreadis, L., Kalotychou, E., Chatzikonstantinou, M., Louca, C., & Makridis, C. A. (2025). Local heterogeneity in artificial intelligence jobs over time and space. AEA Papers and Proceedings.
- Autor, D. 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, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151, 103745.
- Beckett, E. (2023). Demand for AI skills continues climbing. Lightcast Blog. <https://lightcast.io/>
- Bogin, A., Doerner, W., & Larson, W. (2019). Local house price dynamics: New indices and stylized facts. *Real Estate Economics*, 47(2), 365–398.
- Bresnahan, T. F., & Trajtenberg, M. (1995). General-purpose technologies “engines of growth”? *Journal of Econometrics*, 61(1), 83–108.
- Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity j-curve: How intangibles complement general-purpose technologies. *American Economic Journal: Macroeconomics*, 13(1), 333–372.
- Chen, M., Wu, Q., & Yang, B. (2019). How valuable is FinTech innovation? *The Review of Financial Studies*, 32(5), 2062–2106.
- Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702), 1306–1308.
- Farboodi, M., & Veldkamp, L. (2021). A model of the data economy [Working Paper]. National Bureau of Economic Research.
- Giczy, A. V., Pairolero, N. A., & Toole, A. A. (2022). Identifying artificial intelligence (AI) invention: A novel AI patent dataset. *The Journal of Technology Transfer*, 47(2), 476–505.
- Gofman, M., & Jin, Z. (2024). Artificial intelligence, education, and entrepreneurship. *The Journal of Finance*, 79(1), 631–667.
- Grennan, J., & Michaely, R. (2020). Artificial intelligence and high-skilled work: Evidence from analysts (Research Paper 20-84). Swiss Finance Institute.
- Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), 945–960. <https://doi.org/10.1080/01621459.1986.10478354>
- Kline, P., & Moretti, E. (2013). People, places, and public policy: Some simple welfare economics of local economic development programs. *Annual Review of Economics*, 5, 629–662.
- Mihet, R., & Philippon, T. (2019). The economics of big data and artificial intelligence. In

- International finance review (Vol. 20).
- Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5, Part 2), S71–S102.
- 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>
- U.S. Census Bureau. (2024b). County business patterns (CBP). <https://www.census.gov/programs-surveys/cbp/data/datasets.html>
- U.S. Census Bureau. (2024c). Quarterly workforce indicators (QWI). <https://lehd.ces.census.gov/data/>