

# 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†

*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

We replicate Andreadis et al. (2025) on the correlation between the level and growth artificial intelligence job openings and education, innovation, and industry factors across U.S. counties from 2014 to 2023. We successfully reproduce their main results and extend the analysis by employing log-population weights to assess robustness. The core associations persist, though most magnitudes shrink. We also highlight unsupported causal claims in Andreadis et al. (2025), given its observational design.

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

---

\*Solon High School, [jacobkhaykin27@solonschools.net](mailto:jacobkhaykin27@solonschools.net)

†Institute for Globally Distributed Open Research and Education, [dave.kane@gmail.com](mailto:dave.kane@gmail.com)

## 1 Introduction

This paper replicates the analysis of Andreadis et al. (2025) on the local heterogeneity in artificial intelligence (AI) jobs across U.S. counties from 2014 to 2023. Their study investigates how AI-related employment is distributed geographically and how this distribution evolves over time. By focusing on county-level variation, the work sheds light on which regions are gaining or lagging in access to AI-driven labor market opportunities, a subject of growing importance as artificial intelligence reshapes industries.

Andreadis et al. (2025) document substantial variation in AI-related job postings and identify key correlates of this variation. Counties differ widely in their shares of AI-related employment, and the analysis links these differences to educational attainment, local innovation ecosystems, and industry composition. Together, these findings suggest that access to AI employment is not uniform but instead reflects deeper structural characteristics of local economies.

We successfully reproduce the main results of the original paper by implementing the same empirical strategy and data structure. The baseline specification regresses the county-level share of AI jobs on covariates of interest, using the model

$$AI Share_{ct} = \alpha + \beta_1 Education_{ct} + \beta_2 Innovation_{ct} + \beta_3 Industry_{ct} + \gamma X_{ct} + \epsilon_{ct} \quad (1)$$

where  $AI Share_{ct}$  denotes the proportion of AI-related job postings in county  $c$  at year  $t$ , and  $X_{ct}$  represents additional controls. Our replication confirms that the core predictors remain significant and directionally consistent with those reported by Andreadis et al. (2025).

We then extend the analysis by applying log-population weights to assess robustness. The motivation for this adjustment is to reduce the disproportionate influence of very large counties while still reflecting the relative size of different local labor markets. Under this weighting scheme, the central associations persist, though magnitudes attenuate across the board, indicating that the original results are not entirely driven by large-population counties.

Finally, we highlight that some of the causal interpretations offered in the original study are not fully supported by its design. The analysis is based on observational data, which limits causal inference despite the suggestive correlations. While the evidence provides valuable insights into the geography of AI employment, we emphasize the need for caution in drawing strong causal conclusions.

## 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 Andreas 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 Artificial Intelligence Jobs

	Demographics	Innovation	Industry	All Controls	All + State FE
	(1)	(2)	(3)	(4)	(5)
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, z-score		0.013 (0.026)		0.034 (0.026)	0.034 (0.027)
STEM Degrees' share, z-score		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, z-score			-0.057*** (0.012)	-0.048*** (0.014)	-0.034** (0.015)
Turnover Rate, z-score			0.034*** (0.011)	0.016 (0.017)	0.010 (0.016)
<i>Fixed-effects</i>					
Year	Yes	Yes	Yes	Yes	Yes
County	Yes	Yes	Yes	Yes	Yes
State Year					Yes
<i>Fit statistics</i>					
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>	0.052	0.003	0.002	0.086	0.095

\* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

Notes: Sources: Lightcast, American Community Survey, Quarterly Workforce Indicators, 2014-2023. The table reports coefficients from regressions of the share of AI jobs in a county on Demographic, Innovation, and Industry Characteristics. Observations are weighted by log(1 + job postings) and standard errors are clustered at the county-level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.10

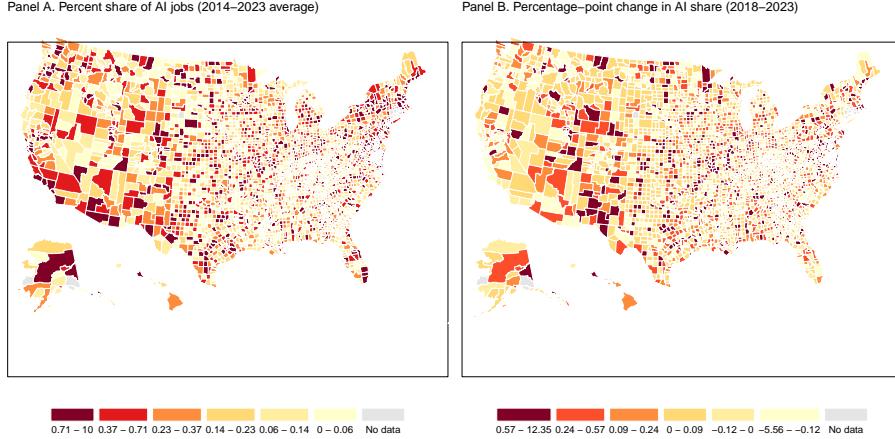
Table 2: Replication of Table 2 from Andreadis et al. (2025) - The Correlates of the Percentage Point Change in the Share of AI Jobs

	Demographics	Innovation	Industry	All Controls	All + State FE
	(1)	(2)	(3)	(4)	(5)
Bachelors,	(0.028)			(0.051)	(0.043)
Black,	(0.020)			(0.033)	(0.034)
Poverty,	(0.037)			(0.070)	(0.075)
Pop. Growth	-0.016			-0.007	-0.013
	(0.023)			(0.057)	(0.057)
House Price Growth z-score in 2017	-0.032*			-0.036	0.045
	(0.017)			(0.037)	(0.042)
Income, Log z-score in 2017	0.124***			0.123	0.195**
	(0.043)			(0.077)	(0.076)
Tightness, z-score in 2017	0.089***			0.178***	0.139
	(0.021)			(0.042)	(0.096)
Patents per employee z-score in 2017		0.008		-0.056	-0.001
		(0.029)		(0.037)	(0.043)
AI Patents' Share z-score in 2017		0.137***		0.108**	0.079
		(0.036)		(0.043)	(0.066)
Degrees awarded per capita, z-score in 2017		0.020		0.048*	0.062
		(0.016)		(0.028)	(0.037)
STEM Degrees' share, z-score in 2017		0.094***		0.083***	0.078**
		(0.027)		(0.027)	(0.036)
Large Establishments, ICT sector Intensity, Manufacturing Intensity, Turnover Rate,			(0.021)	(0.050)	(0.063)
			(0.018)	(0.031)	(0.026)
			(0.019)	(0.046)	(0.043)
			(0.022)	(0.061)	(0.081)
<i>Fixed-effects</i>					
State					Yes
<i>Fit statistics</i>					
Observations	2,751	897	2,723	810	810
R <sup>2</sup>	0.025	0.039	0.009	0.075	0.187
Within R <sup>2</sup>	—	—	—	—	0.065

\* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

Notes: Sources: Lightcast, American Community Survey, Quarterly Workforce Indicators, 2016–2023. The table reports the coefficients associated with regressions of the change in share of AI jobs in a county from 2017–18 (average) to 2022–23 (average) on Demographic, Innovation, and Industry characteristics. Observations are unweighted. Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Figure 1: Replication of Andreadis et al. (2025) - Spatial heterogeneity in AI job share (Panel A, 2014–2023 average) and percentage-point change (Panel B, 2018–2023).



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) frequently employs language that suggests causal relationships, even though the analysis relies on observational county-level data with multiple potentially endogenous predictors. Several passages illustrate this concern.

From the *Introduction*, the authors write:

> “Second, we identify several key drivers of AI job intensity, including demographics, innovation, and industry factors, after controlling for county and year fixed effects. Specifically, 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*, they conclude:

> “Labor market tightness emerges as a key driver, with a positive and highly significant coefficient... highlighting the importance of technical education and local innovation capacity in fostering AI job growth.”

And from the *Conclusion*:

> “Counties with stronger innovation ecosystems, higher STEM degree attainment, and tighter labor markets have seen greater AI job growth, whereas manufacturing-heavy regions and areas with high labor turnover have faced

challenges in integrating AI. These findings point to the role of place-based policies to attract and retain top-tier talent for economic development.”

Each of these statements frames correlational associations as if they were causal mechanisms. Phrases such as “drivers,” “emerges as a key driver,” “underscoring the importance,” and “findings point to the role of policy” imply that manipulating these variables would directly change AI adoption outcomes. However, the study’s empirical design—regressions on observational county characteristics—cannot identify such causal effects, only statistical associations.

### 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 equal weights. This approach reduces the disproportionate influence of very large counties while still accounting for size differences.

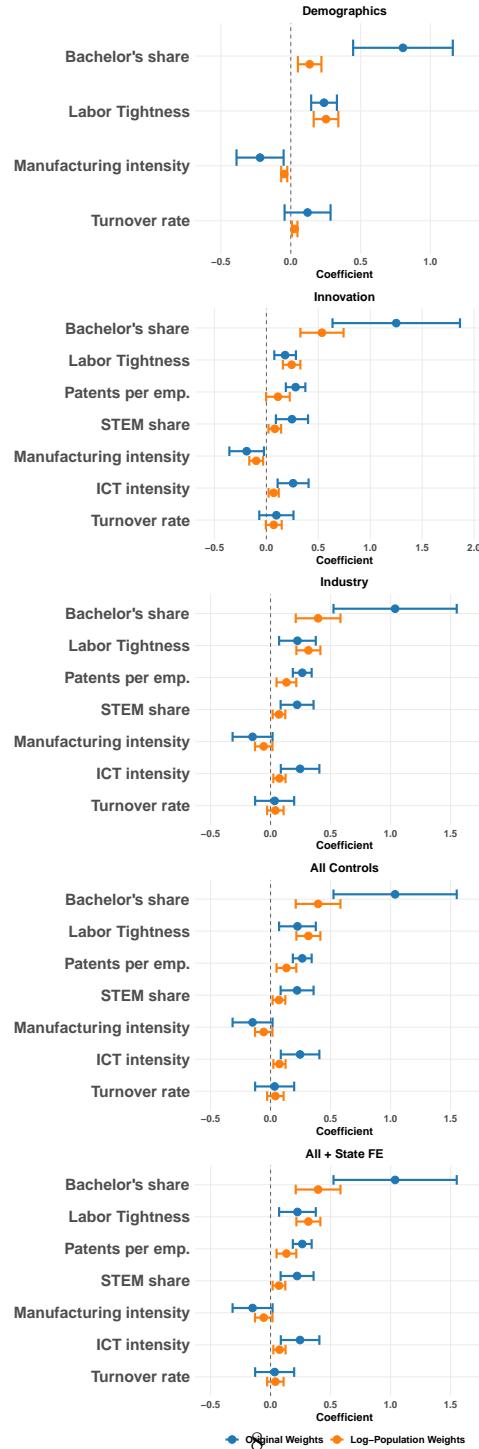
The modified weighting scheme is:  $w_{it} = \log(Population_{it})$

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 equal weights and one using  $\log(\text{population})$  weights. This comparison illustrates how model weighting influences the interpretation of each predictor.

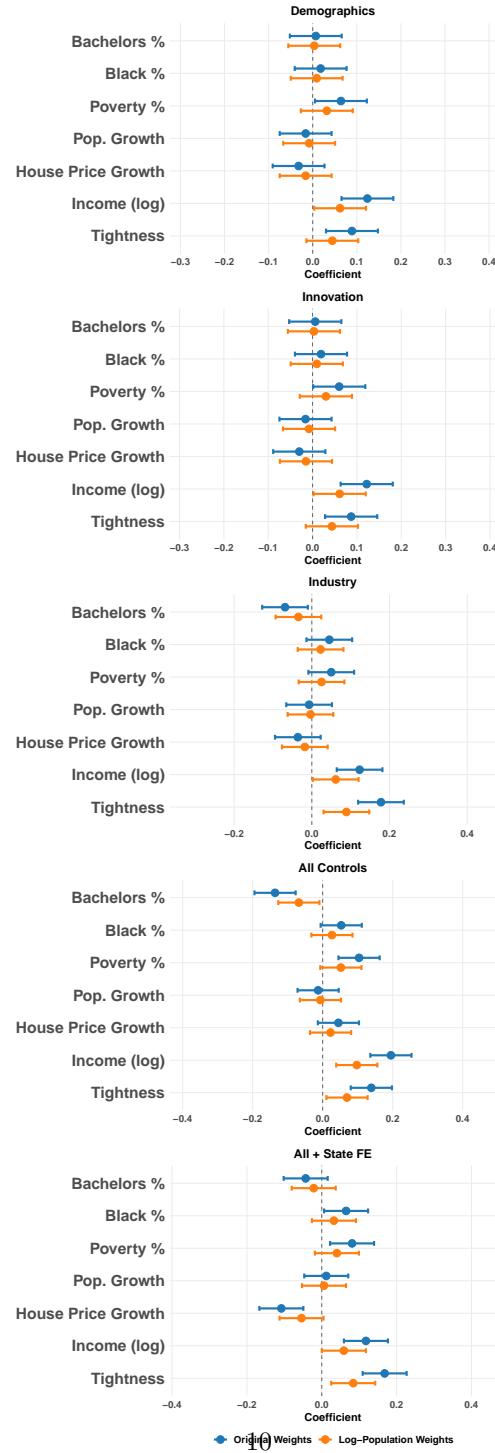
Figure 2: Table 1 — Coefficients under Original equal weights vs Log(population) weights.



### **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 equal-weighted and log(population)-weighted regressions.

Figure 3: Table 2 — Change in AI share under Original equal weights vs Log(population) weights.



### 3.3 Results

The comparison between equal-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 relationships 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 employ 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.

## 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., Pairolo, 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/>