Fall and Rise of Labor after Technological Revolutions: Substitutability and Scalability since 1800

Kevin Hjortshøj O'Rourke[1], Tancredi Rapone[2], Raül Santaeulàlia-Llopis[3]

July 2025, Bellaterra Macro Club

[1]Sciences Po

[2]Barcelona School of Economics (BSE) and Universitat Autònoma de Barcelona (UAB)

[3]New York University Abu Dhabi (NYUAD) and CEPR

Motivation

Motivation

- Rising concerns about the *future* of labor due to the disruption of new technology (Acemoglu and Restrepo 2019, Autor et al. 2020, Karabarbounis and Neiman 2013)
- Literature largely focuses on aggregate labor demand and supply (e.g. will tasks be performed by K or L?)
- In practice occupations are affected differently by technological change (e.g. Autor and Dron 2013) $\rightarrow$ substitutability vs scalability.

Q We revisit this important question (susbitutabiliy vs. scalability) in a context which witnessed among the largest technological disruptions and structural transformations ever observed in human history: the US during the Technological Revolution

Substitutability and Scalability

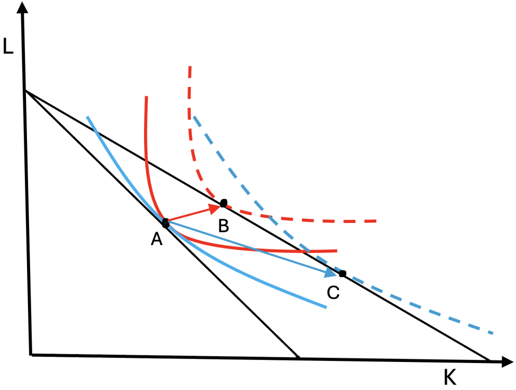


Figure 1: Isoquants from a CES production function for a complementary (red) and substitutable (blue) occupation in response to a fall in the price of capital r.

What We Do

- Build new database: arrival & diffusion of tech. innovations across time & space:
  - ▷ Source: Digitized *local* newspapers in the U.S. (1820-1960) (Library of Congress)
  - ▷ *Equipment ads* with pricing and *Job ads*
- Document the dynamic response of labor to new technology and the evolution of county level aggregates:
  - ▷ Labor: rise and fall of occupations
    - high frequency data; potentially more granular occ. than HISCO/Census.
    - cross-validated with population census data.
  - ▷ County level aggregates: using the census of manufacturing we quantify the effect of technological shocks on productivity, wages, capital, employment and the labor share.

What We Find

- New technologies between 1860–1950 reshaped labor markets by **substituting** and **scaling** occupations.
  - **Heterogeneous** Effects on occupations: fall (e.g. coachmen) and rise (e.g. drivers).
    - ▷ Staggered DID (Time × Space Diffusion in LOC and Population Census)
    - ▷ Instrumenting (TBD)
  - Focusing on Manufacturing Census: On average, new technologies:
    - ▷ Increase employment, wages, and productivity
    - ▷ Unaffected labor share in manufacturing
- Build a model that rationalizes these effects (TBD)
- Help-wanted ads (demand) respond faster Census: (Lack of) Adaptability?
  - ▷ Counties with **higher adaptability**: Are richer and more employed in 2000
  - ▷ Show lower high–low skill polarization
- Overall: technologies **reallocate** labor rather than **replace** it.

A New Historical Database of Technology and Occupations

A New Database

We use two data sources:

1. Library of Congress: *Chronicling America*
    - Public domain repository of over 21 million newspaper pages published between 1756 and 1963.
    - National coverage.
    - ▷ We extract approximately 400 thousand pages containing equipment ads and 550 thousand pages containing help wanted ads.
2. Census of population, manufacturing and agriculture for 1850-1960:
    - Data on employment, wages, output and capital for the manufacturing sector.
    - Detailed information on occupations and industry from the Census of population.

◂ map

Technological Innovations Overview

- **Overview:** 123 technologies spanning 3rd century BCE –early 20th century
  - Force Pump (3rd c. BCE) to the Television (1927).
- **Sectors:**
  - Manufacturing: Steam Engine, Airplane, Assembly Line
  - Agriculture: Seed Drill, Combine Harvester, Steel Plow
  - Services: Telegraph, Telephone, Phonograph
- **Origins:** Europe (England, Germany, France) $\rightarrow$ USA (19th–20th century.)
- **Trends:**
  - Manual/water power $\rightarrow$ steam $\rightarrow$ internal combustion
  - Advent of electricity: Light Bulb, Electric Motor
  - Rise of mass production: Assembly Line, Punch Press

Why this Time Period?

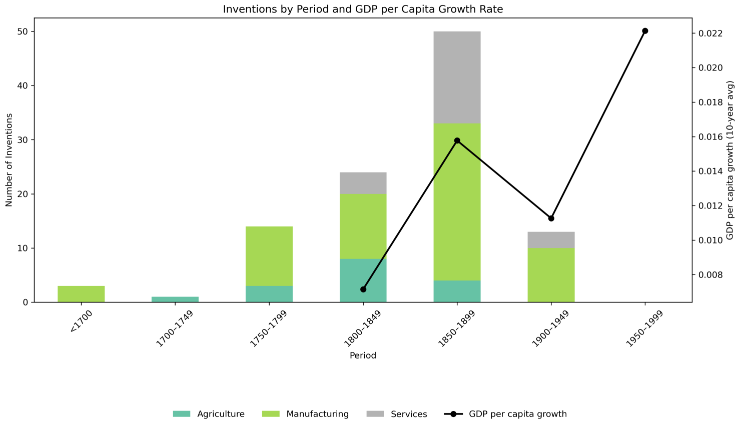


Figure 2: Number and sector of inventions for 50 year time periods shown together with the average GDP per capita growth rate. Data sources: data on inventions is from author's calculations using the Library of Congress (LOC), data on GDP per capita is from Maddison 2023.

Examples: Technology Ad

(a) Saw Replaces 10 Men

(b) Loader Replaces 40 Men

Figure 3: Source: (left) The St. Mary banner. (Franklin, Parish of St. Mary, La.), November 20, 1920.; (right) Montana farmer-stockman. [volume] (Great Falls, Mont.), December 15, 1948.

Examples: Technology Ad (AI-Like)

(a) Loom (AI-Like)

(b) Cultivator (AI-Like)

Figure 4: Source: (left) The Minneapolis journal. [volume] (Minneapolis, Minn.), 01 March 1904.; (right) The log cabin Democrat. [volume] (Conway, Ark.), 02 March 1916.

Source of Variation 1: Technology Diffusion across Time and Space

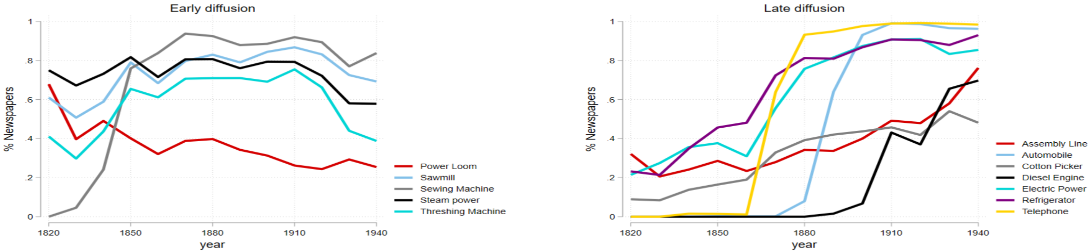
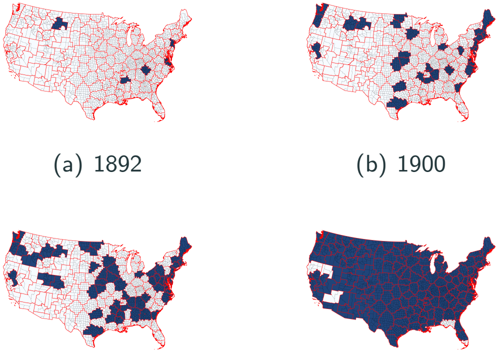


Figure 5: Fraction of newspapers mentioning each technology over time. Source: Author's calculations using Library of Congress, *Chronicling America*.

Source of Variation 2: Technology Diffusion across Time and Space



(a) 1892 (b) 1900

(c) 1908 (d) 1916

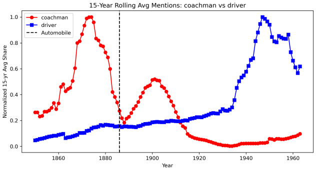
Figure 6: Diffusion of the tractor over time. Newspapers are clustered into 150 groups and each county is assigned to the closest newspaper group. Red lines show the boundaries between groups. Source: Author's calculation using Library of Congress newspapers chroniclingamerica.

▸ Telephone

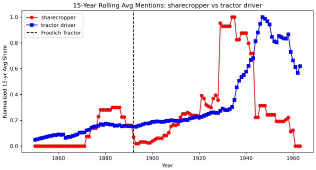
Help Wanted Ads: Labor Demand

- **Occupations & Triggers:**
  - Start with 1950 census occupations (occ1950)
  - Define for each occupation a set of "trigger" words/phrases
- **Text Mining:**
  - Search pages for instances of the occupation or its trigger words.
- **Counting:**
  - For every occupation, flag ads containing its name or any trigger
  - Aggregate counts by year and location (each county is matched to a newspaper group)

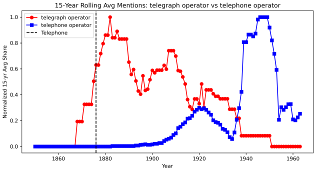
Help Wanted Ads: Descriptives



(a) Automobile



(b) Tractor



(c) Telephone

Figure 7: Normalized 15-year moving average of help wanted ads (1860–1960) for selected professions. Vertical lines show the introduction of relevant technologies. Source: Library of Congress *Chronicling America*.

Empirics

Empirical Strategy 1/2

We are interested in estimating the following regression model:

$$y_{it} = \alpha_i + \delta_t + \sum_{k=-2}^{2} \beta_k \, D_{t>\tau_i - k} + \epsilon_{it}, \tag{1}$$

Where y_{it} is a county level outcome variable. We use the Borusyak et al. (2024) estimator to account for staggered treatment.

Empirical Strategy 2/2

We construct a county level measure of adaptability to technological shocks:

$$M_i = \underbrace{\frac{1}{JT}\Big(\sum_j \sum_{o\in\mathcal{S}_j} \sum_{t=-k}^{k} (s_{o,a,t} - s_{o,c,t})\Big) \mid s_{o,a,t} > s_{o,c,t}}_{M_p} - \underbrace{\frac{1}{JT}\Big(\sum_j \sum_{o\in\mathcal{S}_j} \sum_{t=-k}^{k} (s_{o,a,t} - s_{o,c,t})\Big) \mid s_{o,a,t} \leq s_{o,c,t}}_{M_n}$$

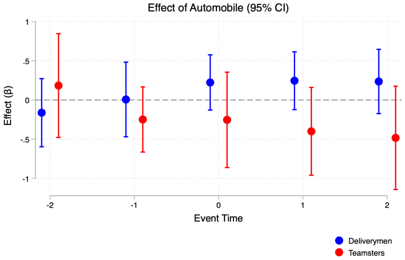
And correlate it with long-run county level outcomes:

$$y_{i,2000} = \alpha_s + \beta\, M_i + \Gamma' X + \epsilon_{it} \tag{2}$$

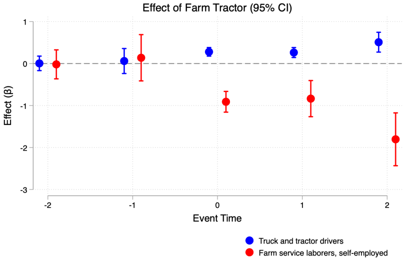
We use copyright registrations between 1840 and 1860 as an IV for M. ▸ Controls

Results

Employment: Granular Treatment Events



(a) Automobile



(b) Tractor

Figure 8: Event study plots of the effect of an innovation on two occupations. Dependent variable is log of employment for that occupation in the census. Data source: US Census of Population 1860–1940 and author's calculations using Library of Congress *Chronicling America*.

▸ Telephone

Employment: Occupation Level Treatment Events

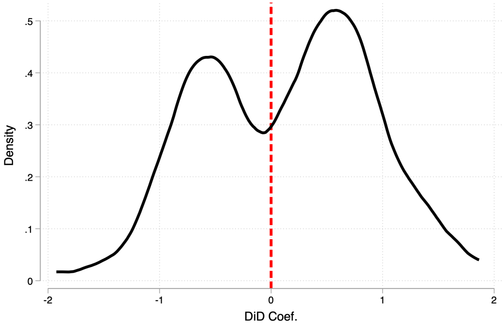


Figure 9: Distribution of $\hat{\beta}$ coefficients estimated from difference-in-differences regressions of model 1. Dependent variable is log of employment at the occupation level. For each innovation we consider 10 relevant occupations which may be negatively or positively affected by the technology. Data source: US census of population 1860-1940 and author's calculations using Library of Congress *Chronicling America*.

▸ Het. Skill ▸ Het. Sector

Manufacturing Census

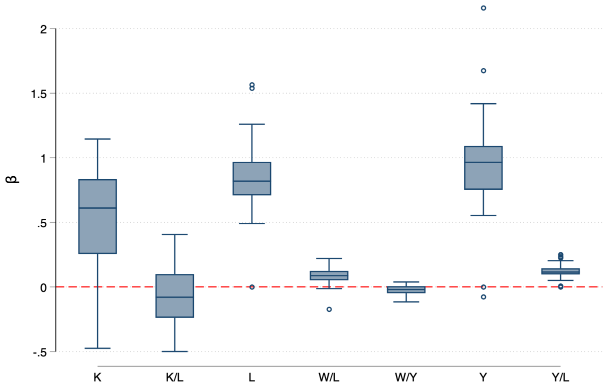


Figure 10: Distribution of the treatment effects for different outcome variables. Each box plot shows the inter quartile range and outliers of $\hat{\beta}$ coefficients estimated from difference-in-differences regressions of model 1 using ten-year census intervals.

Employment: Census vs Ads

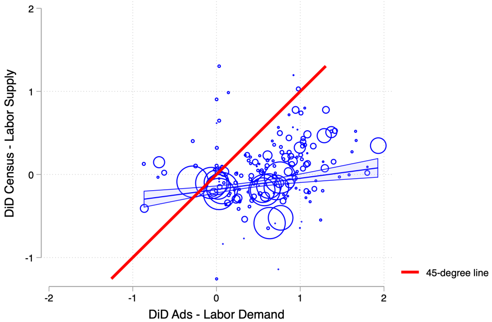


Figure 11: Scatter plot of treatment effects $\hat{\beta}$ from difference-in-differences regressions (model 1), with the log of census (labor supply) and help wanted ads (labor demand) as outcomes. Observations are weighted by final-period labor supply. Each innovation includes 10 occupations potentially affected (positively or negatively). Data: US Census 1860–1940 and author's calculations from Library of Congress *Chronicling America*.

OLS and IV Results: Long-Run Outcomes

Table 1: Effect of Misalignment on Economic Outcomes

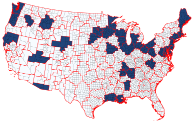
	(1) Log. Income	(2) Unemp.	(3) Share HS	(4) Share College
Panel A: OLS Estimates				
Misalignment	-0.523***	0.003	-0.058***	0.143***
	(-3.89)	(0.31)	(-2.85)	(3.06)
R^2	0.226	0.316	0.513	0.302
Panel B: IV Estimates				
Misalignment	-2.176**	0.148***	-0.470***	0.700***
	(-2.46)	(2.67)	(-3.96)	(3.61)
N	1087	1087	1087	1087
R^2	0.141	0.149	0.276	0.075

t statistics in parentheses. Standard errors computed using the Conley correction with a 100km cutoff.

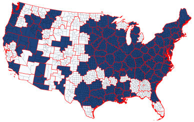
Conclusion

- We construct a county level dataset of technology diffusion and labor demand for the US over 1800/1960 covering 123 technologies.
- Technological innovations between 1860–1950 displaced and scaled different occupations — **reallocation, not replacement**.
- Average effects on labor markets are **positive**: higher wages, productivity, and employment. Labor share unaffected.
- Regions that adapted faster are richer, more polarized, and less unemployed by 2000.
- From a historical perspective, we offer **cautious optimism** about current technological transitions.

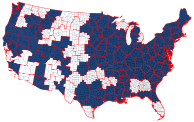
Source of Variation 2: Technology Diffusion across Time and Space



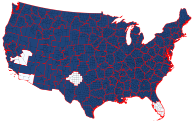
(a) 1876



(b) 1878



(c) 1880



(d) 1900

Figure 12: Diffusion of the teleophone over time. Newspapers are clustered into 150 groups and each county is assigned to the closest newspaper group. Red lines show the boundaries between groups. Source: Author's calculation using Library of Congress newspapers chroniclingamerica.

Back

Map of Newspaper Coverage

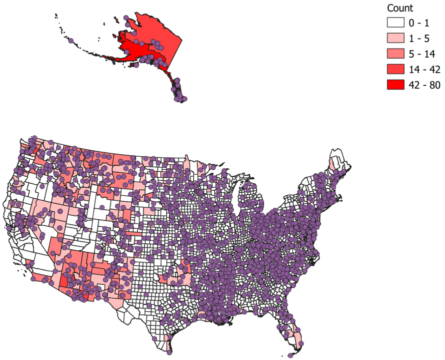


Figure 13: Map of newspapers at the county level. Data source: Library of Congress *Chronicling America*.

Back

Employment: Occupation Level Treatment Events

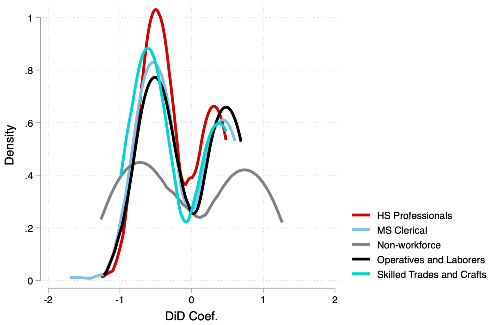


Figure 14: Distribution of $\hat{\beta}$ coefficients estimated from difference-in-differences regressions of model 1. Dependent variable is log of employment at the occupation level. For each innovation we consider 10 relevant occupations which may be negatively or positively affected by the technology. Data source: US census of population 1860-1940 and author's calculations using Library of Congress *Chronicling America*.

Back

Employment: Occupation Level Treatment Events

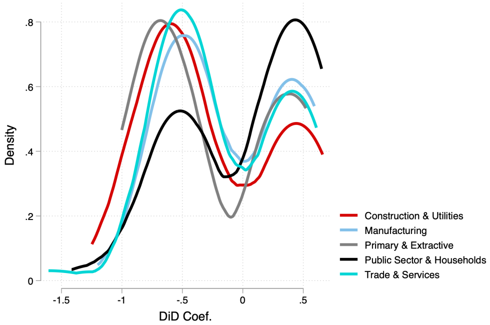


Figure 15: Distribution of $\hat{\beta}$ coefficients estimated from difference-in-differences regressions of model 1. Dependent variable is log of employment at the occupation level. For each innovation we consider 10 relevant occupations which may be negatively or positively affected by the technology. Data source: US census of population 1860-1940 and author's calculations using Library of Congress *Chronicling America*.

Back

Employment: Granular Treatment Events

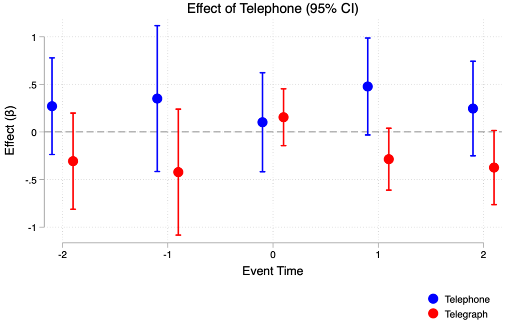


Figure 16: Event study plot showing the effect of the tractor on employment in a selected occupation. Dependent variable is log employment from the census. Data: US Census of Population 1860–1940 and author's calculations using Library of Congress *Chronicling America*.

Back

Controls

Table 2: Summary of Control Variables

Type of Control	**Variables**
Geographic	Potential yields
	Average rainfall
	Average temperature
	Distance to oceans and the Great Lakes
	Terrain ruggedness
Human Capital (measured in 1850)	Literacy rate
	School enrollment rate
Spatial Controls	Third-degree polynomial of longitude and latitude

Back