The Social Consequences of Technological Change: Evidence from U.S. Electrification and Immigrant Labor*

Sara Benetti†

November 2, 2025
Updated Frequently. **Click Here for Most Recent Version.**

Abstract

This paper examines how technological change in production processes affects social cohesion in ethnically diverse societies. I study the early expansion of the electric grid in the United States between 1900 and 1940, when electrification transformed manufacturing and large-scale immigration reshaped the labor force. Using newly digitized maps of the U.S. high-voltage transmission network linked to full-count census data, I exploit the staggered rollout of electrification across counties to estimate its causal effects on the integration of immigrant and native workers. Electrified industries became more diverse and less segregated along ethnic lines. These effects extend beyond the workplace. Electrification is associated with lower residential segregation among manufacturing workers and a partial attenuation of the negative relationship between immigrant presence and local public service provision. Overall, I find that by reshaping production, technological change can also reshape the social fabric, promoting integration both at work and within local communities.

Keywords: Technological Change, Social Cohesion, Electrification, Immigration, Manufacturing. **JEL Codes:** J15, J61, O33, N32, R23.

*I am indebted to my committee members Nathan Nunn, Mauricio Drelichman, Patrick Francois, and Réka Juhász for their extensive advice and guidance throughout this project, as well as to Felipe Valencia for his support. For helpful comments and conversations, I thank Mike Andrews, Siwan Anderson, Pierre-Loup Beauregard, Enrico Berkes, Alberto Bisin, Martina Björkman, Jiwon Choi, Jevan Cherniwchan, Francisco Eslava, Claudio Ferraz, Andy Ferrara, Price Fishback, Martin Fiszbein, Giovanni Gallipoli, Michela Giorcelli, David Green, Stephan Heblich, Kisho Hoshi, Daniel Jaramillo, Taylor Jaworski, Yannis Kastis, Carl Kitchens, Thomas Lemieux, Jeanne Lafortune, Elijah Locke, Eleonora Nillesen, Samuel Norris, Miguel Ortiz, Tommaso Porzio, Pascuel Plotkin, Tzachi Raz, Angela Redish, Luca Repetto, Federico Ricca, Diego Ramos-Toro, Raffaele Saggio, Edson Severnini, Clara Sievert, Henry Siu, Munir Squires, Catherine van der List, Daniela Vidart, and Randall Walsh. For useful feedback, I also thank conference participants at ASREC (Copenhagen 2025), CEA (Montréal 2025), CNEH (Montréal 2025), EHA (Philadelphia 2025), NBER Summer Institute, DAE (2025), and UEA (Montréal 2025), as well as seminar participants at several UBC lunches, UMBC, and the Virtual Economic History Workshop. I am grateful to Raju Ahmed for excellent research assistance, to Alex Alisauskas and Lily Demet of the UBC Geography Department for their help during the digitization process, and to Leanne Koch at the UBC Library for her invaluable help in securing access to the EEI maps. I thank Juan José Céspedes for his constant support and for his assistance in locating historical sources. Financial support from the Economic History Association and the Stone Centre for Wealth and Income Inequality at UBC is gratefully acknowledged.

†University of British Columbia, Vancouver School of Economics. Email: sara.benetti.cespedes@gmail.com.

1 Introduction

Technological change drives economic growth by introducing more efficient means of production. In addition to economic effects, technological transformations can also have profound social consequences. Innovations that alter production processes, in particular, can reshape how firms are organized, how people work, and how they interact within their communities (Tönnies, 1887; Durkheim, 1893; Weber, 1922). A priori, however, it is unclear whether such changes reinforce existing social divisions or foster greater cohesion. Understanding this relationship is especially important in culturally diverse societies, where social cohesion may be important for growth but is challenging to achieve (Alesina et al., 1999; Alesina & La Ferrara, 2000; Easterly et al., 2006). This paper studies how technological change in production processes affects the social cohesion of ethnically diverse societies.

I explore this relationship in the context of the early expansion of the electric grid in the United States between 1900 and 1940. Electrification was the pivotal technological change of the Second Industrial Revolution, and fundamentally transformed production processes in manufacturing (David, 1990). This transformation unfolded during a period of remarkable cultural heterogeneity in the U.S., shaped by decades of open immigration policies and the arrival of millions of immigrants from across Europe (Hatton & Williamson, 1998). Against this background, I examine how electrification affected social cohesion between native and immigrant manufacturing workers. I study whether the new technology changed the ethnic composition of the workforce and reduced occupational segregation within manufacturing industries. I then assess whether electrification influenced the residential segregation of manufacturing workers and local provision of public services, providing insight into its broader impact on community integration.

In the early twentieth century, industrial electrification transformed factory production by replacing centralized steam engines with decentralized electric motors. Under steam power, machines had to be placed close to the central engine to reduce energy loss. Electrification altered this constraint. Once factories could draw affordable and reliable power from external sources, each machine could be equipped with its own motor, allowing production to be arranged around the sequence of operations (Du Boff, 1967; Devine Jr, 1983). As a result, production coalesced around fixed workstations, reducing the need for coordination and movement on the factory floor and giving workers greater task independence (Nye, 2013). The new production system made it easier for employers to hire unskilled and immigrant workers, as tasks required less technical ability, coordination, or shared language. With each worker operating more autonomously, the ethnic background of nearby co-workers mattered less, potentially encouraging more heterogeneous workgroups. As a result, the increased task independence introduced by electrification may have promoted workforce integration within industries, potentially fostering informal interactions and gradual social integration in ethnically mixed communities. Still, greater task independence did not guarantee integration. Employers might have continued to group workers by ethnicity to avoid tensions or because they believed homogeneous teams were more efficient. As a result, ethnic divisions within the industry could have persisted despite the organizational shift. Whether the technological changes in production brought by electrification actually promoted integration is ultimately an empirical question, which I address in this paper.

To measure electrification, I digitize a series of historical maps published by the Edison Electric Institute (EEI, 1962), which trace the expansion of the U.S. high-voltage transmission grid between 1908 and 1946, and maps reporting the locations of major electric power plants, published by the U.S. Department of Commerce (1912) and the U.S. Federal Power Commission (1935). By combining these sources, I classify a county as electrified in a given decade if it is connected to the high-voltage grid or is located near a major power plant.[1] I then exploit the staggered expansion of the U.S. electric grid to identify the causal effect of electrification on multiple dimensions of social cohesion between immigrant and native workers.

To examine the employment consequences of electrification, I use individual-level census data aggregated to the industry–county level for the period 1900–1940 (Ruggles et al., 2024). In this paper, "ethnicity" refers to race for U.S.-natives and to country of birth for foreigners, rather than ancestral origin. For example, the "Italian ethnic group" includes people born in Italy who live in the United States, while U.S.-born individuals of Italian descent are classified as U.S.-born Whites. Accordingly, I group workers as U.S.-born Whites, U.S.-born Blacks, and foreign-born individuals ("immigrants") by country of birth.[2]

I use census information on industry and occupation to measure the ethnic characteristics of the labor force in each manufacturing industry, by county and decade. I calculate the share of workers belonging to each ethnic group. Then, I compute an index of ethnic diversity, which measures the probability that two randomly selected workers within an industry belong to different ethnic groups (Alesina et al., 2016). This index captures overall heterogeneity of workforce but not the degree of integration among groups. Two industries could have similar diversity levels, yet differ greatly in how workers from different ethnic backgrounds are distributed across occupations, and ultimately in the degree of ethnic mixing workers actually experience within the industry. To capture this dimension, I construct an index of ethnic segregation across occupations within each industry, adapted from Alesina and Zhuravskaya (2011). This index measures how unevenly ethnic groups are distributed across occupations relative to their overall share in the industry. It ranges from complete integration, where every occupation mirrors the industry's overall composition, to complete segregation, where each occupation is entirely homogeneous.

To assess the causal effect of electrification on the ethnic composition and integration of the labor force in manufacturing industries, I estimate an event-study and a difference-in-differences (DID) model. The unit of observation is a manufacturing industry in a county and decade. The sample includes 49 manufacturing industries, all counties in the continental United States, and covers the period 1900–1940, yielding to 350 thousand unique observations. The treatment variable is an indicator that equals one once a county becomes electrified and its manufacturing industries gain access

[1] I exclude the largest urban centers in 1900 from the analysis because they had access to local electric transmission through urban power plants by the late 1890s and early 1900s (Cohn, 2017), while the first available map on power plants dates from 1912.

[2] Combining race for U.S.-born individuals and country of birth for immigrants yields 42 distinct ethnic groups.

to the transmission network. The empirical specification includes industry–county fixed effects to control for time-invariant local characteristics –such as geography, natural endowments, or initial industrial structure– that may jointly influence electrification and industrial growth. State–decade fixed effects absorb time-varying state-level shocks and policies, including grid expansion, federal programs, immigration policy changes, and wartime mobilization. Industry–decade fixed effects capture national industry trends and differential rates of electricity adoption across industries. I also include the logarithm of the 1900 county population interacted with decade indicators to account for differences in initial size. This specification isolates plausibly exogenous variation in the timing of electrification by comparing the same industry across counties that gained access to electricity with those that are not yet electrified. I estimate both the event-study and DID models using the imputation-based method of Borusyak et al. (2024), which provides unbiased estimates under staggered treatment adoption. The approach relies on the standard DID assumptions of parallel trends and no anticipation before treatment, for which I provide supporting evidence, and identifies the average treatment effect on the treated (ATT).

My results are consistent with historical accounts suggesting that the reorganization of production brought about by electrification increased the integration of immigrant workers in manufacturing. Electrification had a positive and statistically significant effect on the share of immigrant workers in manufacturing industries. The group that lost relative representation was U.S.-born Whites, while the effect on U.S.-born Blacks is statistically indistinguishable from zero. The gains were concentrated among immigrants who were culturally and linguistically more distant from the U.S.-born majority –particularly those from Southern and Eastern Europe and from non-English-speaking or non-Protestant countries. These groups faced the greatest communication and cultural barriers and thus benefited most from the reduced need for coordination and communication that electrification afforded. As a result, the ethnic diversity of the workforce in manufacturing industries significantly increased.

Electrification made industries more heterogeneous overall. However, the key question for social cohesion is whether it also promoted integration within industries. I find that electrification significantly reduced ethnic segregation across occupations. This indicates that electrification not only increased overall diversity, but also fostered greater mixing within occupations, expanding opportunities for integration of workers of different ethnic backgrounds within manufacturing industries. To my knowledge, this provides the first systematic measurement of ethnic segregation across occupations in U.S. manufacturing and offers new evidence on an important but previously unexplored dimension of labor organization.[3] I find considerable variation across industries in the effect of electrification on the integration of ethnically diverse workers. The reduction in ethnic segregation is stronger in industries that were more energy-intensive and in those with greater scope

[3]Margo (1990) calculates a dissimilarity index between Black and White workers in the U.S. South from 1910 to 1940, computed separately by industry and occupation. Tomaskovic-Devey et al. (2006) and Hellerstein and Neumark (2008) examine workplace segregation starting from the 1960s using matched employer–employee data. Xu and Zhang (2022) and Locke (2025) study ethnic–occupational niches in craft and trade occupations during the Age of Mass Migration.

for reorganization, as they already operated with large establishments before electrification.

I perform several tests to assess the validity of the identification strategy and the robustness of the results. A key concern is that the expansion of the grid was not random, as electrification followed geographic and industrial patterns. To address this, I include industry–county, state–decade, and industry–decade fixed effects, which control for time-invariant local characteristics, regional shocks, and national industry trends. The results remain robust when excluding early-electrified areas, hydro-intensive regions, and large urban centers. I also examine potential endogeneity related to manufacturing intensity and county demographics. Electrification is uncorrelated with pre-existing immigrant shares or the ethnic diversity of the county. It also does not affect the overall ethnic composition of the county population, including the shares of foreign-born and Black population, or ethnic diversity. Thus, the estimated industry-level effects are not driven by broader demographic shifts. The results hold when controlling for baseline and time-varying ethnic composition, as well as for manufacturing intensity, and remain stable when excluding counties with extreme initial shares of manufacturing employment.

Having established that electrification had a causal impact on the integration of ethnically diverse workers within manufacturing industries, I next examine whether it also had effects on social dynamics beyond the workplace. I focus on the community consequences of electrification. A key element of cohesive local communities is the spatial mixing of individuals from different backgrounds, which helps prevent the formation of segregated neighborhoods (Cutler & Glaeser, 1997). Using census data, I calculate an index of residential segregation of manufacturing workers, which captures the degree of ethnic integration across enumeration districts (a proxy for neighborhoods) within each county (Alesina & Zhuravskaya, 2011).

A second key dimension of social cohesion is a community's collective orientation toward the common good, reflected in its willingness to contribute to public goods and services (Schiefer & van der Noll, 2017). Because detailed data on county public finances are unavailable for this period, I proxy local public service provision using the number of workers employed in public service occupations. Using census data, I calculate the number of teachers, doctors, police officers, firefighters, and public administrators per one thousand inhabitants at the county level. Prior research shows that areas with larger immigrant populations tend to provide fewer public goods, both historically and today (Tabellini, 2020; Alesina et al., 2023). This pattern reflects the broader challenge of achieving social cohesion in ethnically diverse societies (Alesina et al., 1999; Alesina & La Ferrara, 2000). Building on my earlier results showing that electrification increased the integration of diverse manufacturing workers, I examine whether it also weakened the negative relationship between immigrant presence and the provision of local public services, particularly in counties with high manufacturing intensity.

To estimate the effect of electrification on the residential segregation of manufacturing workers, I use a difference-in-differences model estimated at the county–decade level, covering all counties in the continental United States in the period 1900-1940. To assess its effect on the provision of local public services, I estimate an OLS regression with the county's immigrant share, a post-electrification indicator, and their interaction as key explanatory variables. Both models include county fixed effects, state-by-decade fixed effects, and interactions of decade indicators with the 1900 county population and manufacturing employment share. These controls account for baseline differences in population size and manufacturing intensity, as well as for time-invariant county characteristics and state-specific time trends.

I find that residential segregation among manufacturing workers significantly declined in electrified counties compared to non-electrified ones. This suggests that electrification was associated with greater integration not only within manufacturing industries but also within the neighborhoods where workers lived. Regarding local public services, I find, at baseline, a negative and statistically significant relationship between the immigrant share and public service employment, consistent with previous evidence. The interaction between immigrant share and electrification is positive and statistically significant, but only in counties with high baseline manufacturing intensity. These findings suggest that electrification partially mitigated the negative association between immigrant presence and local public service provision, consistent with its furnishing greater social cohesion and improving the integration of immigrant manufacturing workers into local communities.

Finally, I examine whether electrification affected the cultural assimilation of immigrant manufacturing workers. A priori, it is unclear how immigrants will adjust their cultural identity as they become more economically and socially integrated. On one hand, immigrants experiencing reduced social and economic distance from natives may be more likely to assimilate culturally (Fouka et al., 2022; Abramitzky et al., 2024). On the other hand, assimilation may be a response to discrimination, so that well-integrated immigrants may feel less pressure to signal assimilation and instead choose to preserve their ethnic identity (Fouka, 2019, 2020). Accordingly, results on assimilation are mixed. I find a modest increase in intermarriage between immigrants and natives in electrified counties, consistent with greater social integration. However, I find no systematic change in the naming patterns of children born to immigrant parents. Overall results suggest that while electrification may have fostered closer social ties, it did not necessarily lead to broader changes in cultural assimilation.

Taken together, my results show that by removing barriers that keep workers segregated on the job, technological change in production can promote integration at work and in the community, strengthening social cohesion within local communities.

Related Literature and Contribution. This paper contributes to several strands of literature. First, it builds on research showing that radical technological change often requires a deep reorganization of production and the adoption of complementary innovations to realize productivity gains. These dynamics have been documented in the diffusion of the steam engine and the rise of the factory system (Sokoloff, 1984; Mokyr, 2001; Juhász et al., 2024), electricity (Du Boff, 1967; Devine Jr, 1983; David, 1990; Atkeson & Kehoe, 2007), and modern information and communication technologies (Brynjolfsson & Hitt, 2000; Syverson, 2011; Brynjolfsson et al., 2025).[4] Building

[4]Radical technological changes that affect a wide range of sectors and industries (e.g., steam, electricity, internal combustion engine, information technologies), and that can transform both household life and firm organization, are

on this established insight, I provide empirical evidence that changes in production processes can also influence how workers are integrated in the workplace, with broader implications for social cohesion.[5]

My paper connects two strands of research: the growing literature on the economic consequences of electrification in the United States and the extensive work on immigrant integration during the Age of Mass Migration. Regarding the first, recent work provides empirical evidence that electrification increased manufacturing productivity, agricultural modernization, and structural transformation (Kitchens & Fishback, 2015; Lewis & Severnini, 2020; Gaggl et al., 2021; Fiszbein et al., 2024).[6] Other scholars emphasize that the production paradigm introduced by electrification, characterized by high division of labor and the rise of assembly-line methods, increased demand for both low-skill and high-skill workers, while reducing employment in semi-skilled occupations (Goldin & Katz, 1998; Gray, 2013; Jayes et al., 2025). I make a direct data contribution to this literature by introducing newly digitized maps that trace the decennial expansion of the U.S. high-voltage grid during the first half of the twentieth century.[7]

Turning to the integration of immigrants in U.S. society, the literature shows that assimilation generally improves labor market outcomes and intergenerational mobility (Abramitzky & Boustan, 2017; Abramitzky et al., 2021, 2024). However, assimilation is a complex process shaped by both immigrants' choices and the surrounding economic and social context. Previous studies highlight the role played by native backlash, often driven by perceived labor market competition or cultural distance, as well as government policies (Goldin, 1994; Fouka, 2019, 2020; Tabellini, 2020; Fouka et al., 2022; Abramitzky et al., 2023; Fouka, 2024; Medici, 2025).

My paper offers a new perspective by examining technological change as a channel that can shape immigrant assimilation and social integration. I show that the early diffusion of electrification created new opportunities to incorporate immigrant workers into manufacturing, with implications that extended beyond the workplace. Electrification not only increased productivity and transformed industrial organization but also influenced how ethnically diverse workers interacted and integrated, both at work and within their communities. These findings highlight that technological change has social as well as economic dimensions, affecting the structure and cohesion of local communities.[8]

referred to as "general-purpose technologies" (David, 1990; Bresnahan & Trajtenberg, 1995; Helpman, 1998; David & Wright, 2003; Jovanovic & Rousseau, 2005; Mokyr, 2005).

[5] A large empirical literature studies the effects of technological change on workers and labor markets, focusing on wage inequality, job polarization, and the skill composition of the labor force (Katz & Murphy, 1992; Goldin & Katz, 1998; Caroli & Van Reenen, 2001; Card & DiNardo, 2002; Autor et al., 2003, 2006; Acemoglu & Autor, 2011; Autor & Dorn, 2013; Acemoglu & Restrepo, 2019, 2020; Jaimovich & Siu, 2020).

[6] Contemporary evidence from developing countries shows that electrification enhances productivity, industrialization, and living standards, but its benefits rely on affordable access, reliable supply, and users' capacity to adopt the technology effectively, choing the adoption challenges faced in the historical U.S. context (Dinkelman, 2011; Lipscomb et al., 2013; Allcott et al., 2016; K. Lee et al., 2020; Kassem, 2024).

[7] Most previous studies on U.S. electrification define local access using the location of hydroelectric plants or a single cross-section of the grid. In Section 3, I provide an overview of previous measures used in the literature.

[8] A large body of qualitative work in sociology and political science examines how major technological changes related to industrialization transform social relations, community structures, and the organization of work (Marx & Engels, 1848; Tönnies, 1887; Durkheim, 1893; Polanyi, 1944; Thompson, 1967; Braverman, 1974; Hounshell, 1984;

This contribution advances the emerging literature that examines how technological change in the way industries operate and produce can shape social dynamics.[9] Acemoglu and Wolitzky (2025) develop a model linking workplace relations and community interactions, showing that technologies which increase workplace monitoring can weaken informal ties and reduce community cooperation, ultimately eroding local social capital. In contrast, my paper provides empirical evidence that production technologies that reduce barriers between diverse workers can promote integration. Most empirical work in this area has focused on gender, showing that electrification and automation increased female labor participation and improved women's economic and social status (Cortés et al., 2024; Feigenbaum & Gross, 2024; Vidart, 2024; Forslund et al., 2025; Vidart, 2025; Ager et al., *forthcoming*). I extend this line of research by examining the effects of technological change on immigrant workers, another historically marginalized group.

Paper Outline. The remainder of the paper is organized as follows. Section 2 provides the historical background. Section 3 describes the data on the expansion of electrification and defines the treatment. The core of the paper is divided into two parts. Section 4 examines the employment consequences of electrification. It presents the main variables and data sources, outlines the empirical strategy, and reports results at the industry-county-decade level. Section 5 explores heterogeneity across industries. Section 6 analyzes the community effects of electrification. It describes the data and variables, details the empirical approach, and reports results at the county-decade level. Section 7 concludes.

2 Historical Background

This section provides historical context on the Age of Mass Migration and the expansion of the U.S. electric grid, with a focus on the electrification of the manufacturing sector. It also examines the role of immigrant workers in manufacturing and presents historical evidence on how the transformation of production processes driven by electrification affected them.

2.1 The Age of Mass Migration

Between 1850 and 1920, about 30 million Europeans migrated to the United States. This influx, which represented an average annual arrival rate of 0.75% of the U.S. population, greatly increased the share of foreign-born residents and reshaped the American economy (Carpenter, 1927; Hutchinson, 1956). In 1910, at the peak of immigration, 14.9 percent of the U.S. population and 20.5 percent of its workforce were foreign-born (Appendix Table A.1).

Mokyr, 1992; Nye, 1992; Inglehart & Baker, 2000; Landes, 2003). Other studies show that technological progress related to ICT and digital innovations continues to reshape work, social relations, and civic engagement (Bell, 1973; Castells, 1996; Putnam, 2000; DiMaggio et al., 2001).

[9] This focus differs from a well-established literature in economics showing that specific communication technologies, such as the television or the radio, can directly affect social capital, political behavior, and intergroup relations, across both developed and developing countries (Gentzkow, 2006; Ferraz & Finan, 2008; Olken, 2009; Enikolopov et al., 2011; DellaVigna et al., 2014; Yanagizawa-Drott, 2014; Adena et al., 2015; Martin & Yurukoglu, 2017; Blouin & Mukand, 2019; Durante et al., 2019; Wang, 2021; Hornuf et al., 2023; Armand et al., 2024; Russo, 2024).

Early migrants came mainly from Northern and Western Europe, including the United Kingdom, Ireland, Germany, and Scandinavia. After the 1890s, falling transatlantic transportation costs and worsening economic conditions in parts of Europe shifted migration toward poorer countries in Southern, Central, and Eastern Europe (Hatton & Williamson, 1998). In 1850, more than 90% of foreign-born residents in the U.S. came from Northern and Western Europe, but by 1910 this share had fallen below 50% as increasing numbers of Italians, Poles, and Eastern European Jews arrived (Abramitzky & Boustan, 2017).[10] These new immigrants differed sharply from earlier arrivals in skills, language, and religion, bringing significant cultural and economic change to American society.

This "Age of Mass Migration", marked by open borders and a steady inflow of immigrants, ended in the early twentieth century, first with the outbreak of World War I and later with the introduction of restrictive immigration laws in the United States. Growing concerns about the cultural distance between newcomers and natives, fears of job competition, and national security motivated this shift in political attitudes toward migration (Goldin, 1994). The Immigration Act of 1917 was the first major federal restriction, introducing a literacy test for immigrants. Its impact was limited, as most European immigrants were literate and enforcement was weak. In 1921, Congress passed the Emergency Quota Act, which capped annual immigration at 3% of each nationality's U.S. population based on the 1910 Census. The Immigration Act of 1924 further reduced the quota to 2% and based it on the 1890 Census, favoring earlier immigrants from Northern and Western Europe (Hutchinson, 1981; Abramitzky et al., 2023). These laws effectively ended the era of open immigration, leading to a long period of restricted inflows that lasted until the Immigration and Nationality Act of 1965.

2.2 The Expansion of the Electric Grid and the Electrification of U.S. Manufacturing

Centralized electricity generation in the United States began in the 1880s with the opening of the Pearl Street Station in Manhattan and the completion of the hydroelectric project at Niagara Falls (Hughes, 1979). Yet by the end of the nineteenth century, electricity had only a limited impact on the U.S. economy. In 1899, electric motors accounted for less than 5% of total mechanical horsepower in manufacturing, with steam power still dominant (Mowery & Rosenberg, 1999, p. 103).[11] However, factories already produced and consumed more than half of all electricity produced in the country (Nye, 1992, p. 186). Electricity soon emerged as the pivotal technology of the Second Industrial Revolution (Jevons, 1931; Mokyr & Strotz, 1998; Rosenberg, 1998). In the early 1900s, many manufacturers replaced waterwheels and steam engines with in-house electric generators,

[10]See Appendix Figure A.1. During the 1850s, a total of 268 thousand immigrants arrived in the United States, with 94.8% coming from Northern and Western Europe. In the 1890s, of the 368 thousand immigrants who arrived, 49.6% were from North-West Europe, while 47.7% originated from Central, Southern and Eastern Europe. The 1900s saw the peak at 817 thousand arrivals, only 22.2% of whom came from the "old" sending countries, with 71.3% now originating from the "new" regions (figures calculated from the Historical Statistics of the United States, Millennial Edition). Appendix Figure A.2 reports maps of the share of immigrants over the total population in U.S. counties in 1900 and 1940, as well as their distribution between "old" and "new" stock.

[11]In the same year, only 3% of all U.S. homes and 8% of urban dwellings used electricity (David, 1990, p. 356).

which they viewed as more reliable and cost-effective than power purchased from central stations (Cohn, 2017). Over time, however, cost comparisons shifted in favor of external supply (Devine Jr, 1983). Industrial electrification gained momentum between 1914 and 1917, as the diffusion of central power plants and the resulting decline in energy prices allowed utilities to produce about half of all U.S. electricity. During this period, central stations also overtook isolated plants in generating capacity (Du Boff, 1967).

The key driver of the growing availability and falling cost of electricity was the expansion of the high-voltage transmission grid –an interconnected network that carried power from generating plants to substations and, ultimately, to consumers, operating in near-perfect synchrony (Cohn, 2017, p. 2).[12] By the 1910s, engineers viewed connecting hydroelectric plants as essential to fully exploit water resources, conserve coal, and reduce urban pollution. They began to envision "a network of high-tension lines connecting together, with efficient distribution, all the waterpowers capable of development" across the United States (Electrical World, 1912, p. 859). The grid expanded gradually through numerous independent public and private initiatives rather than a single coordinated plan. This decentralized process fostered continuous technical innovation aimed at improving reliability and ultimately led to coast-to-coast integration, achieved only by the late 1960s (Cohn, 2017).

These interlinked systems allowed utilities to operate more efficiently and deliver electricity over long distances to a broader range of customers. By the late 1910s, utilities were rapidly expanding their industrial client base, and isolated plants were replaced by the growing centralized network. Coal shortages during World War I further accelerated this shift, and by the end of the war, most manufacturers relied on "rented electric power". By the late 1920s, manufacturing had become the largest electricity-consuming sector in the U.S. economy, accounting for about half of total use. The share of horsepower generated by electric motors in manufacturing rose from 23 percent in 1909 to 77 percent in 1929 (Appendix Table A.3).

The adoption of electricity in manufacturing was driven mainly by the goal of reducing production costs. Although the technology offered major advantages, the shift to electrification was gradual, and its effects on productivity growth remained difficult to detect for several decades (David & Wright, 1999). The slow pace reflected the high economic and organizational costs typical of transitions to new technological paradigms (David, 1990; Atkeson & Kehoe, 2007). Appendix Figure A.4 contrasts the traditional layout of factories powered by steam or water with the new organization made possible by external electrification. Before electrification, factories relied on a central power source –usually a steam engine– that transmitted energy through overhead shafts and belts to individual machines. This system required heavy structures and forced machinery to be placed near the power source to minimize energy loss (Licht, 1995). Electrification reduced transmission costs by eliminating these inefficiencies and allowed each machine to connect directly to its own motor. Machines could be turned off when idle, saving energy and lowering costs, while factories could be built with lighter, more flexible structures. This decentralization of power en-

[12]For a detailed account of the early development of the U.S. electric grid, see Hughes (1983) and Cohn (2017).

abled more efficient factory layouts, with machines arranged flexibly according to production needs rather than mechanical constraints (Du Boff, 1967; Devine Jr, 1983).

2.3 Immigrant Workers in Manufacturing and the Role of Electrification

Between the late nineteenth and early twentieth centuries, immigrants played a central role in the industrialization of the United States. Their willingness to work for lower wages and take on physically demanding jobs enabled firms to expand production and keep costs low, increasing profitability (Carter & Sutch, 1998). Support for immigration restrictions partly stemmed from concerns that immigrant labor depressed wages and displaced native workers, assuming that the two groups were close substitutes (Goldin, 1994; Hatton & Williamson, 1998). However, later research challenged this view, showing that complementarities between less-skilled immigrant workers and more-skilled native workers often increased overall productivity and supported higher wages for natives in complex tasks (Hirschman, 2005; Ottaviano & Peri, 2012). Beyond their role as workers, immigrants also contributed to population growth in rapidly urbanizing areas and stimulated demand for manufactured goods, further fueling industrial expansion (Hirschman & Mogford, 2009).

Between 1880 and 1920, manufacturing employment increased from 14% to approximately 25% of the workforce in the United States, expanding from 2.5 to 10 million workers. Within manufacturing, the industries with the highest growth were metallurgy and machinery (Hounshell, 1984; Hirschman & Mogford, 2009). As highlighted above, the expansion of the electric network, which ensured reliable and affordable power, played a key role in the growth of the manufacturing sector during this period and the following decades (Du Boff, 1967). The rapid growth of manufacturing also relied heavily on immigrant labor (Kuznets, 1971).

Before electrification, factories used a central engine to transmit power through a system of belts and shafts connected to machines (Devine Jr, 1983; Hounshell, 1984). This setup required frequent worker movement and close coordination around shared equipment. Skilled machinists typically stayed at fixed stations to operate machines, while laborers handled raw materials, cleaned, and kept production moving. These tasks were often performed by *teams of laborers* –known as gangs– who loaded, unloaded, and carried materials across the shop floor. Foremen or gang bosses usually oversaw coordination (Graziosi, 1981). These gangs tended to be *ethnically homogeneous*. This could happen when a well-connected worker helped hire acquaintances from the same background, or when timekeepers needed to quickly assemble a crew and turned to familiar contacts, often resulting in a fairly homogeneous crew (Dillingham et al., 1911). Historians note that nationality and language homogeneity was common in factory labor organization in the early twentieth century. For example, Commons (1935, p. xxv) described visiting a factory in 1904 and observing gangs of workers made up entirely of Swedes. When asked why, the agent replied, "It is only for this week. Last week we employed Slovaks. *We change about among different nationalities and languages*". When work groups were more ethnically diverse, language barriers made team coordination and understanding complex instructions especially difficult (Graziosi, 1981). This environment posed particular challenges for immigrant workers, especially recent arrivals with limited industrial experience or English knowledge. Many were confined to manual, low-skilled tasks and were often assigned to different stages of production along ethnic lines (Dillingham et al., 1911).

The spread of electricity as the main source of industrial power led to a gradual but fundamental reorganization of factory production. Independent electric motors replaced centralized drive systems, allowing machines to be arranged according to the flow of manufacturing operations. This change supported the development of fixed workstations, where materials and products moved through sequential stages of production. Continuous-flow processes reduced the need for workers to have technical skills or maintain machines—tasks that had previously been part of the operator's job (Nuwer, 1988). Machine design and labor practices shifted toward narrow, repetitive tasks, increasing specialization and the division of labor (Devine Jr, 1983; Sonenblum, 1990). These changes lowered barriers to entry for unskilled and immigrant workers. As Commons (1904, p. 6) noted, the first goal of this new division of labor was that "*cheaper men –unskilled and immigrant labor– could be utilized in large numbers.*"

The reduced need for coordination on the factory floor following electrification made it easier to employ immigrant labor and may also have supported the integration of workers from diverse cultural backgrounds. The new production system reduced the importance of shared language and cultural alignment by enabling greater task independence. Electrification could allow workers from different ethnic groups to perform their duties side by side with limited direct coordination. This physical and functional proximity could create opportunities for informal social contact and gradual acculturation at work, potentially fostering greater cohesion across ethnic lines. For example, Barrett (1984, p. 44) noted that, by the late 1910s, work groups "were quite *mixed ethnically, racially, and in terms of skills* [...] This allowed for an *informal process of acculturation*". *The same work process* and labor market *which seemed to divide workers* from one another also [...] offered a *basis for unity*" (Barrett, 1984, p. 45). These patterns of integration in the workplace and shared labor experiences may have supported cross-ethnic solidarities and contributed to a common working-class culture (Gutman, 1973).

However, although the new organization of production may have allowed for greater ethnic mixing, the extent of actual integration remains unclear. Other historical sources emphasize persistent ethnic divisions. For instance, despite the adoption of the division of labor, Bodnar (1977, pp. 38–39) emphasized that real earnings "tended to *follow racial and ethnic lines.* Native-born Whites and Western Europeans generally earned more than Croats, Serbs, Bulgarians, and other newcomers from Southern Europe". In many cases, "*ethnic and racial segregation by department minimized inter-ethnic social contact on the job*" (Barrett, 1984, p. 45). These persistent divisions may have reflected employers' belief that homogeneous work groups were still more productive even with the new production process, particularly if ethnic tensions existed or if shared language and cultural background were still viewed as important for efficiency (Williams Jr, 1947; Blalock et al., 1967; Bonacich, 1972; Lang, 1986; Lazear, 1995).

In cases where historical sources point to increased ethnic mixing due to changes in factory organization, they offer insight into how greater workplace contact may have facilitated broader social integration. According to Barrett (1984, p. 43), during the early twentieth century, "*workmen overcame labor market segmentation and significan social and cultural barriers*". Similarly, Brody (1960, p. 96) observed that while "recent [immigrant] arrivals dominated the bottom ranks of the [manufacturing] industry", they gradually improved their status. Over time, they "sensed [...] the *disappearance of clear lines of class and status*" as they "did rise", "*learned to speak English* with fair fluency", and eventually merged with skilled native workers (Brody, 1960, pp. 106–108). Despite demanding work schedules, gangs and work groups still spent part of the day "*standing around talking*", which offered informal opportunities for acculturation (Barrett, 1984, p. 44). Saloons near the plants provided another point of social contact. These spaces were "were quite mixed, tending to *draw workers from particular departments in a plant regardless of skill or ethnicity*" (Barrett, 1984, p. 45). Outside the factory, neighborhoods also created opportunities for interaction. Many manufacturing workers lived in densely populated, ethnically diverse areas near industrial centers (Conzen, 1979). Most blocks –and, importantly, *individual tenement buildings–were ethnically mixed*. This community structure created frequent chances for inter-ethnic contact and integration (Barrett, 1990).

Building on these varied historical narratives, this paper empirically examines whether electrification contributed to the employment and community integration of ethnically diverse manufacturing workers, linking a major technological shift in industrial production to changes in local social cohesion.

3 Data on U.S. Electrification Expansion

In this study, I combine data on the gradual expansion of the high-voltage electric grid with ethnicity-related measures at the industry–county and county levels, constructed from the full-count U.S. population census between 1900 and 1940. This section describes the data sources used to document the spread of electrification across the continental United States during this period.

High-Voltage Electric Grid. To identify the timing at which each county gained access to high-voltage electricity from a central power plant, I rely primarily on a series of maps that trace the expansion of the high-voltage grid across the continental United States. These maps were published by the Edison Electrical Institute (EEI, 1962) as part of a report on the status of U.S. interconnections and the pooling of electric utility systems, based on data collected up to 1959. I digitize the maps reported in the EEI document for the years 1908, 1918, 1928, 1933, 1940, and 1946.[13]

Although all these maps were produced by the same organization, the recorded location of grid lines may contain slight inaccuracies, as is common with historical maps (Rumsey & Williams,

[13]The original maps are shown in Appendix Figure B.1. The same maps also appear in Cohn (2017), which provides a detailed recount of the origin and expansion of the electric grid in the U.S..

2002). To minimize errors related to the precise placement of the electric grid over time, I follow an approach similar to that used by Sequeira et al. (2020) for digitizing railway lines. I begin by georeferencing and digitizing the 1946 electric grid map to create a shapefile representing the grid at its widest territorial extent.[14] I then georeference the map from the preceding period (1940), overlay it on the 1946 shapefile, and manually remove segments that did not exist in 1940.[15] This process is repeated sequentially for every earlier map –1933, 1928, 1918, and 1908. In each case, I remove from the latest shapefile the lines that did not exist in the previous period. In the rare cases where grid lines appear in earlier maps but are missing from the more recent shapefile, I manually add these segments using the georeferenced image of the corresponding historical map.

Electricity Power Plants. As shown in the first map published by the Edison Electrical Institute (1962), displayed in Appendix Figure B.1a, the high-voltage grid was still in its infancy in 1908, with limited transmission lines near the Niagara Falls power plant, along the California mountain range, and in the Carolinas (Cohn, 2017). However, by 1910, numerous power plants were operating independently of manufacturing establishments across the United States. Locations in proximity of power plants could obtain electricity through local transmission lines, even before the development of high-voltage connections (Fiszbein et al., 2024). During this period, hydroelectricity emerged as the dominant form of commercial power production by electric utilities (Severnini, 2023).

To capture the spatial distribution of early electrification, before the expansion of the high-voltage electric grid, I digitize a 1912 map published by the U.S. Department of Commerce (1912, p. 37), which shows the location of hydroelectric central stations reporting waterpower of at least 1,000 horsepower.

To track the subsequent expansion of large-scale power generation along with the high-voltage grid, I also digitize a 1935 map published by the U.S. Federal Power Commission (1935, p. 5), which reports electric generating stations with annual output exceeding 100 million kilowatt hours. This later map only reports plants substantially larger than those in 1912.[16] Therefore, in the 1935

[14]The 1946 map is the most recent and most comprehensive version available in the EEI report that remains comparable to earlier maps. Sequeira et al. (2020) begin their mapping process using a geo-referenced shapefile of the current U.S. railway network from the Department of Transportation. A comparable map of the present high-voltage electric grid is available from the U.S. Energy Information Administration, U.S. Energy Atlas, Electricity Energy Infrastructure and Resources. However, the electric grid undergoes continuous changes due to decommissioning of aging transmission and distribution lines, upgrades or replacements of existing infrastructure, and the construction of new lines to serve emerging energy sources. These occur along with changes in institutional and regulatory frameworks that can affect the evolution of the grid (Elmes, 1996; Borenstein & Bushnell, 2015; Cohn, 2019). As a result, matching the historical high-voltage network depicted in the EEI Report with the current network is extremely challenging. For this reason, to accurately illustrate the evolution of the grid in the first half of the twentieth century, I utilize the latest historical map from the EEI Report that is consistent with earlier maps as the starting point of my mapping process.

[15]As an illustration of the mapping process, Appendix Figure B.2 shows the 1946 and 1940 digitized and georeferenced electric grid network overlaid on the original paper maps from which the data were obtained.

[16]For reference, 1 horsepower (hp) corresponds to 0.7457 kilowatts (kW). The plants reported on the 1912 map had a capacity of at least 1,000 hp, which corresponds to 745.7 kilowatts. Under full utilization, the 1,000 hp station would generate about 6.5 million kilowatt hours (kWh) per year. This means that the minimum size of a power plant reported in the 1935 map is approximately 15 times larger (in annual energy output) than the minimum size plant reported in the 1912 map.

geolocated dataset, I include all plants from 1912, along with the additional larger stations that appeared by that year.[17]

Definition of Treatment: Electrified Counties. I defined a *county* as *treated* once it gets *access to electricity* from central power stations. The analysis proceeds in decennial intervals. For 1910, I use the 1908 electric grid map and the 1912 power plant data; for 1920 and 1930, the 1918 and 1928 electric grid maps, respectively, along with the 1912 power plants; and for 1940, I use the 1912 and 1935 power plant data and the 1940 grid map. Figure 1 illustrates the expansion of electrification between 1910 and 1940.[18]

I define a county c as *electrified* in decade t if at least one of the following conditions is met: (i) the county lies within a 50 km (ca. 30 miles) radius of a central power station, or (ii) the county is intersected by a 5 km (ca. 3 miles) buffer around a high-voltage transmission line. In robustness checks, I vary these parameters to test the sensitivity of the results to the definition of electrification. Specifically, I use alternative radii of 25 km (ca. 15 miles) and 75 km (ca. 45 miles) around power plants –the latter corresponding to the upper bound of electricity transmission reach without high-voltage infrastructure, as identified in previous studies (e.g. Fiszbein et al., 2024; Vidart, 2024). I also test alternative buffer zones around grid lines of 1, 10, and 20 km (ca. 0.5, 5, and 10 miles).

Decennial Expansion of Electrification. Appendix Table B.1 illustrates the evolution of the number of electrified counties between 1910 and 1940, based on my definition of treatment. In 1910, 35.6 percent of U.S. counties were electrified, mostly those located near power plants (Figure 1a). Between 1910 and 1920, the grid expanded slowly, mainly by connecting existing plants, raising the share of electrified counties only slightly to 39.6 percent (Figure 1b). After the 1920s, grid expansion accelerated, extending from the coasts toward the center of the country and increasing network density (Figures 1c–1d). By 1940, 73.6 percent of U.S. counties were electrified. Howevr, regional differences were large. For example, in New England, 95.5 percent of counties were electrified by 1910 due to the abundance of hydroelectric plants, while in the West North Central region, nearly 60 percent of counties remained unelectrified even in 1940. These patterns align with the documented historical development of the U.S. electric grid (Hughes, 1983; Cohn, 2017).

[17] The original maps are shown in Appendix Figure B.3.

[18] Most previous studies on U.S. electrification define local access using the location of hydroelectric plants or a single cross-section of the grid. Lewis and Severnini (2020), Fiszbein et al. (2024), and Clay et al. (2025) use the location of hydroelectric plants. The electrification measure in Gray (2013) is the share of total horsepower in a state–year derived from electricity, reported in the Census of Manufacturers. Kitchens and Fishback (2015) digitize a 1935 map of the U.S. electric transmission grid and generation plants published by the Federal Power Commission (1935), which is comparable to the EEI maps for that year. Gaggl et al. (2021) digitize maps compiled by the U.S. Army Corps of Engineers in the 1960s documenting grid expansion over time. Their approach is the closest to mine, but their maps are not readily available for consultation. Vidart (2024) uses changes in county-level power capacity between 1911 and 1919 based on directories of central stations. While this plant-level data provides rich detail on local capacity (the intensive margin), my approach focuses on the extensive margin, identifying the decade when each county first connects to the central power system. This allows me to cover a longer period and mitigates concerns about differences in the precision or coverage of plant-level data. Our approaches are essentially complementary.

Selection into Electrification. Given the staggered expansion of electrification between 1910 and 1940, I examine which non-geographic county characteristics were associated with earlier adoption. I compare counties that became electrified in a given decade with those that became electrified in the following decade. For each comparison, counties electrified in decade t are assigned a treatment value of one, and those electrified in decade $t+1$ are assigned zero. I repeat this process for all consecutive decade pairs (1910–1920, 1920–1930, 1930–1940, and 1940–non-electrified) and stack the resulting datasets. I then regress county characteristics on the treatment indicator, including state-by-year and decade-pair fixed effects, with standard errors clustered at the county level. The analysis focuses on non-geographic characteristics such as population size, urbanization, literacy, ethnic composition, diversity, and the share of employment in manufacturing.

Panel A of Table 1 reports the results for key county characteristics. Counties electrified earlier were more populated, urban, and literate, with higher shares of foreign-born residents, greater ethnic diversity, and larger manufacturing employment shares. These patterns are consistent with historical accounts of grid expansion (Cohn, 2017). However, my identification strategy focuses on differences in *growth rates rather than levels* between treated and control units, as the difference-in-differences design relies on parallel pre-treatment trends rather than identical initial conditions. Panel B of Table 1 reports the results for the *growth rates* of county characteristics, measured as percentage changes between the decades before and after electrification. For most variables, there is no statistically significant difference in growth rates between counties electrified earlier and those electrified later. The only exceptions are population and manufacturing employment growth, both lower in early-electrified counties. This indicates that counties electrified earlier were larger and more industrialized but grew more slowly than those electrified later. In the empirical analysis, I account for these differences and perform robustness checks to confirm that the results are not driven by these county-level patterns.

Having defined treatment based on the expansion of electrification, I study its effects on the employment and community integration of ethnically diverse manufacturing workers in the United States between 1900 and 1940.

4 Employment Consequences of Electrification

I start by assessing the employment consequences of electrification. In this section, the unit of analysis is a manufacturing industry q located in county c during decade t (1900-1940). To address changes in county boundaries over time, I harmonize geographic units using 1900 county borders throughout the analysis. In this section, I first describe the data sources and define the key outcome variables. Second, I specify the empirical strategy and the estimation procedure. Finally, I present the main results, followed by a summary of the results related to robustness tests and additional analysis. The data, empirics, and results on the community consequences of electrification are provided in Section 6.

4.1 Data and Definitions

In this section, I describe the sources of the industry–county–decade data, define the key variables used in the empirical analysis, and present summary statistics. Supplementary variables employed in robustness checks and additional analyses are introduced in the relevant sections of the paper.

Definition of "Ethnicity". In this project, I define *ethnicity* as a classification that combines race for individuals born in the United States and country of birth for foreign-born individuals. Specifically, I consider the following ethnic groups: White U.S. natives, Black U.S. natives, and foreign-born individuals, each assigned to a group based on their country of birth (Alesina et al., 2016). That is, among the foreign-born, each country of birth defines a distinct ethnic group. Unless otherwise specified, I define *immigrants* as foreign-born individuals (i.e., first-generation), and treat all American-born individuals, regardless of parental origin, as *natives*. The dataset includes 42 ethnic groups: two based on the race of U.S. natives and 40 based on the countries of birth of first-generation immigrants.

My analysis covers the period 1900–1940, a time of substantial geopolitical change in Europe, with shifting national borders due to the two World Wars. To ensure consistency over time, I apply standard consolidations to the country-of-birth information reported in the census data. These procedures are detailed in the Data Appendix (Section B.2).

U.S. Individual Full-Count Census. To construct measures at the industry–county-decade level, I use the decennial, full-count individual U.S. Census data from 1900 to 1940, available through the Integrated Public Use Microdata Series (IPUMS) (Ruggles et al., 2024). Each census observation corresponds to an individual and includes information on demographic characteristics, place of birth (U.S. state for native-born, country for foreign-born), race, and county of residence c. As specified above, I combine information on place of birth and race to assign individuals to their corresponding *ethnic group* k.

For individuals in the labor force, the data report both the *industry* q in which they are employed and their *occupation* j. These two concepts are distinct: industry refers to the sector or type of economic activity, while occupation captures the specific job, profession, or task performed by the worker.[19] To ensure a clearer conceptual separation between industry and occupation and to align more closely with standardized classifications, I consolidate certain categories within each variable. For industry, I compare the IPUMS variable IND1950 with the Standard Industrial Classification (SIC) Manual developed by the U.S. Department of Labor. I construct a crosswalk between the

[19]Despite this conceptual distinction, both variables are challenging to compare over time, as the Census Bureau's classification schemes changed significantly across census years, often tailored to the needs of contemporaneous researchers. To address these inconsistencies, IPUMS researchers have harmonized the original census responses, reclassifying them into the 1950 industrial and occupational coding schemes, recorded as the variables IND1950 and OCC1950. Details on the harmonization process are available at the IPUMS USA website. In line with the original Census Bureau approach, the harmonized variables maintain a relationship between occupation and industry: the same occupation may be classified differently depending on the industry in which it occurs. For example, OCC1950 distinguishes between "attendants, physician's and dentist's office" (302) and "attendants, hospital and other institution" (730), despite the similarity of tasks and skill requirements.

two systems, with the aim of preserving the structure of the IPUMS classification while enhancing alignment with the SIC codes. This results in 135 industry categories, including 49 within the manufacturing sector. For occupation, I compare the IPUMS variable OCC1950 with the 1958 International Standard Classification of Occupations (ISCO-58) developed by the International Labor Organization. I construct a crosswalk that retains the level of detail in OCC1950 while minimizing dependence on industry-specific context, consistent with the ISCO framework. The main objective of this reclassification is to reduce noise arising from variation in the level of detail in self-reported census responses, since the original occupational variables are derived from open-ended entries.[20] Additionally, the reclassification helps mitigate the artificial multiplication of occupations that can result from technological change and more detailed division of labor over time (Kastis & Vipont, 2025).[21] The final classification includes 108 occupation categories. Some are specific to particular manufacturing industries (e.g., spinner in textiles, pressman in printing), while others are common across sectors (e.g., bookkeeper, engineer). More details on industry and occupation reclassification are provided in the Data Appendix (Section B.2). In the analysis, I only focus on industries in the *manufacturing sector*.

Using the reclassified categories for ethnicity and employment, I construct labor force measures at the level of industry q in county c and decade t. To ensure comparability over time, I fix geographic units to 1900 county borders, the first period covered in the analysis. To obtain consistent measures of the total number of workers and of workers by ethnic group in each industry–county–decade cell, I follow the procedure in Ferrara et al. (2024). First, I calculate the relevant totals using the decennial census and county–state identifiers. Second, I link each decennial county to its corresponding 1900 county using the crosswalks provided by the authors. In some cases, this is a one-to-one match; in others, counties have been split or merged, resulting in multiple linkages across time. Third, I apply the M4 weights from the crosswalk to the decennial values. These population-based weights adjust for urban–rural composition and topographic suitability.[22] Finally, I aggregate the weighted values by the 1900 county–state identifiers to obtain time-consistent measures.

Using these consistent data over time, I compute three main outcomes at the industry–county-decade level, each capturing different aspects of the ethnic composition of the labor force in U.S. manufacturing. I consider 49 manufacturing industries in total, located in 2,843 counties over

[20] For example, ISCO-58 has a unique category for "engineers", but there are several categories for it in OCC1950. In the reclassification, I aggregate "engineers" in one category to avoid picking up the fact that one individual would report only "engineer" as a response for occupation while another would provide a more detailed description and be classified in a more specific category such as "engineer, industrial" or "engineer, metallurgical".

[21] For example, in the historical census (OCC1950 classification), "typesetters" (512) and "electrotypers and stereotypers" (520) are listed as separate occupations, but I merge them into a single category, "typers", to avoid double-counting occupations whose distinction reflects technological changes in the printing process, and not necessarily differences in task content. For reference, Autor and Dorn (2013) classify the two occupations more generally as "machine operators". The minor label of the ISCO-58 classification is an aggregated category for "press workers", which includes those and others like "bookbinders". I keep distinct occupations like "typers" and "bookbinders" separate to mirror the level of detail in the original IPUMS classification.

[22] To obtain the M4 population-based weights, the authors divide the county area into urban and rural areas, after excluding non-inhabitable areas, with additional weighting for topographic suitability (i.e., elevation). The construction procedure is detailed in Ferrara et al. (2024). See also Posch et al. (2024), who use the same procedure and M4 weights.

5 decades. Throughout this section, I use the term *ethnic composition* to describe the relative representation of different ethnic groups in the labor force, measured by their respective shares within a manufacturing industry in a county and decade. I use the term *ethnic integration* to describe the overall ethnic heterogeneity of the labor force (diversity) and the degree of mixing among workers of different ethnic backgrounds (segregation) within the manufacturing industry.

Share of Workers by Ethnicity. For each manufacturing industry q in county c in decade t, I calculate the share of the labor force that belongs to broad "ethnic categories", that is, White U.S.-born, Black U.S.-born, and foreign-born (regardless of country of birth).

$$Workers\ Share\ By\ Ethnic\ Category_{qct} = \frac{Num.\ Workers\ Ethnic\ Category_{qct}}{Tot.\ Num.\ Workers_{qct}} \tag{1}$$

This variable measures the relative importance of each broad ethnic category in the labor force of the manufacturing industry. The definition follows several papers that focus on the role of immigrants during the Age of Mass Migration in the U.S., whose main measure of immigrant presence at the county level is the fraction of immigrants over the population (e.g., Sequeira et al., 2020; Tabellini, 2020; Medici, 2025). I use these three measures of relative representation of ethnic groups, broadly defined, to assess how electrification affects the *overall ethnic composition* of the labor force in manufacturing industries.

Ethnic Diversity of Workers. To capture the *overall ethnic heterogeneity* of the labor force in a manufacturing industry q in county c and decade t, I calculate an index of ethnic diversity of workers, following the definition in Alesina et al. (2016):

$$Ethnic\ Diversity_{qct} = 1 - \sum_{k=1}^{K} \pi_{kqct}^2 \tag{2}$$

where $k = 1, \ldots, K$ indicate the ethnic groups of the workers in the industry-county-decade cell, and π_{kqct} is the fraction of workers that belong to ethnic group k over the total labor force in the industry-county-decade cell. For the computation of this index, I consider 42 ethnic groups: White U.S.-natives, Black U.S.-natives, and immigrants assigned to a group by country of birth.

This measure follows the standard definition of the fractionalization index, which captures the probability that two randomly selected individuals belong to different ethnic groups (Alesina et al., 2003). The index ranges from zero, indicating a completely homogeneous industry with no ethnic diversity, to one, indicating maximum diversity, where each worker belongs to a different ethnic group.

Ethnic Segregation of Workers Across Occupations Within Industry. The diversity index captures the overall ethnic heterogeneity of the labor force in a manufacturing industry, but does not indicate how integrated workers are within the industry. For example, two industries may have the same overall diversity, yet differ in their internal composition: in one, workers from different ethnic groups are mixed across occupations, resulting in high diversity at all levels; in the other, workers are segregated into occupations by ethnic group, producing low diversity within occupations despite the same overall industry diversity.

To capture the level of *segregation of workers along ethnic groups within a manufacturing industry*, I adapt the segregation index from Alesina and Zhuravskaya (2011) and calculate the ethnic segregation of workers across occupations within a manufacturing industry q in county c and decade t as follows:[23]

$$Ethnic\ Segregation_{qct} = \frac{1}{K-1} \sum_{k=1}^{K} \sum_{j=1}^{J} \pi_{jqct} \frac{(\pi_{kjqct} - \pi_{kqct})^2}{\pi_{kqct}} \tag{3}$$

where $k = 1, \ldots, K$ indicate the ethnic groups of the workers and $j = 1, \ldots, J$ indicate the occupations within the industry; π_{jqct} is the fraction of workers employed in occupation j in the industry-county-decade cell; π_{kjqct} is the fraction of workers of ethnic group k employed in occupation j over the total number of workers in occupation j within the industry-county-decade cell; and π_{kqct} is the fraction of workers who belong to ethnic group k over the total labor force in the industry-county-decade cell. As before, for the computation of this index, I consider 42 ethnic groups: White U.S.-natives, Black U.S.-natives, and immigrants assigned to a group by country of birth.

The segregation index ranges from zero to one. A value of zero indicates that there is no ethnic segregation within the industry, which means that each occupation is as ethnically diverse as the industry overall. A value of one indicates a completely segregated industry along ethnic lines, where each ethnic group is confined to a separate occupation and each occupation is entirely homogeneous. A decline in the index indicates that occupations within the industry are becoming more ethnically diverse and more similar to the overall ethnic heterogeneity of the industry. In contrast, an increase indicates that occupations are becoming more segregated by ethnicity.

Summary Statistics. Panel A of Appendix Table 2 reports the summary statistics for the main variables at the industry-county-decade level. On average, each manufacturing industry within a county in a decade has 88 workers and represents 0.4% of the county labor force. There is variation in the dimension and relative importance of manufacturing industries in employment within the county, but the dimensions remain relatively small throughout the sample, with 95% of manufacturing industries having less than 250 workers and representing less than 1.4% of county employment. Across the period 1900-1940, on average, immigrant workers are 10.4% of the labor force in a manufacturing industry, while White and Black U.S. natives comprise 81.2% and 8.4%, respectively. In 1900, at the initial measurement point, on average, 19.26% of workers in manufacturing industries were foreign-born. By 1940, this share fell to 4.58%. This decline matches the historical pattern of the end of the Age of Mass Migration. In the late nineteenth and early twentieth centuries, the United States received a large number of immigrants thanks to unrestricted immigration policies.

[23] In their paper, Alesina and Zhuravskaya (2011) calculate the country-level spatial segregation of ethnic groups across sub-national regions within the country. I apply their procedure, considering a manufacturing industry (in a county and decade) as the unit of analysis and the occupations within the industry as the sub-units.

Therefore, at the beginning of the twentieth century, the U.S. had a vast stock of foreign-born workers. During the 1920s, the introduction of strict immigration quotas led to a substantial decline in new arrivals. With fewer new immigrants coming to the U.S., the share of foreign-born workers in manufacturing steadily decreased.

Between 1900 and 1940, on average, manufacturing industries have an ethnic diversity index of 0.147 and an ethnic segregation index of 0.202. Importantly, the two indices are weakly correlated, with a correlation coefficient of -0.088 (Appendix Table B.2). This suggests that the diversity and segregation index captures different aspects related to the distribution of ethnic groups within the industry. In 1900, the average ethnic diversity of workers in manufacturing industries was 0.210. The general trend of ethnic diversity in the manufacturing workforce was declining in this period. By 1940, the average diversity index was 0.099, less than half the baseline value. In 1900, the average ethnic segregation of workers across occupations within manufacturing industries was 0.111. Unlike the previous measure, the ethnic segregation of workers in manufacturing industries was increasing in this historical period, and by 1940, the average index was 0.255, more than double the baseline value.

4.2 Empirical Strategy

My goal is to assess the *causal effect* of electrification on the ethnic composition and ethnic integration of the labor force within manufacturing industries. The unit of analysis is a manufacturing industry q in a county c and decade t. I take advantage of the staggered expansion of the electric grid over time across the continental United States to implement a difference-in-differences (DID) estimation strategy.

As detailed in Section 3, combining information on the early expansion of the high-voltage electric grid and the location of early major electric power plants, I define the decade in which a U.S. county gets access to reliable and affordable electricity from the network. This defines *treatment* at the county level. Once a county has been electrified, it remains treated for the following periods, as the electrical transmission infrastructure is in place.

Every manufacturing industry located in an electrified county is considered as *treated*. Historical sources suggest that the development of electric utilities in charge of providing energy from power stations was "particularly stimulating to electrification of manufacturing plants" and "by 1920 manufacturing had become the largest electricity-using sector in the economy" (Du Boff, 1967, p. 510). Detailed data on electricity use by county and industry are not readily available for this period. However, the 1930 Census of Manufacturers provides data on energy consumption for the manufacturing industries located in the 57 largest cities in the United States. It reports the total horsepower by industry and divides it between prime movers –which comprise factory sources of mechanical power that convert energy from sources like coal (steam engines) or water (water wheels) into rotational motion– and electric motors. In large U.S. cities, by 1930, on average, 86.4% of horsepower in manufacturing industries came from electric motors, and all manufacturing industries partially used electricity to power their operations, ranging from 38.9% in "pulp, paper, and paperboard mills" to 100% in industries related to the production of "watches, clocks, and clockwork-operated devices".[24]

Econometric Framework. I estimate the causal effect of electrification on the ethnic composition of the labor force in manufacturing industries by estimating the following *event-study* model. This approach compares treated industries –those in counties that have been electrified– to control industries in counties that are not yet electrified. The regression model is specified as follows:

$$Y_{qct} = \alpha_{qc} + \phi_{s(c)t} + \delta_{qt} + \sum_{h=-4}^{3} \beta_h \, \mathbb{I}\{K_{ct} = h\} + \mathbf{X}'_{ct} \, \Gamma + \varepsilon_{qct} \tag{4}$$

In the model, Y_{qct} represents the outcome of interest for manufacturing industry q located in county c in decade t. Specifically, Y_{qct} can be the fraction of White U.S.-born, Black U.S.-born, or foreign-born workers in the industry, as well as the index of ethnic diversity of the industry's workforce and the index of ethnic segregation of workers across occupations within the industry. Detailed definitions of these outcome variables are provided in Section 4.1.

To ensure that baseline differences in county size do not drive the comparison between treated and control industries, the control vector $\mathbf{X}_{ct}$ includes the logarithm of the county population in 1900, interacted with indicator variables for each decade. Furthermore, the event-study model incorporates a comprehensive set of fixed effects to control for confounding factors and capture unobserved heterogeneity. Industry-by-county fixed effects, α_{qc}, absorb all time-invariant characteristics specific to each industry in each county. They control for factors like geography or natural resource endowments that could influence the presence or growth of certain industries in particular locations. State-by-year fixed effects, $\phi_{s(c)t}$, capture any shocks or trends that are common to all counties within a state in a given decade. By including them, the model accounts for state-level conditions (for example, state-specific booms or recessions) that might affect industrial outcomes in that decade. Industry-by-year fixed effects, δ_{qt}, control for national trends that affect specific industries over time, accounting for the fact that different manufacturing industries may experience distinct national patterns or respond differently to economy-wide events. The estimation is weighted by the share of county employment in industry q, reflecting the relative size of the industry within the county economy.

Including this set of fixed effects makes the empirical specification comparable to a two-way fixed-effects (TWFE) model, where the unit of observation is the industry-county pair and the temporal trends are allowed to vary by both state and industry. This framework allows for comparison of the evolution of the ethnic composition of the labor force within the same manufacturing industry across counties. It contrasts the industry located in electrified counties with the industry in counties not yet electrified, while controlling for a wide range of potential unobserved confounding factors.

To implement the event-study, I define event-time indicator variables that track the time rel-

[24] Energy consumption data at the industry-city level are from Lafortune et al. (2019), which provides the digitized data from the 1930 U.S. Census of Manufacturers and links the industries reported in the census of manufacturers to the industries reported in the individual census (variable IND1950).

ative to the electrification of each county. The sample period covers decades five decades, $t = 1900, \ldots, 1940$, and $\mathit{Electrified}_{ct}$ denotes the decade when county c first became electrified. I define the event time as $K_{ct} = t - \mathit{Electrified}_{ct}$; this measures the number of decades since electrification for county c at time t. For each integer value h, I include an indicator variable $\mathbb{I}\{K_{ct} = h\}$ in the regression. $\mathbb{I}\{K_{ct} = 0\}$ indicates the period in which the county becomes electrified (the first treatment period). Negative values of h correspond to pre-treatment periods (decades before the county was electrified), whereas positive values of h correspond to post-treatment periods (decades after electrification). The analysis covers four pre-treatment periods and four post-treatment periods. Standard errors, ε_{qct}, are clustered at the county level –the level of treatment– to account for serial autocorrelation within industries and potential correlation across industries in the same geographical area. I classify counties with population levels above the 99th percentile in 1900 as always treated and exclude them from the analysis. These counties include the largest U.S. cities, which were likely already served by local urban power plants prior to the earliest period for which the electric grid can be observed in my data.

The primary parameters of interest are the coefficients of the event-time indicators, denoted by β_h. Under the standard parallel trends assumption (which posits that, in the absence of treatment, treated and control groups would have followed similar outcome trajectories), each coefficient β_h can be interpreted as the *causal effect* of electrification on the outcome Y_{qct} at event time h (average treatment effect on the treated, ATT). In other words, β_h measures how much electrification changed the outcome h decades after electrification (for $h \geq 0$). It also indicates whether there are differences in trends h decades before electrification (for $h < 0$), which serves as a check on the validity of the parallel trends assumption.

I complement the analysis by estimating the following *difference-in-differences* (DID) model. This model assesses the change in the ethnic composition of the labor force in manufacturing industries before and after electrification, comparing industries located in counties that have already been electrified with those located in counties that have not yet been electrified. The regression model is specified as follows:

$$Y_{qct} = \alpha_{qc} + \phi_{s(c)t} + \delta_{qt} + \beta_t \, \mathbb{I}\{t \geq \mathit{Electrified}_{ct}\} + \mathbf{X}'_{ct}\,\Gamma + \varepsilon_{qct} \tag{5}$$

The outcome variables, set of controls, fixed effects, and other model specifications are identical to those described for Equation 4. The only difference in this DID setup is the regressor of interest. In the DID model, $\mathbb{I}\{t \geq \mathit{Electrified}_{ct}\}$ is an indicator variable that takes the value of one for each decade t in or after the decade when county c becomes electrified for the first time (denoted by $\mathit{Electrified}_{ct}$). This indicator takes the value 0 for all the decades before electrification. Under the standard parallel trends assumption, the coefficient β_t represents the average treatment effect on the treated (ATT). In other words, it measures the overall causal effect of electrification on the ethnic composition of the labor force in manufacturing industries located in treated counties (electrified) relative to control counties (not-yet electrified).

Issues with Staggered Treatment. The conventional practice of implementing event studies via two-way fixed-effect ordinary least squares (TWFE OLS) regressions has recently come under scrutiny in settings with staggered difference-in-differences designs, where treatment occurs for multiple units at different times. In these scenarios, units treated earlier inadvertently serve as controls for units treated later. This overlap can bias estimates when treatment effects differ across groups or over time. The TWFE estimator effectively computes a weighted average of group-specific treatment effects, and some of these weights can be negative (Borusyak & Jaravel, 2018; de Chaisemartin & d'Haultfoeuille, 2020; Goodman-Bacon, 2021). In practice, this means that the overall estimate of the TWFE regression can be underweight or even subtracted from the true impact of certain treatments. For example, it might produce a near-zero or negative overall coefficient even if all groups actually experience positive effects. Furthermore, when leads and lags of the treatment are included (as in an event-study specification), the TWFE coefficients at each event time can be contaminated by effects from other periods. This contamination can create misleading patterns in the estimated dynamics. For instance, apparent pre-treatment trends or spurious long-term effects may arise solely due to the heterogeneity of the treatment effect rather than any actual violation of the parallel trends assumption (Sun & Abraham, 2021). Together, these issues undermine the reliability of the conventional TWFE approach in staggered adoption designs, motivating the development of new estimation methods that remain valid under staggered treatment and treatment heterogeneity.[25]

Estimation Procedure. To address these issues, I implement the estimation method of Borusyak, Jaravel, and Spiess (BJS, 2024). This approach introduces a simple *imputation procedure* that solves the problems of staggered treatment designs and provides a robust and efficient estimator. The key idea is to reconstruct the counterfactual outcome for each treated unit –what its outcome would have been had it not yet been treated– and then compare it with the observed outcome to identify treatment effects. By construction, this imputation estimator isolates the causal effect of treatment in a transparent way: it contrasts observed outcomes of treated units with model-predicted counterfactuals derived from untreated units, using the information from all available control observations to inform those counterfactual predictions. This approach yields event-study estimates that are free of the contamination and weighting biases inherent in the standard TWFE estimator, since no treated unit is inappropriately used as a control for another once its own counterfactual is explicitly accounted for.

In practice, the estimation procedure comprises three steps. First, the fixed effects and the controls are fitted by running a regression using only the sub-sample of untreated observations at a given time. This step captures the normal evolution of the outcome in the absence of treatment, accounting for unit differences and common time trends. Second, the estimated fixed effects and

[25] A large recent literature covers the issues related to the implementation of TWFE estimation in staggered treatment adoption settings and proposes ways to address these problems (e.g., de Chaisemartin & d'Haultfoeuille, 2020; Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021; Sun & Abraham, 2021; Roth, 2022; Roth et al., 2023; Borusyak et al., 2024; de Chaisemartin & d'Haultfoeuille, 2024).

controls are used to impute the expected outcome for each treated observation as if that unit had remained untreated at that time (counterfactual). In other words, for every treated unit and period, this step predicts what the outcome would have been without treatment, using the controls and fixed effects estimated from the untreated data. Finally, the third step is to compute the treatment effect for each treated observation by taking the difference between its actual observed outcome and the imputed no-treatment outcome. These individual-level effects are then aggregated (taking an appropriate weighted average) to obtain the overall estimate of the average treatment effect on the treated (ATT).

The imputation approach proposed by Borusyak et al. (2024) provides an estimator that remains unbiased under arbitrary treatment-effect heterogeneity, addressing the core problem of TWFE with staggered treatment. It also offers both conceptual and practical advantages compared to other recently proposed methods to address the issues related to staggered treatment (Callaway & Sant'Anna, 2021; Sun & Abraham, 2021; de Chaisemartin & d'Haultfoeuille, 2024). The BJS estimator achieves efficiency by leveraging all available untreated observations across all periods to estimate counterfactual outcomes, rather than relying on only a single baseline period or a limited control group for each cohort. Using the full set of untreated observations, this approach extracts more information and thus achieves lower variance in the estimated effects.[26] This leads to higher precision without sacrificing validity. The imputation method does not require stronger identifying assumptions than the alternative methods.[27] Furthermore, the BJS imputation approach ensures transparency. It directly links the estimator with the counterfactual outcomes computed from the data, making it clear how the estimates are obtained and where the identifying assumptions enter the analysis. In contrast, some alternative procedures involve multiple steps of differencing or balancing across groups and time periods, which can make the estimates harder to interpret.

I implement the BJS imputation procedure to estimates the event-study model from Equation 4 and the DID model from Equation 5.

Identification. The BJS imputation method assumes the standard DID conditions for the identification of causal effects: *parallel trends* in the absence of treatment and *no anticipation* of treatment. The imputation procedure described above explicitly relies on parallel trends and no anticipation to justify those imputations. A key benefit of this approach is that it provides a straightforward

[26] The key difference between the estimator proposed by Callaway and Sant'Anna (2021) (CS) and the estimator proposed by Borusyak et al. (2024) (BJS) relates to how the two approaches use pre-treatment periods. CS makes all the comparisons relative to the *last pre-treatment period*, whereas BJS makes comparisons relative to the *average of the pre-treatment periods*. Averaging over multiple pre-treatment periods can increase precision; however, the BJS approach imposes parallel trends for all groups and time periods. The comparison between the two approaches is discussed in detail in Roth et al. (2023).

[27] Borusyak et al. (2024) (BJS) observe that the assumptions in Callaway and Sant'Anna (2021) (CS) may appear weaker than the assumptions in their imputation-based approach. In principle, the CS approach requires the outcome trends to be parallel between the treated unit and some reference group only since the period before treatment (and not necessarily in earlier periods). However, if all not-yet treated cohorts are used as the reference group, one implicitly assumes that the untreated potential outcomes of those cohorts also follow parallel trends over time (Marcus & Sant'Anna, 2021). In practice, then, the requirement of parallel trends is effectively equivalent across the BJS and CS methods in this case.

procedure for testing the validity of these identifying assumptions.

The approach provides a clear separation between estimation and testing of identifying assumptions.[28] This is achieved by estimating a separate OLS regression on untreated observations only. This regression augments the specification with fixed effects and controls by adding a series of event-time lead indicator variables for $h = 1, \dots, k$ periods before the treatment onset.[29] These indicator variables capture any systematic differences in outcomes between treated and control observations prior to treatment. The k-th period before treatment serves as the reference group, where the difference between treated and control units is set to zero.[30]

A cluster-robust Wald test is then conducted to check if these pre-treatment lead coefficients are jointly zero. That is, this is a test of the null hypothesis $\beta_h = 0$ for all $h = -k, \dots, -1$ from Equation 4. Not rejecting this null hypothesis implies non-significant pre-treatment coefficients, which supports the validity of the parallel trends and no-anticipation assumptions. The rejection of the null hypothesis is an indication that the treated units had already diverged or responded before treatment. This implies that the parallel trends or no anticipation assumption does not hold in the model.

In the rest of this section, I present the empirical results on the employment consequences of electrification, estimated applying the empirical strategy discussed above.

4.3 Electrification and the Ethnic Composition of Manufacturing Labor Force

My first set of results assesses the consequences of electrification on the *ethnic composition* of the labor force in manufacturing industries. The DID results estimated from Equation 5 are reported in Table 3. Electrified manufacturing industries have, on average, a 1.61 percentage-point *higher* share of **immigrant workers** than non-electrified industries in the post-treatment period. This increase corresponds to 8.57% of the baseline mean in 1900.[31] The ethnic category whose relative representation decline after electrification is **White U.S.-native workers**. Their share falls by 1.61 percentage points in manufacturing industries located in electrified counties relative to non-electrified ones, a decline of 2.2% compared with the 1900 baseline mean. The effect of electrification on the share of **Black U.S.-native workers** is statistically indistinguishable from zero. For all three outcomes, the pre-treatment test shows that the pre-treatment coefficients are jointly indistinguishable from zero, supporting the validity of the parallel-trends and no-anticipation

[28] Borusyak et al. (2024) discuss how this approach is advantageous compared to conventional OLS regressions with leads and lags of treatment and to placebo-based tests. In particular, it avoids the contamination of the tests by treatment-effect heterogeneity, as well as inference problems after pre-treatment testing (details on these issues are provided by Sun and Abraham (2021) and Roth (2022), respectively).

[29] Borusyak et al. (2024) note that the optimal choice of k is a difficult problem. In my setting, the panel is relatively short, and I rely on all the available data for estimation. Therefore, I set $k = -4$, which corresponds to the earliest pre-treatment period observed.

[30] The usual approach in event-study analysis is to set the period before treatment as the reference group. Instead, the BJS procedure sets the k-th period before treatment as the reference group. Due to this definition, the standard errors are expected to be the largest for the last period before treatment ($h = -1$); this is opposite to some other conventional tests for parallel trends.

[31] This corresponds to 14.58% of the mean immigrant share in non-electrified industries over the entire period.

assumptions. The event-study plots are displayed in Appendix Figure C.3.

The non-significant effect of electrification on the share of Black native workers in manufacturing holds when the outcome is defined as the share of Black workers among all U.S.-born workers, rather than among the total labor force (Appendix Table C.3). Interestingly, I also find no statistically significant effect on the ratio of occupational income scores across ethnic groups within manufacturing industries (Appendix Table C.4).[32] Taken together, these results suggest that electrification *causes* a change in the composition of the labor force in manufacturing, increasing the relative representation of immigrant workers while decreasing that of White U.S.-native workers. However, on average, electrification did not translate into changes in the relative economic standing of ethnic groups within the manufacturing industries.

Given the effect of electrification on the share of foreign-born workers within manufacturing industries, I then assess which types of immigrant groups are mostly affected by the technological shock. The results of this analysis are reported in Appendix Table C.5, and the corresponding event study graphs are in Appendix Figure C.4. Electrification does not have a statistically significant effect on the share of immigrant workers from Northern and Western Europe (the "old stock"), but it has a positive and statistically significant effect on the share of immigrants from Southern and Eastern Europe (the "new stock"). It is worth noting that in 1900, foreign-born workers from Northern and Western Europe accounted for 14.0% of the manufacturing labor force, whereas those from Southern and Eastern Europe represented only 1.7%.[33] The estimated DID coefficient for the share of Southern and Eastern European immigrant workers is 1.1 percentage points, corresponding to a 64.7% post-electrification increase relative to the baseline mean.[34] Similar patterns emerge when examining the effects of electrification on immigrant workers from English-speaking versus non-English-speaking European countries, and from Protestant-majority versus non-Protestant European countries. After electrification, the share of immigrants from non-English-speaking European countries in manufacturing increases by 9.3% relative to the baseline mean of 11.8%. The share of immigrant workers from non-Protestant European countries rises by 37.5% relative to the baseline mean of 3.2%. Overall, these results suggest that electrification increases the share of

[32] Wages are recorded in the census only starting in 1940. I proxy for the average income of workers by ethnic category (immigrants, native Whites, and native Blacks) using their average occupational score within the manufacturing industry in a county and decade. This is a standard approach in the literature (e.g., Abramitzky et al., 2012, 2014). Occupational scores are defined by occupation and assign each worker the median income of their job category in 1950, serving as a proxy for lifetime earnings. Therefore, the results indicate that electrification did not change the relative average income scores of occupations held by different ethnic groups. However, this measure does not capture wage differences between workers of different ethnic backgrounds within the same occupation.

[33] On average, in 1900, 3.2% of the county population was foreign-born from Northern and Western Europe. This means that "old stock" immigrants were represented in manufacturing industries at a rate 4.4 times higher than in the overall county population. In the same year, 0.6% of the population was foreign-born from Southern and Eastern Europe, implying that "new stock" immigrants were represented in manufacturing industries at a rate 2.8 times higher than in the total county population.

[34] By 1940, foreign-born workers from Northern and Western Europe made up 2.0% of the manufacturing labor force and 0.8% of the county population, indicating a representation rate 2.4 times higher in manufacturing than in the overall population. Immigrants from Southern and Eastern Europe accounted for 1.4% of the manufacturing labor force and 0.6% of the county population, also corresponding to a representation rate 2.4 times higher in manufacturing than in the total county population.

immigrant workers who are *more distant* from the native majority in terms of culture, language, and religion. These groups faced the greatest language and cultural barriers and were therefore most likely to benefit from the reduced need for coordination and communication in manufacturing made possible by electrification.

4.4 Electrification and the Ethnic Integration of Manufacturing Labor Force

Having established that electrification affected the relative share of immigrant workers in manufacturing industries, I examine its impact on the *ethnic integration* of the labor force within the industry. As described in Section 4.1, I analyze two related dimensions. First, I study the ethnic diversity of workers, which captures the overall heterogeneity of the labor force. Second, I assess ethnic segregation across occupations, which reflects the extent to which workers of different ethnic backgrounds are mixed within the industry. The event-study estimates from Equation 4 are shown in Figure 2, and the difference-in-differences results from Equation 5 are reported in Table 4.[35]

Figure 2a presents the event-study estimates of the *causal effect* of electrification on the **ethnic diversity** of workers in manufacturing industries.[36] The estimates show no statistically significant differences between treated and control industries before electrification. Electrification leads to an *increase* in ethnic diversity in treated industries relative to untreated ones. This positive effect is statistically significant in all post-treatment periods.

Appendix Figure C.5a shows the average trend in the ethnic diversity index for treated and control industries constructed from the raw data. Overall, ethnic diversity declines during this period, and both groups display a similar downward trend. This pattern reflects the final phase of the Age of Mass Migration in the United States, when immigrant inflows fell sharply while the native-born population continued to expand. As a result, the relative share of foreign-born individuals declined, leading to a general reduction in ethnic diversity. However, in industries exposed to electrification, the decline is less steep. The estimated positive effect thus reflects a slower reduction in diversity after electrification, implying that the gap in ethnic diversity between treated and control industries widens following the shock.

The corresponding DID estimates, shown in Column 1 of Table 4, indicate that electrified industries have an ethnic diversity index that is 3.4 points higher than non-electrified industries. This difference equals 16.83% of the baseline mean in 1900 of 20.1 points.[37] The pre-treatment test shows that the pre-treatment coefficients are jointly indistinguishable from zero ($F = 0.24$, $p = 0.87$), supporting the validity of the parallel-trends and no-anticipation assumptions.

Figure 2b presents the event-study estimates of the *causal effect* of electrification on the **ethnic**

[35]Appendix Figure C.5 shows the average trends of the main outcomes constructed from the raw data. These trends help illustrate the trajectory of each variable and provide context for interpreting the event-study results. Unlike the event-study estimates, these trends are not adjusted for unobserved factors, which are accounted for in the regression analysis through fixed effects and covariates.

[36]The estimated coefficients are reported in Appendix Table C.6.

[37]This effect also represents 25.63% of the mean ethnic diversity index in non-electrified industries over the study period.

segregation of workers across occupations within manufacturing industries.[38] Electrification *reduces* ethnic segregation in treated industries relative to control industries, and this negative effect is statistically significant in all post-treatment periods.

Appendix Figure C.5b shows the average trends in the ethnic segregation index for treated and control manufacturing industries constructed from the raw data. In contrast to ethnic diversity, this period is characterized by a general rise in ethnic segregation. Workers became more separated across occupations along ethnic lines, leading to more homogeneous occupational groups within manufacturing industries. This pattern is a novel finding made possible by my data, which uses individual-level census records to reconstruct key ethnic characteristics of the manufacturing labor force during this historical period. Both electrified and non-electrified industries follow this upward trend, but the increase is slower in electrified industries. As a result, the gap in segregation between treated and control industries widens after electrification.

Column 3 of Table 4 reports the corresponding DID estimate. Electrified industries have an ethnic segregation index 3.5 points lower than control industries, a reduction equal to 30.48% of the baseline mean in 1900 of 11.4 index-points.[39] The pre-treatment test shows that the pre-treatment coefficients are jointly statistically indistinguishable from zero ($F = 1.30$, $p = 0.27$), supporting the validity of the parallel trends and no-anticipation identifying assumptions.[40]

The models in Columns 1 and 3 of Table 4 represent my preferred specification. Since White U.S.-born workers are the largest group, their relative share could influence both indices. Columns 2 and 4 confirm that electrification increases ethnic diversity and reduces ethnic segregation, even after controlling for the share of White U.S.-born workers within the industry. When this control is added, the magnitude of the estimated coefficient for ethnic diversity decreases to 10.4% of the 1900 baseline mean, while the coefficient for ethnic segregation remains unchanged. Results in Column 5 show that the negative and statistically significant effect of electrification on ethnic segregation persists even after accounting for the overall level of ethnic diversity within the industry.[41] The magnitude of the estimated coefficient for ethnic segregation decreases and corresponds to 22.8% of the 1900 baseline mean. Column 6 shows similar results when both the share of White U.S.-born workers and ethnic diversity within the manufacturing industry are included as controls.

Overall, the industry-county-decade results show that electrification had a significant impact on the manufacturing workforce. It increased the overall ethnic heterogeneity of workers within the industry and increased the average ethnic diversity within occupations, leading to a lower

[38] The estimated coefficients are reported in Appendix Table C.6.

[39] This effect also corresponds to -18.89% of the mean ethnic segregation index in non-electrified industries over the study period.

[40] For the ethnic segregation index, the estimated coefficient in the period immediately before treatment is positive and marginally significant ($p = 0.07$). Although the pre-treatment coefficients are jointly indistinguishable from zero, this marginally significant estimate at event time $h = -1$ (the decade before electrification) may suggest some anticipation effects in industries about to be electrified. If present, these effects would bias the results in the opposite direction of the treatment effect, since β_{-1} is positive while all post-treatment coefficients are negative. However, to ensure that the main estimates are not affected by this issue, I re-estimate the model treating $h = -1$ as the treatment period. The results, shown in Appendix Table C.14, confirm the robustness of the main findings.

[41] In their analysis, Alesina and Zhuravskaya (2011) include both diversity (or fractionalization) and segregation indices in the same regression, so that segregation captures the degree of mixing conditional on overall heterogeneity.

segregation of workers along ethnic lines. These findings are consistent with the historical narrative discussed in Section 2.3. In this historical context of the United States, electrification created the opportunity to reorganize the manufacturing production process around unit electric drives and fixed workstations, introducing a finer division of labor. This reorganization reduced the barriers to the employment of immigrant workers. As tasks became simpler, more repetitive, and more autonomous, linguistic and cultural differences became less relevant. This transformation could allow for a more ethnically heterogeneous workforce, as foreign-born workers who were culturally and linguistically distant from the native majority could more easily work alongside others. When this shift spread across occupations, it likely reduced segregation and promoted greater ethnic integration within manufacturing.

4.5 Discussion of Validity and Robustness Checks

I perform several checks to verify the validity of the estimated causal effects of electrification on the ethnic composition and integration of the manufacturing labor force. My identification strategy exploits the staggered expansion of the electric grid across the United States between 1900 and 1940. The standard identifying assumptions of the difference-in-differences (DID) framework apply: parallel trends between treated and control observations in the pre-treatment period and no anticipation of treatment. As discussed in Section 4.2, an advantage of using the imputation method by Borusyak et al. (2024) is that it allows a direct test of these assumptions. All result tables report the outcome of this test, and in the event-study figures, the validity of the assumptions can be visually assessed from the pre-treatment coefficients. Nevertheless, some concerns about the validity of the identification strategy may remain. In this section, I discuss potential threats to identification and summarize the results of a series of tests that confirm the robustness of the main findings.

Location and Timing of Treatment. The DID framework does not require random treatment assignment for valid causal estimation. It only requires that, conditional on covariates, treated and control units follow similar growth trajectories before treatment. However, it is important to note that the location of hydroelectric plants and the expansion of the electric grid were not random. In the early twentieth century, most electricity in the United States came from hydroelectric power or coal combustion (Cohn, 2017; Severnini, 2023). The inclusion of *Industry* $\times$ *County* fixed effects controls for geographic conditions and natural endowments that could favor electricity generation, such as proximity to rivers suitable for hydroelectric power or to coal deposits (Lipscomb et al., 2013; Gaggl et al., 2021). It also accounts for other county characteristics (such as access to transportation networks during the nineteenth century, historical industrial specialization, and proximity to major markets) that could influence the location and growth of specific industries. I also show that the results remain robust when excluding New York and California, the states that first developed high-voltage transmission networks thanks to their ideal topography. The findings are likewise robust when excluding the New England states,[42] which had a high density of hydroelectric plants by the 1910s and had long been the manufacturing core of the nation (Appendix Tables C.7–C.9).

Starting in the 1910s, discussions on developing an interconnected national electricity network intensified. The process involved multiple public and private actors and evolved gradually through many independent projects rather than a single coordinated effort. As noted by Cohn (2017, p. ix), the U.S. electric grid "grew out of the aggregation of hundreds of different projects initiated by an equal number of different entities, over decades". The inclusion of *State* $\times$ *Decade* fixed effects captures this regional heterogeneity in grid development and the varying roles of local institutions and actors over time. Moreover, the fact that an infrastructure project like the grid takes a long time to build introduces a degree of randomness to its decade-by-decade evolution. For example, a project designed to connect a coal mining area with a major urban center could span several decades, during which intermediate counties gain access to electricity as the grid expands.[43]

The expansion of the grid accelerated after World War I, as manufacturing industries faced coal shortages and utilities reduced electricity costs. During the war, the need to increase power production and transmission for defense industries further stimulated electrification. Similar patterns of network expansion to serve strategic industries continued in later decades (Cohn, 2017). The inclusion of *Industry* $\times$ *Decade* fixed effects controls for nationwide industry-level trends and for differences in the growth rates of manufacturing industries with distinct characteristics. It also accounts for potential differences in electricity adoption by industry.

Until the 1930s, the federal regulation of electric utilities was minimal, and thousands of companies operated independent local power systems. The Great Depression marked a turning point: the federal government expanded regulation and invested heavily in energy infrastructure, financing new dams, power plants, and transmission lines, and promoting rural electrification (Kline & Moretti, 2014; Kitchens & Fishback, 2015; Cohn, 2017; Lewis & Severnini, 2020). The inclusion of *State* $\times$ *Decade* fixed effects also captures these temporal trends by allowing for differential trajectories across states. Given the massive expansion of the grid under New Deal programs, I further show that the results remain robust when excluding counties electrified between 1930 and 1940 (Appendix Table C.10). The inclusion of *State* $\times$ *Decade* fixed effects also accounts for other shocks that could confound the results, such as industrial mobilization during major conflicts and changes in immigration policy. To address the latter, given the potential impact of the 1920s immigration quotas on the composition of the labor force, I show that the results remain robust when excluding counties electrified after 1920 (Appendix Table C.10).

County Manufacturing Intensity. Historical evidence indicates that the expansion of the high-voltage grid was partly driven by the goal of delivering more power to industrial centers. This was especially true after World War I, when "postwar proposals for large-scale interconnected systems specifically focused on delivering electricity to the growing industrial markets for power" (Cohn,

[42] Maine, Vermont, New Hampshire, Massachusetts, Connecticut, Rhode Island.

[43] This idea of "time to build" goes back to Kydland and Prescott (1982) and relates to recent approaches that use construction delays to identify the causal effects of infrastructure projects (Borusyak & Hull, 2023).

2017, p. 38). A possible concern, therefore, is that counties with a stronger manufacturing base before treatment were more likely to be electrified earlier. The inclusion of fixed effects helps mitigate this issue, but I also examine it directly.[44]

I test whether electrification affected the share of manufacturing employment at the county level. The DID results, reported in Appendix Table C.1, reject the joint null that pre-treatment coefficients are indistinguishable from zero, suggesting that the parallel-trends assumption may not hold.[45] The event-study plot in Appendix Figure C.1a shows a positive and statistically significant coefficient only in the decade before electrification, consistent with historical accounts of grid expansion.

To account for this, I re-estimate the model treating the *decade before electrification as the treatment period*. Both the event-study and DID results confirm no pre-treatment effects under this adjustment. The post-treatment effect of electrification on the manufacturing share remains positive and statistically significant, leading to an increase of 1.1 percentage points, or 17.7% of the 1900 baseline mean of 6.2%. This finding supports the view that electrification acted as a positive shock to the manufacturing sector (Gaggl et al., 2021; Fiszbein et al., 2024).

To ensure that the industry-county-decade results are not driven by anticipation effects, I conduct several robustness checks. I include additional controls for the baseline manufacturing share in 1900, interacted with decade dummies, and for the time-varying county manufacturing share. Results are robust to both specifications. They also hold when excluding counties with extreme manufacturing employment shares at baseline or those that remained entirely rural throughout the study period (Appendix Tables C.11–C.13). Finally, while the estimated effects vary slightly, the main conclusions on the ethnic integration in manufacturing industries remain the same when using the decade before actual electrification as the treatment period (Appendix Table C.14).

Ethnic Composition of County Population. A potential concern for the validity of the main results is that electrification may be correlated with changes in the overall ethnic composition of the county population. First, counties with a higher share of foreign-born residents might have been electrified earlier or later. Second, electrification could have affected migration patterns by attracting more foreign-born individuals.[46]

Appendix Figure C.2a shows the effect of electrification on the share of foreign-born individuals in the county population. The estimates reveal no statistically significant differences between treated and control counties before electrification, supporting the validity of the parallel-trends and no-

[44] To address this concern, I estimate a DID model at the *county-decade* level, using the county's decade of electrification as the treatment. The model includes county fixed effects, state-decade fixed effects, the log of the 1900 county population interacted with decade dummies as controls, and standard errors clustered at the county level. The event-study estimates use the same specification.

[45] The issue persists when adding the 1900 share of manufacturing employment interacted with decade dummies as controls.

[46] To address these concerns, I estimate a DID model at the *county-decade* level, using the county's decade of electrification as the treatment. The model includes county fixed effects, state-decade fixed effects, the log of the 1900 county population interacted with decade dummies as controls, and standard errors clustered at the county level. The event-study estimates use the same specification.

anticipation assumptions. This indicates no systematic relationship between the pre-treatment share of immigrants and the timing of electrification. Moreover, electrification does not affect the overall share of immigrants in the county population after treatment. Similar results hold for the share of Black U.S.-born residents (Appendix Figure C.2b). Appendix Figure C.2c shows a slightly negative pre-treatment relationship between electrification and the county-level ethnic diversity index, but no post-treatment effects. Appendix Table C.2 confirms that electrification had no lasting impact on overall ethnic diversity at the county level, nor did it affect the county shares of immigrants or native Black residents. Taken together, these findings reduce concerns about reverse causality between pre-existing county demographics and electrification, as well as about electrification altering the ethnic composition of the local population.

To further rule out confounding from the ethnic characteristics of the county population, I re-estimate the main models including baseline county variables in 1900, interacted with decade dummies, and time-varying county covariates. Results are robust to both checks (Appendix Tables C.15–C.17). Regarding the possibility that the results are driven by the internal mobility of immigrants in response to electrification, it is important to note that the baseline analysis excludes the largest U.S. cities, where immigrants were more likely to settle during this period due to larger ethnic communities and better labor market opportunities (Hutchinson, 1956). To reinforce this point, I further exclude counties with a population size in 1900 above stricter thresholds, which remove additional large and medium-sized urban centers. The results remain unchanged (Appendix Table C.18). Finally, the findings also remain robust when excluding Southern states,[47] where the share of immigrants was historically much lower and the share of native Black residents much higher than elsewhere in the country (Appendix Tables C.7–C.9).[48]

Figure 3 reports a summary of the main robustness checks for the results of ethnic diversity and ehnitic segregation of workers within manufacturing industries.

5 Heterogeneity

As discussed in Section 2.3, historical sources point to mixed evidence. In some cases, electrification supported the integration of ethnically diverse workers in manufacturing; in others, ethnic divisions persisted within industries. The main result in Section 4.4 supports the former view: on average, electrification significantly reduced ethnic segregation across occupations within manufacturing. In this section, I examine whether this effect varied across industries.

Figure 4 presents the estimated effects of electrification on ethnic segregation by industry, aggregated at the 2-digit SIC level.[49] Most coefficients are negative and statistically significant, though

[47] Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.

[48] See maps in Appendix Figure A.2.

[49] These estimates are based on the difference-in-differences model in equation 5, using the imputation-based method proposed by Borusyak et al. (2024). For the heterogeneity analysis, the first step remains the same: fixed effects and controls are estimated using only untreated units. These fitted values are then used to impute a counter-

the size of the effect varies across industries. The most substantial declines in segregation occur in electrical equipment and chemical products, where electrification reduced ethnic segregation across occupations by 40.8 percent and 45.3 percent of the 1900 baseline mean, respectively (coefficients of –0.108 and –0.107). In contrast, the industrial equipment sector shows a positive and significant effect: electrification increased ethnic segregation by 52.1 percent relative to the 1900 baseline mean (coefficient of 0.023). In this case, electrification appears to have reinforced, rather than reduced, occupational segregation among ethnically diverse workers. The effect for fabricated metal products is close to the overall average, with a 28.7 percent relative to the baseline mean (coefficient of –0.038). Given the variation in electrification's effect on ethnic segregation across industries, I next examine which industry characteristics account for this heterogeneity.

Energy Use Intensity. In highly energy-intensive industries, the search for fuel was not only about finding cheap energy, but also about securing power that was reliable and available in large quantities. For these industries, electricity offered an ideal power source. As electricity costs declined with the growth of utilities and the expansion of the grid, adoption accelerated. Typical examples include metallurgy (or "primary metals") and chemicals (Rosenberg, 1998). With widespread electrification, these industries also experienced major changes in production: unit motors powered individual machines, equipment was reorganized along the production flow, and the division of labor made workers more independent in performing their tasks (Du Boff, 1967, 1979). This shift in factory organization is key to understanding how electrification could promote greater integration of ethnically diverse workers. I therefore use energy intensity as the main industry characteristic to explain differences in the effect of electrification on workplace segregation.

Due to limited historical data on electricity adoption by industry over time, I construct several measures of energy use intensity across different periods. I start with industries' *baseline energy intensity* in 1900. This is based on data from the 1900 Census of Manufacturers, which covers establishments in major urban centers.[50] I define baseline energy intensity as the ratio of fuel and energy expenses to the total wage bill –i.e., the number of dollars spent on fuel and energy per dollar spent on wages.[51] Next, I construct a *short-term energy intensity* measure for 1930, roughly two decades after the initial expansion of the electric grid. This is based on the 1930 Census of Manufacturers, which also covers major cities. I define this measure as the ratio of total horsepower to the number of wage workers –i.e., the amount of energy used per worker. [52] Finally, I measure *long-term energy intensity* in 1963, once the electric grid covered nearly the entire country and electricity had become the dominant power source for both industry and households (Cohn, 2017;

factual outcome for each treated observation. The treatment effect is calculated as the difference between the actual outcome and the imputed counterfactual. In the heterogeneity analysis, these effects are aggregated by the relevant category to compute the heterogeneous average treatment effect on the treated.

[50]Energy consumption data at the industry-city level come from Lafortune et al. (2019), which digitizes the 1900 and 1930 U.S. Census of Manufacturers and links reported industries to individual census categories (IND1950).

[51]Fuel and energy expenses are the only energy-related variables available in the 1900 census.

[52]The 1930 census also reports fuel and energy expenses. When I use the fuel-to-wage ratio instead, results are very similar to those using the 1900 measure. The correlation between fuel and energy expenses in 1900 and 1930 is 0.90.

Lewis & Severnini, 2020). I use data from the 1963 U.S. Input-Output Table, which provides the first detailed breakdown of electricity use by industry.[53] I calculate the share of electricity used by each 2-digit SIC manufacturing industry relative to the total used by the entire manufacturing sector. This offers a direct measure of electricity intensity under full grid coverage. For each of these three periods, I compute the average energy intensity for each 2-digit SIC industry. I then classify industries as *high* energy-intensive if their value is above the median and *low* if it is below. The correlation matrix of these measures of energy use intensity is reported in Appendix Table B.4.

Columns 1–6 of Table 5 report the results. The effect of electrification on ethnic segregation is negative and statistically significant across all subgroups. Regardless of the measure used, electrification consistently reduced ethnic segregation more in high energy-intensive industries. In industries with higher expenditure in fuel and energy at baseline, before the expension of electrification, the new technology reduced ethnic segregation by 39.1 percent of the 1900 baseline mean. In contrast, the reduction was 22.2 percent in industries with lower energy expenses. Similarly, in industries with a high horsepower-to-worker ratio in 1930, segregation declined by 31.6 percent of the baseline mean, compared to 26.5 percent in less energy-intensive industries. Industries that became the most electricity-intensive in the long run also show the strongest effects of early electrification on reducing ethnic segregation between 1900 and 1940. In these sectors, segregation declined by 36.6 percent of the 1900 baseline mean, compared to 23.3 percent in industries that ultimately used less electricity.

Taken together, these findings suggest that electrification reduced ethnic segregation most in industries where energy use was highest. These industries likely adopted electricity earlier and more intensively due to its advantages in reliability, availability, and affordability. As a result, they were more likely to reorganize production around individual workstations and reduce the need for coordination on the factory floor, key conditions that enabled greater mixing of ethnically diverse workers.

Scope for Reorganization. To further test whether the decline in ethnic segregation within manufacturing industries was linked to the reorganization of production enabled by electrification, I examine variation by average establishment size in 1900. The key idea is that industries with larger establishments before electrification had greater scope to reorganize production once they adopted electricity as their main power source (Rosenberg, 1998). With a larger workforce, these industries could reassign tasks and mix ethnically diverse workers more extensively once production shifted to fixed workstations and reduced coordination needs on the factory floor. In contrast, industries with smaller establishments faced more limited opportunities to reorganize labor or increase ethnic mixing, even if they adopted electricity, simply because fewer workers were involved.

I compute average establishment size using data from the 1900 Census of Manufactures, defined as the total number of workers (wage earners and officers) divided by the number of establish-

[53]The 1963 U.S. Input-Output Table is available from the Bureau of Economic Analysis of the U.S. Department of Commerce. The detailed information is available in the "367-industry level intra-industry transactions matrix with producing industries, consuming industries, and total requirements coefficients" file.

ments. I calculate this at the 2-digit SIC-code level and classify industries as having a *high* or *low* establishment size based on whether they are above or below the median.[54] Columns 7–8 of Table 5 show that electrification reduced ethnic segregation significantly more in industries with larger establishments at baseline. In these industries, segregation declined by 34.4 percent of the 1900 mean, compared to just 14.4 percent in industries with smaller establishments. These findings offer additional support for the idea that opportunities to integrate ethnically diverse workers depended on the scale of operations and the potential to reorganize production after electrification.

Having established the effects of electrification on the integration of ethnically diverse workers in manufacturing industries, the next section examines whether these effects extended to communities, with broader implications for social cohesion beyond the workplace.

6 Community Consequences of Electrification

In Section 4, I examine how electrification affected employment in manufacturing industries through changes in the ethnic composition and integration of the labor force. I show that industries in electrified counties employed a larger share of immigrant workers and had a more ethnically diverse labor force than those in non-electrified counties. Electrification also reduced ethnic segregation across occupations within manufacturing, meaning that workers of different ethnic backgrounds were more likely to hold similar jobs. This pattern aligns with historical evidence showing that electrification allowed factories to reorganize production, increasing the division of labor and reducing coordination needs among workers. These changes could potentially have lowered barriers for employers to hire a more diverse workforce and to integrate workers of different ethnic backgrounds within the same occupations.

Having established these results at the industry–county level, I now examine whether employment integration had broader social consequences at the community level. The unit of observation is a county c in decade t (1900-1940). As before, county borders are harmonized using the 1900 borders. In this section, I first describe the data sources and define the main outcome variables. Next, I outline the empirical strategy and estimation approach. Finally, I present the main results and summarize the robustness checks.

6.1 Data and Definitions

In this section, I describe the sources of the industry–county–decade data, define the key variables used in the empirical analysis, and present summary statistics. I construct the community-level outcomes from the decennial, full-count individual U.S. Census data from 1900 to 1940, available through IPUMS (Ruggles et al., 2024). Supplementary variables used in robustness checks are introduced in the relevant sections of the paper.

Residential Segregation. A key factor for cohesive local communities is the spatial mixing of

[54] Appendix Table B.4 reports the correlation between baseline establishment size and energy use intensity.

individuals from different backgrounds, which prevents the formation of segregated neighborhoods (Cutler & Glaeser, 1997). Therefore, residential segregation is a straightforward measure of integration outside the workplace. Residential patterns also provide a valuable measure of the social integration of immigrants. Historically and today, immigrants from the same ethnic group tend to live close to one another.[55] When immigrants live in more ethnically mixed neighborhoods, it indicates greater assimilation into the host community, while persistent clustering in ethnic enclaves reflects slower integration and stronger ethnic boundaries (Waldinger, 1987; Borjas, 2000; Waters & Jiménez, 2005; Damm, 2009; Abramitzky et al., 2024).

I measure residential segregation at the county level using full-count U.S. Census data from 1900 to 1940, covering 2,843 counties over five decades. I adapt the segregation index from Alesina and Zhuravskaya (2011), following the same logic used previously to measure ethnic segregation across occupations within manufacturing industries (see Section 4.1). Each household is linked to a county (c) and an enumeration district (e), which represents a small geographic area covered by one census enumerator and serves as a proxy for a *neighborhood*.[56] By leveraging enumeration districts as the geographic units, I can capture fine-grained residential patterns across both urban and rural parts of each county.[57]

To build this measure, I restrict the sample to working-age individuals (15 years and older) who have a recorded occupation in a manufacturing industry, including both men and women. I do not make a distinction between manufacturing industries in this case. In other words, I focus on the *residential segregation of the county manufacturing labor force*. This ensures that I am comparing the locations of economically active people employed in manufacturing industries over time, which is consistent with the earlier employment analysis. I consider the same 42 *ethnic groups* as defined before: White U.S.-born, Black U.S.-born, and foreign-born assigned to a group based on their country of birth. I calculate the residential segregation within a county c in decade t as follows:

$$\textit{Residential Segregation}_{ct} = \frac{1}{K-1} \sum_{k=1}^{K} \sum_{e=1}^{E} \pi_{ect} \frac{(\pi_{kect} - \pi_{kct})^2}{\pi_{kct}} \tag{6}$$

[55] These areas, known as ethnic enclaves, are neighborhoods with a high concentration of co-ethnics or countrymen. Common examples include Chinatown and Little Italy.

[56] An enumeration district was the area that a single census enumerator could cover within two weeks in urban areas or four weeks in rural areas. Between 1880 and 1940, U.S. counties were subdivided into many such districts; for example, even medium-sized counties often contained dozens of enumeration districts, and large cities had hundreds.

[57] Earlier studies of historical segregation often relied on city wards, which are available for a longer period in the census but vary greatly in size and are not available for rural areas. Logan and Parman (2017) introduce a page-based approach to measure residential segregation between Blacks and Whites from 1880 to 1940. This method exploits the fact that census enumerators recorded households in the order they visited them door to door, numbering them sequentially on the census sheets. As a result, households listed consecutively on the page serve as good proxies for next-door neighbors. Using this approach, Eriksson and Ward (2019) measure segregation between native- and foreign-born households from 1850 to 1940, focusing on whether foreign-born households had at least one native-born neighbor. Although informative, the page-based method is cumbersome to implement and best suited for comparisons between two groups (e.g., black-white or foreign-native). In contrast, I adopt the segregation index of Alesina and Zhuravskaya (2011), which allows me to account for 42 ethnic groups. This index is straightforward to implement and interpret, and it provides a measure that is conceptually consistent with the index of ethnic segregation across occupations calculated for manufacturing industries (see Section 4.1).

where $k = 1, \ldots, K$ indicate the ethnic groups and $e = 1, \ldots, E$ indicate the enumeration districts within a county-decade cell; π_{ect} is the fraction of the county's manufacturing workers living in district e; π_{kect} is the fraction of manufacturing workers in district e who belong to ethnic group k; and π_{kct} is the fraction of the total county's population working in manufacturing that belongs to ethnic group k. Intuitively, this index captures how unevenly each ethnic group is distributed throughout the county districts, compared to the overall share of that group in the county.

The resulting segregation index ranges from zero to one. An index value of 0 indicates that there is no residential segregation by ethnicity within the county, and each enumeration district is as ethnically diverse as the county as a whole. In this case, manufacturing workers by ethnic groups are evenly distributed across all neighborhoods, without any concentration of groups in a particular space. An index value of 1, in contrast, denotes complete segregation, meaning each enumeration district is inhabited by manufacturing workers belonging to only a single ethnic group and there is no ethnic mixing at the local level. A decline in the segregation index over time means that workers of different ethnic origins have become more spatially mixed within the county, while an increase would imply growing ethnic clustering in specific neighborhoods.

Local Public Services. Since my ultimate goal is to evaluate how technological change in production affected social cohesion, I focus on the provision of local public services, which is a central aspect of community integration (Schiefer & van der Noll, 2017). For the historical period under study, detailed data on local public finances are not available at the county level, nor are there records on the number of public facilities such as schools and hospitals.[58] However, I can proxy for the level of local provision of public services with the *number of workers employed in occupations linked to provision of public service*, which I can recover from the census.[59] In particular, I consider the number of individuals employed as teachers, doctors, police officers, firefighters, and public administration officials and administrators.[60] These occupations are directly linked to the categories of local expenditure associated with public services considered in Alesina et al. (1999) (health and hospitals, education, police protection, fire protection, and general expenditures for welfare).[61] I calculate the number of workers in these occupations relative to the county population (per 1,000

[58]Tabellini (2020) uses data on public spending and city finances from the 1910-1930 Financial Statistics of Cities, which cover only cities with more than 30,000 inhabitants.

[59]I leverage the relationship between employment in public service occupations and the level of public expenditure on those services, as government programs –such as law and order, public administration, education, and health and social care– tend to be highly labor-intensive and rely more on human labor than on capital (Ross, 1985). This approach has been applied in previous work. For example, Fulford et al. (2022) measure local public goods provision before 1960 using the share of the county population employed in teaching and policing, and after 1960 using local expenditure data in those categories. Similarly, Putnam (1995) and Posch and Raz (2025) measure civic engagement at the county level using the number of residents employed in industries related to civic organizations, public administration, and recreation, based on census data.

[60]The occupation categories correspond to the following IND1950 codes: Teachers, 93; Doctors, 75; Police officers, 771 (*marshals and constables*), 773 (*policemen and detectives*), 782 (*sheriffs and bailiffs*); Firefighters, 762 (*firemen*); Public administration officials and administrators (not elsewhere classified), 250.

[61]The authors use data on the share of local government expenditure devoted to public services (health, education, police, fire protection, and welfare) and public goods (roads, highways, sewerage, and sanitation) for the 1990s. These data, from the County and City Data Books, are available starting in 1944.

inhabitants), as follows:

$$\textit{Public Service Occupation (PSO)}_{ct} = 1000 \cdot \frac{\textit{Num. Workers in PSO}_{ct}}{\textit{Tot. Population}_{ct}} \tag{7}$$

Cultural Assimilation of Immigrants. The last dimension I examine regarding the community consequences of electrification concerns the cultural assimilation choices of immigrant manufacturing workers. I focus on two key aspects of immigrants' cultural assimilation. The first is the **foreign-native intermarriage rate**, one of the strongest indicators of cultural assimilation of immigrants, as it weakens the transmission of foreign ethnic traits across generations (Blau et al., 1984; Pagnini & Morgan, 1990). Because marriage is a long-term and intimate union, intermarriage reflects both the willingness of immigrants to integrate and the openness of natives to accept culturally different partners (Gordon, 1964; Alba & Golden, 1986; Wildsmith et al., 2003). Therefore, intermarriage signals mutual acceptance between groups and the broader social legitimacy of cross-ethnic unions.

To measure the intermarriage rate, I use census data and identify married couples in which *at least one spouse is employed in a manufacturing industry.* I further restrict the sample to couples where the younger spouse is 25 years old or younger at the time of enumeration. Because the census does not report the year of marriage, this restriction helps ensure that the observed *marriage occurred within the past decade.*[62] I define intermarriage as a *union between one foreign-born and one U.S.-born spouse.* The intermarriage rate in county c and decade t is calculated as the share of such marriages over the total number of marriages that include at least one U.S.-born spouse.

$$\textit{Intermarriage Rate}_{ct} = \frac{\textit{Num. couples with one foreign-born and one US-born}_{ct}}{\textit{Tot. num. couples with at least one US-born}_{ct}} \tag{8}$$

The second measure of cultural assimilation concerns the **foreignness of names given to children** by foreign-born parents. A large literature in sociology and economics uses children's names to infer cultural traits and identity (e.g., Lieberson & Bell, 1992; Lieberson, 2000; Bazzi et al., 2020). For immigrants, the first name chosen for a child reflects cultural identity and the degree of assimilation into the host society (Abramitzky et al., 2020, and references therein).

[62] I focus on intermarriage among households in which at least one spouse works in manufacturing. While marriage is a one-time decision, labor force participation and occupational affiliation can change more frequently. To reduce discrepancies between the timing of marriage and employment, I restrict the sample to recently formed marriages. Since the census does not record the year of marriage, I limit the analysis to couples in which one spouse is young enough that the marriage likely occurred between the current and previous census. I set the age cutoff at 25. Under American common law, the marriageable age was 12 for females and 14 for males, a standard that continued into the early 1900s (Hamilton, 2012). A 1919 nationwide digest by the Russell Sage Foundation reported that 17 states had no statutory minimum; in those states, the common-law ages applied. Where statutory minimums existed, they ranged from 15 to 18 for males (with 18 the most common) and from 12 to 18 for females (with 15 or 16 the most common) (Hall & Brooke, 1919). Between 1919 and 1940, two states introduced statutory thresholds: Tennessee set a minimum of 16 in 1937, and Maryland established 18 for males and 16 for females in 1939 (Dale et al., 1946). The 25-year threshold also aligns with the average age at marriage during 1900–1940, which was 26.3 years for men and 22.9 years for women (Haines, 1996).

Naming is a deliberate cultural choice and, as such, it "provides a window into parental visions of the ethnic identity of their children" (Sue & Telles, 2007, p. 1383). Selecting a distinctively foreign name signals a stronger attachment of the parents to their ethnic heritage, while choosing a common American name reflects a shift toward the host culture. Because naming is a voluntary and meaningful choice, it offers a powerful indicator of assimilation.[63]

To measure the degree of foreignness of names by county and decade, I first assign a foreignness value to each name recorded in a census year. I collect the universe of first names from each full-count U.S. census and apply the cleaning and standardization algorithm developed by Abramitzky et al. (2012, 2014, 2019).[64] Then, for each first name recorded in the decennial censuses from 1900 to 1940, I calculate the foreignness index separately for males and females as follows:[65]

$$Foreignness\ Index(Name) = \frac{\frac{Num.\ Foreign(Name)}{Tot.\ Foreign}}{\frac{Num.\ Foreign(Name)}{Tot.\ Foreign} + \frac{Num.\ Native(Name)}{Tot.\ Native}} \tag{9}$$

where *Num. Foreign*$(Name)$ and *Num. Native*$(Name)$ are the number of foreign-born and U.S.-native individuals with a given name in the census-decade, respectively; *Tot. Foreign* and *Tot. Native* are the total number of foreign- and U.S.-born individuals in the census-decade, respectively. The resulting *name's foreignness index* ranges from 0 to 1.[66] A higher value indicates a more distinctively foreign name. If all individuals with a given name are native-born, the index equals 0; if all are foreign-born, it equals 1. Thus, greater Americanization of immigrant children's names appears as a *decline* in the foreignness index.

Next, I return to the full-count censuses from 1900 to 1940. I focus on two-generation households with at least one parent and one U.S.-born child living together, where *both parents are foreign-born* and *at least one parent is employed in manufacturing*.[67] Within these households, I consider children aged 10 or younger to ensure they were *born within the past decade*. I assign each child the foreignness index corresponding to their first name (by gender) from the same census. Finally, I aggregate these individual measures by county c and decade t, computing the *average name foreignness index* for young children in immigrant households where at least one parent worked in

[63] Parents face a trade-off between preserving their cultural traditions and helping children integrate into the host society. The names they choose reveal how they navigate this balance (e.g., Goldstein & Stecklov, 2016).

[64] The Stata command *abeclean* is provided by the authors, as a preliminary step to standardize first names, before applying the Abramitzky–Boustan–Eriksson (ABE) matching method.

[65] The index of name foreignness can be calculated as a relative probability:

$$R(Name) = \frac{\frac{Num.\ Foreign(Name)}{Tot.\ Foreign}}{\frac{Num.\ Native(Name)}{Tot.\ Native}}$$

This measure has a straightforward interpretation (for example, $R = 2$ means that a name is twice as likely to be used by foreign-born individuals as by natives), but it is sensitive to outliers. Therefore, I follow Abramitzky et al. (2020), and adopt the normalization index introduced by Fryer Jr and Levitt (2004) to measure distinctively black names. In practice, the standardized foreignness index is equivalent to $\frac{R}{1+R}$.

[66] Abramitzky et al. (2020) multiply the index by 100 so that it ranges from 0 to 100. I keep it between 0 and 1 for consistency with the other variables in the paper.

[67] To capture naming decisions made by immigrants in the United States, I restrict the sample to children born in the U.S. to two foreign-born parents.

manufacturing.

$$\textit{Foreignness Name of Children}_{ct} = \frac{\sum_{n=1}^{N_{ct}} \textit{Foreignness Index}(\textit{Name})_n}{N_{ct}} \tag{10}$$

where $n = 1, \ldots, N_{ct}$ indexes children aged 10 or younger who were born in the United States to foreign-born parents and reside in county c in decade t.

Summary Statistics. The summary statistics of the county-level variables are reported in Panel B of Appendix Table 2. The corresponding correlation matrix is reported in Appendix Table B.3.

6.2 Empirical Strategy

In this section, I examine the community consequences of electrification on manufacturing workers. The unit of analysis is county c in decade t. As described in Section 3, I combine data on the gradual expansion of the electric grid to identify the decade in which each U.S. county first became electrified. This defines the *treatment* at the county level. Once a county becomes electrified, it remains treated in all subsequent periods.

Econometric Framework: DID Analysis. Having established the causal effects of electrification on the ethnic integration within industries of the manufacturing labor force, I now turn to its impact on residential integration. I focus on workers employed in manufacturing. To analyze this, I use a *difference-in-differences* framework at the county–decade level, which closely parallels the earlier analysis at the industry–county–decade level. In this setting, I compare residential segregation among immigrant manufacturing workers in electrified counties with that in counties that are not yet electrified. The regression model is specified as follows:

$$Y_{ct} = \alpha_c + \phi_{s(c)t} + \beta_t \, \mathbb{I}\{t \geq \textit{Electrified}_{ct}\} + \mathbf{X}'_{ct} \, \Gamma + \varepsilon_{ct} \tag{11}$$

Here, Y_{ct} denotes the outcome of interest for county c in decade t: the residential segregation of manufacturing workers. Later, I apply the same empirical framework to outcomes related to the cultural assimilation of immigrant manufacturing workers, namely the native–foreign intermarriage rate and the foreignness of names given to children. Detailed definitions of the outcome variables are provided in Section 6.1.

The variable $\mathbb{I}\{t \geq \textit{Electrifiedct}\}$ is an indicator that equals one starting from the decade when county c first becomes electrified (denoted by $\textit{Electrified}_{ct}$) and in all subsequent decades. It equals zero in all decades before electrification. The control vector X_{ct} includes the logarithm of the 1900 county population, interacted with decade indicators, to account for baseline differences in county size. Because previous robustness checks show that the county share of manufacturing employment slightly anticipates electrification, I also include the 1900 manufacturing employment share, again interacted with decade indicators. County fixed effects, α_c, absorb all time-invariant local characteristics, such as geographic features or historical development patterns, that could jointly influence electrification and community integration. State-by-decade fixed effects, $\phi_{s(c)t}$, control for decade-specific shocks or trends common to all counties within a state. This specification produces a two-way fixed-effects model, with counties as the unit of observation and temporal trends allowed to vary by state. Standard errors, ε_{ct}, are clustered at the county level. Estimation is weighted by each county's share of manufacturing employment to reflect the sector's importance in the local economy. This is appropriate because the analysis focuses on the community integration of manufacturing workers, the group most directly affected by electrification. Following the earlier analysis, I exclude counties with 1900 populations above the 99th percentile, which correspond to the largest U.S. cities. These urban counties likely had access to electricity from local power plants before the earliest period observed in the data and are therefore classified as always treated.

Under the standard assumptions of parallel trends and no anticipation, the coefficient β_t captures the average treatment effect on the treated (ATT). It measures the causal effect of electrification on the residential segregation (and cultural assimilation) of immigrant workers in treated (electrified) counties relative to control (not-yet-electrified) counties. I estimate the DID model using the method proposed by Borusyak et al. (2024), which corrects common biases in staggered adoption settings. A detailed description of this estimator is provided in Section 4.2. The BJS approach also includes a direct test for the identifying assumptions, which I apply in each estimated model to ensure the validity of the causal interpretation of β_t.

Econometric Framework: OLS Analysis. As discussed in Section 6.1, a core element of social cohesion is a community's orientation toward the common good and the willingness of its members to contribute to public goods and services that benefit all (Schiefer & van der Noll, 2017, for a review). In ethnically diverse societies, achieving social cohesion is often more difficult. Differences in perspectives and values can weaken trust, reduce participation, and lower social capital (Alesina & La Ferrara, 2000; Putnam, 2000). Moreover, ethnic groups may hold different preferences over the type and level of public good provision, making agreement harder to reach and ultimately influencing local public finances (Rubinfield et al., 1987; Lieberman, 1993). Another mechanism that may affect the provision of local public goods is fiscal competition across groups: voters may prefer lower public spending when a substantial share of tax revenues collected from one group is used to fund goods shared with other ethnic groups (Alesina et al., 1999).

Consistent with this, evidence from the United States shows that areas with a higher presence of immigrants tend to exhibit less support for redistribution and lower provision of public goods, both historically and today (Tabellini, 2020; Alesina et al., 2023). Therefore, I build on this literature and assess whether electrification could *mitigate* the negative relationship between immigrant presence and provision of local public services. The underlying idea is that when workers are less segregated within industries along ethnic lines, the native majority may perceive immigrants more favorably, reducing social distance and potential resentment. To do so, I estimate the following ordinary least squares (OLS) regression model:

$$\begin{aligned} Y_{ct} = \alpha_c + \phi_{s(c)t} + \beta_1\, \mathbb{I}\{t \geq \mathit{Electrified}_{ct}\} + \beta_2\, \mathit{Immigrant\ Share}_{ct} \\ + \beta_3\, \mathit{Immigrant\ Share}_{ct} \times \mathbb{I}\{t \geq \mathit{Electrified}_{ct}\} + \mathbf{X}'_{ct}\, \Gamma + \varepsilon_{ct} \end{aligned} \tag{12}$$

The unit of analysis is county c in decade t, and the dependent variable Y_{ct} is the number of workers employed in occupations related to local provision of public services per 1,000 inhabitants. As described in Section 6.1, these occupations include teachers, doctors, police officers, firefighters, and public administrators. Higher values of Y_{ct} indicate greater provision of local public services through employment in essential services.

The key explanatory variables are the electrification indicator, the immigrant share, and their interaction. As in previous models, $\mathbb{I}\{t \geq \mathit{Electrified}_{ct}\}$ equals one in the decade when a county first becomes electrified (denoted by *Electrifiedct*) and in all subsequent decades, and zero in all decades before electrification. The variable $\mathit{Immigrant\ Share}_{ct}$ measures the share of county residents who are foreign-born. It is relevant to note that the robustness checks discussed in Section 4.5 show that electrification is *unrelated* to the share of the immigrant population at the county level. The coefficient $\beta 2$ represents the relationship between immigrant presence and employment in public service occupations in non-electrified counties. A negative estimate of β_2 would suggest that, absent electrification, counties with more immigrants employed fewer workers in public service occupations, consistent with the insight from previous work. The interaction term, $\mathit{Immigrant\ Share}_{ct} \times \mathbb{I}\{t \geq \mathit{Electrified}_{ct}\}$, allows this relationship to differ in electrified counties. The coefficient β_3 measures whether electrification modifies the relationship between immigrant share and employment in public service occupations. A positive β_3, opposite in sign to β_2, would imply that electrification attenuates the negative association between immigrant share and provision of local public services. By contrast, a β_3 with the same sign as β_2 would suggest that electrification amplifies the negative relationship.

The estimation model follows the same structure as in previous equations. It includes county fixed effects (α_c), which absorb time-invariant characteristics such as geography, initial infrastructure, or long-standing cultural attitudes. It also includes state-by-decade fixed effects ($\phi_{s(c)t}$), which capture shocks or policies affecting all counties within a state in a given decade. The control vector ($\mathbf{X}_{ct}$) contains the logarithm of the 1900 county population and the 1900 manufacturing employment share, each interacted with decade indicators. These controls account for differences related to county size and industrial base. Standard errors (ε_{ct}) are clustered at the county level to allow for serial correlation within counties over time. As in the earlier analyses, I exclude counties whose 1900 population exceeded the 99th percentile of the national distribution. These large urban counties likely had access to electricity from local power plants before the earliest period covered by my electrification data. Finally, I estimate the regression separately for counties with manufacturing employment shares above and below the 1900 median. This approach links the analysis of electrification and local public service provision to the county's manufacturing intensity, and thus to the degree of employment integration of immigrant workers within manufacturing industries. Even though this OLS specification does not allow for a causal interpretation of the results, it provides suggestive evidence on how the presence of immigrants relates to local provision of public services and whether electrification may moderate this relationship.

6.3 Electrification and the Residential Segregation of Manufacturing Labor Force

I begin by examining the effect of electrification on the *residential segregation* of the manufacturing labor force within counties. The DID estimates from Equation 11 are presented in Table 6. Counties that became electrified show a residential segregation index for manufacturing workers that is 1.0 point *lower* than that of non-electrified counties. This reduction corresponds to a 11.1% decline relative to the 1900 baseline mean of 9.0 index points. The pre-treatment test indicates that the pre-treatment coefficients are jointly indistinguishable from zero ($F = 0.59, p = 0.62$), supporting the identifying assumptions of parallel trends and no anticipation. Column (1) reports my preferred specification. Columns (2) and (3) show that the results remain very similar when I control for the county share of U.S.-born Whites and for the overall ethnic diversity of the population.

In studying immigrants' residential choices, previous research has emphasized the role of interactions between newly arrived immigrants and established co-ethnic communities. Newly arrived foreigners often settle in ethnic enclaves where compatriots already reside (e.g., Borjas, 2000; Eriksson & Ward, 2019; Xu, 2020). To capture this dynamic in the measurement of residential segregation, in Columns (4)–(6) of Table 6, I redefine "immigrants" to include both foreign-born individuals (*first-generation*) and U.S.-born individuals with at least one foreign-born parent (*second-generation*).[68] Under this definition, ethnic groups include U.S.-born Whites (born to two U.S.-born parents), U.S.-born Blacks (born to two U.S.-born parents), and immigrants (first and second generation combined) categorized by their country of origin. Grouping the two generations by shared origin allows me to account for ethnic communities that extend across generations. Results are consistent, and the main conclusion remains valid. When considering both first- and second-generation immigrants to calculate residential segregation, I find that electrification leads to a *decline* of 0.8 points, which corresponds to 8.8% of the 1900 baseline mean of 9.1 index-points (Column 4). Results are similar when accounting for the county share of U.S.-native Whites, U.S.-native with foreign-born parents, and the overall ethnic diversity of the population.

6.4 Electrification and the Provision of Local Public Services

Next, I examine the effect of electrification on the provision of local public services, proxied by the number of workers in related occupations –teachers, doctors, police officers, firefighters, and public administrators. I estimate this relationship using the OLS specification described in Equation 12. Results are presented in Table 8. Column (1) reports the estimates for the full sample, while

[68] For example, an Italian-born father and his U.S.-born son are both classified as part of the Italian ethnic group. This differs from the baseline definition, which includes only foreign-born individuals (in this case, only those born in Italy). For U.S.-born individuals with parents from different countries, I assign the ethnic origin based on the father's country of birth. This ensures that each person is assigned to a single, clearly defined immigrant group.

Columns (2) and (3) split the sample by counties with manufacturing employment shares in 1900 above and below the median.

Overall, the results show a **negative** and statistically significant relationship between the county share of foreign-born residents and employment in public service occupations. This finding is consistent with previous research. For the same historical period, Tabellini (2020) documents a negative association between immigrant presence and public spending on local goods in major U.S. cities. My results extend this evidence to smaller and more rural counties, showing that the relationship was also present outside large urban centers and was particularly strong in areas with higher manufacturing intensity in 1900.

Turning to the role of electrification, I find evidence of an *attenuation effect*. In other words, results suggest that electrification weakens the negative link between immigrant presence and public service employment. In the full sample, the coefficient on the interaction term indicates an increase of 1.8 additional public service workers per thousand inhabitants in electrified counties. This represents 23.8% of the 1900 baseline mean (7.8 workers per thousand inhabitants) and 18.5% of the magnitude of the negative immigrant effect in non-electrified counties. The attenuation effect is concentrated in counties with high manufacturing intensity at baseline. In these counties, electrification is associated with 2.9 more public service workers per thousand inhabitants, corresponding to 36.0% of the baseline mean and 23.7% of the negative relationship observed in non-electrified counties. By contrast, in counties with low initial manufacturing intensity, the coefficient on the interaction term is negative but imprecisely estimated, suggesting no clear moderating effect of electrification (and possibly even a mild reinforcement of the negative relationship).

6.5 Electrification and the Cultural Assimilation of Immigrant Manufacturing Workers

Having shown that electrification promoted both employment and residential integration among manufacturing workers, I finally examine whether it also influenced their cultural assimilation. I focus on two outcomes: intermarriage between immigrant and native spouses, and the foreignness of names given to children by foreign-born parents. In both cases, I restrict the sample to households in which at least one spouse or parent is employed in manufacturing. I estimate the DID model described in Equation 11, and report the results in Table 9.

I find weak effects of electrification on the cultural assimilation of immigrant manufacturing workers. First, I examine intermarriage between immigrant and native spouses, focusing on marriages likley celebrated in the previous decade. Electrified counties show a positive, though only marginally significant, increase in intermarriage compared with non-electrified ones. Although the coefficient is not precisely estimated, the magnitude is meaningful, corresponding to an 11.2% increase relative to the 1900 baseline mean. Intermarriage reflects both the willingness of the foreign-born spouse to integrate into the host society and the willingness of the native spouse to enter a long-term relationship with someone from a different background. It also captures the broader social acceptance of such unions, making it a measure closely linked to social cohesion.

However, intermarriage is an extreme outcome of integration, as marrying across ethnic or racial lines is often viewed as the ultimate marker of assimilation (Wildsmith et al., 2003). This could explain the weak results.

I also study changes in the naming patterns of children born in the United States to foreign-born parents during the previous decade. Adopting a more common American name is typically seen as a sign of cultural assimilation into the host society (Abramitzky et al., 2020). However, a priori, it is unclear how naming choices would respond to greater employment and residential integration. More integrated immigrants might choose less ethnically distinctive names for their children as a sign of assimilation, or more distinctive ones if they feel less pressure to avoid discrimination once integration has reduced social barriers (Fouka, 2019, 2020; Fouka et al., 2022; Abramitzky et al., 2024; Fouka, 2024). Consistent with this ambiguity, I find no statistically significant difference in naming practices between electrified and non-electrified counties.

7 Conclusions

This paper examines how technological change in production affects the social cohesion of ethnically diverse societies. Using newly digitized data on the early expansion of the electric grid in the United States between 1900 and 1940, I implement a DID analysis to study how electrification influenced the employment and community integration of immigrant and native manufacturing workers.

The results show that electrification affected both workplace and community dynamics. After electrification, manufacturing industries became more diverse overall and more integrated across occupations. Accordingly manufacturing workers experienced more ethnic diversity not only in their industry overall, but also in their occupations. Workers also lived in more ethnically mixed neighborhoods, suggesting that integration extended beyond the factory floor. In counties with high manufacturing intensity at baseline, electrification weakened the negative relationship between the presence of immigrants and the provision of local public services. This pattern is consistent with diverse groups becoming more socially integrated, strengthening overall social cohesion.

By reducing barriers that kept workers segregated in production, the technological change studied here fosted employment and community integration, and strengthened the social fabric of a diverse society. While this paper examines the effects of electrification in manufacturing on the integration of ethnically diverse workers, the underlying mechanisms are not unique to this context. For example, the adoption of containerized shipping standardized handling tasks and reduced reliance on informal ethnic labor networks that had long dominated port work (Levinson, 2016). Similarly, the spread of office mechanization –and later digital communication technologies– standardized and codified tasks, formalized communication within offices, and routinized workflows through written documentation and formal systems, reducing the importance of shared language, cultural background, and social proximity (Chandler Jr & Cortada, 2000; Beniger, 2009).

Ultimately, understanding the specific features of technological change that promote integration and those that deepen existing social divisions remains an important question for future research.

Identifying these mechanisms can help explain when and how technological progress contributes not only to economic growth, but also to more cohesive and inclusive societies.

References

Central electric light and power stations and street and electrical railways with summary of the electrical industries. (1912), Department of Commerce, U.S. Bureau of the Census.

Proposed national commission to solve water-power problems (No. *17*). (1912), Electrical World.

National power survey interim report: Power series no. 1. (1935), U.S. Federal Power Commision.

Report on the status of interconnections and pooling of electric utility systems in the united states (based on data compiled in 1959). (1962), Edison Electrical Institute.

Abramitzky, R., Ager, P., Boustan, L., Cohen, E., & Hansen, C. W. (2023). The effect of immigration restrictions on local labor markets: Lessons from the 1920s border closure. *American Economic Journal: Applied Economics*, *15*(1), 164–191.

Abramitzky, R., & Boustan, L. (2017). Immigration in american economic history. *Journal of Economic Literature*, *55*(4), 1311–1345.

Abramitzky, R., Boustan, L., & Connor, D. S. (2024). Leaving the enclave: Historical evidence on immigrant mobility from the industrial removal office. *Journal of Economic History*, *84*(2), 352–394.

Abramitzky, R., Boustan, L., & Eriksson, K. (2020). Do immigrants assimilate more slowly today than in the past? *American Economic Review: Insights*, *2*(1), 125–141.

Abramitzky, R., Boustan, L., & Eriksson, K. (2019). To the new world and back again: Return migrants in the age of mass migration. *ILR Review*, *72*(2), 300–322.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, *102*(5), 1832–1856.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy*, *122*(3), 467–506.

Abramitzky, R., Boustan, L. P., Jácome, E., & Pérez, S. (2021). Intergenerational mobility of immigrants in the us over two centuries. *American Economic Review*, *11*(2), 580–608.

Acemoglu, D., & Autor, D. H. (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (pp. 1043–1171, Vol. 4B). Elsevier.

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, *33*(2), 3–30.

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, *128*(6), 2188–2244.

Acemoglu, D., & Wolitzky, A. (2025). *Employment and community: Socioeconomic cooperation and its breakdown*, NBER Working Paper No. 32773.

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., & Zhuravskaya, E. (2015). Radio and the rise of the nazis in prewar germany. *Quarterly Journal of Economics*, *130*(4), 1885–1939.

Ager, P., Goñi, M., & Salvanes, K. G. (*forthcoming*). Gender-biased technological change: Milking machines and the exodus of women from farming. *American Economic Review.*

Alba, R. D., & Golden, R. M. (1986). Patterns of ethnic marriage in the united states. *Social forces*, *65*(1), 202–223.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics*, *114*(4), 1243–1284.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, *8*(2), 155–194.

Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, *21*, 101–138.

Alesina, A., & La Ferrara, E. (2000). Participation in heterogeneous communities. *Quarterly Journal of Economics*, *115*(3), 847–904.

Alesina, A., Miano, A., & Stantcheva, S. (2023). Immigration and redistribution. *Review of Economic Studies*, *90*(1), 1–39.

Alesina, A., & Zhuravskaya, E. (2011). Segregation and the quality of government in a cross section of countries. *American Economic Review*, *101*(5), 1872–1911.

Allcott, H., Collard-Wexler, A., & O'Connell, S. D. (2016). How do electricity shortages affect industry? evidence from india. *American Economic Review*, *106*(3), 587–624.

Armand, A., Atwell, P., Gomes, J., Musillo, G., & Schenk, Y. (2024). *It's a bird, it's a plane, it's superman! using mass media to fight intolerance*, IZA Discussion Papers No. 17098.

Atkeson, A., & Kehoe, P. J. (2007). Modeling the transition to a new economy: Lessons from two technological revolutions. *American Economic Review*, *97*(1), 64–88.

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, *103*(5), 1553–1597.

Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). The polarization of the us labor market. *American Economic Review*, *96*(2), 189–194.

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

Barrett, J. R. (1984). Unity and fragmentation: Class, race, and ethnicity on chicago's south side, 1900-1922. *Journal of Social History*, *18*(1), 37–55.

Barrett, J. R. (1990). *Work and community in the jungle: Chicago's packinghouse workers, 1894-1922*. University of Illinois Press.

Bazzi, S., Fiszbein, M., & Gebresilasse, M. (2020). Frontier culture: The roots and persistence of "rugged individualism" in the united states. *Econometrica*, *88*(6), 2329–2368.

Bell, D. (1973). *The coming of post-industrial society: A venture in social forecasting*. Basic Books.

Beniger, J. (2009). *The control revolution: Technological and economic origins of the information society*. Harvard University Press.

Blalock, H. M., et al. (1967). *Toward a theory of minority-group relations* (Vol. 325). Wiley.

Blau, P. M., Beeker, C., & Fitzpatrick, K. M. (1984). Intersecting social affiliations and intermarriage. *Social Forces*, *62*(3), 585–606.

Blouin, A., & Mukand, S. W. (2019). Erasing ethnicity? propaganda, nation building, and identity in rwanda. *Journal of Political Economy*, *127*(3), 1008–1062.

Bodnar, J. E. (1977). *Immigration and industrialization: Ethnicity in an american mill town, 1870-1940*. University of Pittsburgh Press.

Bonacich, E. (1972). A theory of ethnic antagonism: The split labor market. *American Sociological Review*, *37*(5), 547–559.

Borenstein, S., & Bushnell, J. (2015). The us electricity industry after 20 years of restructuring. *Annual Review of Economics*, *7*(1), 437–463.

Borjas, G. J. (2000). Ethnic enclaves and assimilation. *Swedish Economic Policy Review*, *7*(2), 89–122.

Borusyak, K., & Hull, P. (2023). Nonrandom exposure to exogenous shocks. *Econometrica*, *91*(6), 2155–2185.

Borusyak, K., & Jaravel, X. (2018). *Revisiting event study designs, with an application to the estimation of the marginal propensity to consume*, Working Paper, Harvard University.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, *91*(6), 3253–3285.

Braverman, H. (1974). *Labor and monopoly capital: The degradation of work in the twentieth century*. Monthly Review Press.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics*, *65*(1), 83–108.

Brody, D. (1960). *Steelworkers in america: The nonunion era*. University of Illinois Press.

Brynjolfsson, E., & Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, *14*(4), 23–48.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative ai at work. *Quarterly Journal of Economics*, *140*(2), 889–942.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Card, D., & DiNardo, J. E. (2002). Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics*, *20*(4), 733–783.

Caroli, E., & Van Reenen, J. (2001). Skill-biased organizational change? evidence from a panel of british and french establishments. *Quarterly Journal of Economics*, *116*(4), 1449–1492.

Carpenter, N. (1927). *Immigrants and their children, 1920: A study based on census statistics relative to the foreign born and the native white of foreign or mixed parentage*. Department of Commerce, U.S. Bureau of the Census.

Carter, S., & Sutch, R. (1998). Historical background to current immigration issues. In J. S. Smith & B. Edmonston (Eds.), *The immigration debate: Studies on the economic, demographic, and fiscal effects of immigration* (pp. 289–366). National Academy Press.

Castells, M. (1996). *The rise of the network society*. Blackwell Publishers.

Chandler Jr, A. D., & Cortada, J. W. (2000). *A nation transformed by information: How information has shaped the united states from colonial times to the present*. Oxford University Press.

Clay, K., Hernandez-Cortes, D., Jha, A., Lewis, J. A., Miller, N. S., & Severnini, E. R. (2025). *The social lifecycle impacts of power plant siting in the historical united states*, NBER Working Paper No. 34109.

Cohn, J. A. (2017). *The grid: Biography of an american technology*. MIT press.

Cohn, J. A. (2019). When the grid was the grid: The history of north america's brief coast-to-coast interconnected machine [scanning our past]. *Proceedings of the IEEE*, *107*(1), 232–243.

Commons, J. R. (1935). Introduction to volumes iii and iv. In J. R. Commons, D. J. Saposs, H. L. Summer, E. B. Mittelman, H. E. Hoagland, J. B. Andrews, & S. Perlman (Eds.), *History of labor in the united states, 1896–1932. volumes i–iv.* (pp. ix–xxx). The Macmillan Company.

Commons, J. R. (1904). Labor conditions in meat packing and the recent strike. *Quarterly Journal of Economics*, *19*(1), 1–32.

Conzen, K. N. (1979). Immigrants, immigrant neighborhoods, and ethnic identity: Historical issues. *The Journal of American History*, *66*(3), 603–615.

Cortés, P., Feng, Y., Guida-Johnson, N., & Pan, J. (2024). *Automation and gender: Implications for occupational segregation and the gender skill gap*, NBER Working Paper No. 32030.

Cutler, D. M., & Glaeser, E. L. (1997). Are ghettos good or bad? *Quarterly Journal of Economics*, *112*(3), 827–872.

Dale, L. H., Spencer, L. L., Wheeler, H. E., Amundson, P. S., & Urani, J. M. (1946). The legal status of women in the united states of america: United states summary, cumulative supplement, 1938-1945. *Women's Bureau Bulletin, No. 157-A.*

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, *27*(2), 281–314. https://doi.org/10.1086/599336

David, P. A. (1990). The dynamo and the computer: An historical perspective on the modern productivity paradox. *American Economic Review*, *80*(2), 355–361.

David, P. A., & Wright, G. (1999). *Early twentieth century productivity growth dynamics: An inquiry into the economic history of "our ignorance"y*, University of Oxord, Discussion Papers in Economic and Social History, No. 33.

David, P. A., & Wright, G. (2003). General purpose technologies and surges in productivity. In P. A. David & M. Thomas (Eds.), *The economic future in historical perspective* (pp. 135–166). Oxford University Press.

de Chaisemartin, C., & d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.

de Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., & Zhuravskaya, E. (2014). Cross-border media and nationalism: Evidence from serbian radio in croatia. *American Economic Journal: Applied Economics*, *6*(3), 103–132.

Devine Jr, W. D. (1983). From shafts to wires: Historical perspective on electrification. *Journal of Economic History*, *43*(2), 347–372.

Dillingham, W. P., Lauck, W. J., Cance, A. E., & Millis, H. A. (1911). *Immigrants in industries (in twenty-five parts)* [U.S. Immigrant Commission (1907–1910)]. Government Printing Office.

DiMaggio, P., Hargittai, E., Neuman, W. R., & Robinson, J. P. (2001). Social implications of the internet. *Annual Review of Sociology*, *27*(1), 307–336.

Dinkelman, T. (2011). The effects of rural electrification on employment: New evidence from south africa. *American Economic Review*, *101*(7), 3078–3108.

Du Boff, R. B. (1979). *Electric power in american manufacturin, 1889–1958*. Arno Press.

Du Boff, R. B. (1967). The introduction of electric power in american manufacturing. *Economic History Review*, *20*(3), 509–518.

Durante, R., Pinotti, P., & Tesei, A. (2019). The political legacy of entertainment tv. *American Economic Review*, *109*(7), 2497–2530.

Durkheim, E. (1893). *The division of labour in society* (W. D. Halls, Trans.) [Translated work published in 1984]. The Macmillan Press.

Easterly, W., Ritzen, J., & Woolcock, M. (2006). Social cohesion, institutions, and growth. *Economics & Politics*, *18*(2), 103–120.

Elmes, G. (1996). The changing geography of electric energy in the united states: Restrospect and prospect. *Geography*, *81*(4), 347–360.

Enikolopov, R., Petrova, M., & Zhuravskaya, E. (2011). Media and political persuasion: Evidence from russia. *American Economic Review*, *101*(7), 3253–3285.

Eriksson, K., & Ward, Z. (2019). The residential segregation of immigrants in the united states from 1850 to 1940. *Journal of Economic History*, *79*(4), 989–1026.

Feigenbaum, J., & Gross, D. P. (2024). Answering the call of automation: How the labor market adjusted to mechanizing telephone operation. *Quarterly Journal of Economics*, *139*(3), 1879–1939.

Ferrara, A., Testa, P. A., & Zhou, L. (2024). New area-and population-based geographic crosswalks for us counties and congressional districts, 1790–2020. *Historical Methods*, *57*(2), 67–79.

Ferraz, C., & Finan, F. (2008). Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, *123*(2), 703–745.

Fiszbein, M., Lafortune, J., Lewis, E. G., & Tessada, J. (2024). *Powering up productivity: The effects of electrification on us manufacturing*, NBER Working Paper No. 28076.

Forslund, E., Meriläinen, J., & Zipfel, C. (2025). *Engines of empowerment: Cattle tending, the milking machine, and women in politics*, Working Paper, Stockholm School of Economics, Mistra Center for Sustainable Markets.
Fouka, V. (2020). Backlash: The unintended effects of language prohibition in us schools after world war i. *Review of Economic Studies*, *87*(1), 204–239.
Fouka, V. (2019). How do immigrants respond to discrimination? the case of germans in the us during world war i. *American Political Science Review*, *113*(2), 405–422.
Fouka, V. (2024). State policy and immigrant integration. *Annual Review of Political Science*, *27*, 25–46.
Fouka, V., Mazumder, S., & Tabellini, M. (2022). From immigrants to americans: Race and assimilation during the great migration. *Review of Economic Studies*, *89*(2), 811–842.
Fryer Jr, R. G., & Levitt, S. D. (2004). The causes and consequences of distinctively black names. *Quarterly Journal of Economics*, *119*(3), 767–805.
Fulford, S. L., Petkov, I., & Schiantarelli, F. (2022). *Ancestry diversity and local public spending in the us since 1870*, Working Paper, Northeastern University.
Gaggl, P., Gray, R., Marinescu, I., & Morin, M. (2021). Does electricity drive structural transformation? evidence from the united states. *Labour Economics*, *68*, 101944.
Gentzkow, M. (2006). Television and voter turnout. *Quarterly Journal of Economics*, *121*(3), 931–972.
Goldin, C. D. (1994). The political economy of immigration restriction in the united states, 1890 to 1921. In *The regulated economy: A historical approach to political economy* (pp. 223–258). University of Chicago Press.
Goldin, C. D., & Katz, L. F. (1998). The origins of technology-skill complementarity. *Quarterly Journal of Economics*, *113*(3), 693–732.
Goldstein, J. R., & Stecklov, G. (2016). From patrick to john f. ethnic names and occupational success in the last era of mass migration. *American Sociological Review*, *81*(1), 85–106.
Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.
Gordon, M. M. (1964). *Assimilation in american life: The role of race, religion, and national origins*. Oxford University Press.
Gray, R. (2013). Taking technology to task: The skill content of technological change in early twentieth century united states. *Explorations in Economic History*, *50*(3), 351–367.
Graziosi, A. (1981). Common laborers, unskilled workers: 1880–1915. *Labor History*, *22*(4), 512–544.
Gutman, H. G. (1973). Work, culture, and society in industrializing america, 1815-1919. *American Historical Review*, *78*(3), 531–588.
Haines, M. R. (1996). Long-term marriage patterns in the united states from colonial times to the present. *The History of the Family*, *1*(1), 15–39.
Hall, F. S., & Brooke, E. W. (1919). *American marriage laws in their social aspects. a digest*. Russel Sage Foundation.
Hamilton, V. E. (2012). The age of marital capacity: Reconsidering civil recognition of adolescent marriage. *Boston University Law Review*, *92*, 1817–1863.
Hatton, T. J., & Williamson, J. G. (1998). *The age of mass migration: Causes and economic impact*. Oxford University Press, USA.
Hellerstein, J. K., & Neumark, D. (2008). Workplace segregation in the united states: Race, ethnicity, and skill. *Review of Economics and Statistics*, *90*(3), 459–477.
Helpman, E. (1998). *General purpose technologies and economic growth*. MIT Press.
Hirschman, C. (2005). Immigration and the american century. *Demography*, *42*(4), 595–620.

Hirschman, C., & Mogford, E. (2009). Immigration and the american industrial revolution from 1880 to 1920. *Social Science Research*, *38*(4), 897–920.
Hornuf, L., Rieger, M. O., & Hartmann, S. A. (2023). Can television reduce xenophobia? the case of east germany. *Kyklos*, *76*(1), 77–100.
Hounshell, D. (1984). *From the american system to mass production, 1800-1932: The development of manufacturing technology in the united states*. Johns Hopkins University Press.
Hughes, T. P. (1979). The electrification of america: The system builders. *Technology and Culture*, *20*(1), 124–161.
Hughes, T. P. (1983). *Networks of power: Electrification in western society, 1880-1930*. Johns Hopkins University press.
Hutchinson, E. P. (1956). *Immigrants and their children, 1850-1950*. Wiley.
Hutchinson, E. P. (1981). *Legislative history of american immigration policy, 1798-1965*. University of Pennsylvania Press.
Inglehart, R., & Baker, W. E. (2000). Modernization, cultural change, and the persistence of traditional values. *American Sociological Review*, *65*(1), 19–51.
Jaimovich, N., & Siu, H. E. (2020). Job polarization and jobless recoveries. *Review of Economics and Statistics*, *102*(1), 129–147.
Jayes, J., Molinder, J., & Enflo, K. (2025). Power for progress: The impact of electricity on individual labor market outcomes. *Explorations in Economic History*, 101702.
Jevons, H. S. (1931). The second industrial revolution. *The Economic Journal*, *41*(161), 1–18.
Jovanovic, B., & Rousseau, P. L. (2005). General purpose technologies. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of economic growth* (pp. 1181–1224, Vol. 1). Elsevier.
Juhász, R., Squicciarini, M. P., & Voigtländer, N. (2024). Technology adoption and productivity growth: Evidence from industrialization in france. *Journal of Political Economy*, *132*(10), 3215–3259.
Kassem, D. (2024). Does electrification cause industrial development? grid expansion and firm turnover in indonesia. *Journal of Development Economics*, *167*, 103234.
Kastis, Y., & Vipont, H. (2025). *Organizational practices and technology adoption: Evidence from jewish immigration and the tailoring industry in england*, Working Paper.
Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics*, *107*(1), 35–78.
Kitchens, C., & Fishback, P. (2015). Flip the switch: The impact of the rural electrification administration 1935–1940. *Journal of Economic History*, *75*(4), 1161–1195.
Kline, P., & Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *Quarterly Journal of Economics*, *129*(1), 275–331.
Kuznets, S. (1971). The contribution of immigration to the growth of the labor force. In R. W. Fogel & S. L. Engerman (Eds.), *The reinterpretation of american economic history* (pp. 396–401). Harper & Row New York.
Kydland, F. E., & Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica*, 1345–1370.
Lafortune, J., Lewis, E., & Tessada, J. (2019). People and machines: A look at the evolving relationship between capital and skill in manufacturing, 1860–1930, using immigration shocks. *Review of Economics and Statistics*, *101*(1), 30–43.
Landes, D. S. (2003). *The unbound prometheus: Technological change and industrial development in western europe from 1750 to the present*. Cambridge University Press.
Lang, K. (1986). A language theory of discrimination. *Quarterly Journal of Economics*, *101*(2), 363–382.

Lazear, E. P. (1995). *Personnel economics*. MIT press.

Lee, E. (2003). *At america's gates: Chinese immigration during the exclusion era, 1882-1943*. University of North Carolina Press.

Lee, K., Miguel, E., & Wolfram, C. (2020). Experimental evidence on the economics of rural electrification. *Journal of Political Economy*, *128*(4), 1523–1565.

Levinson, M. (2016). *The box: How the shipping container made the world smaller and the world economy bigger*. Princeton University Press.

Lewis, J., & Severnini, E. (2020). Short-and long-run impacts of rural electrification: Evidence from the historical rollout of the us power grid. *Journal of Development Economics*, *143*, 102412.

Licht, W. (1995). *Industrializing america*. Johns Hopkins University Press.

Lieberman, M. (1993). *Public education: An autopsy*. Harvard University Press.

Lieberson, S. (2000). *A matter of taste: How names, fashions, and culture change*. Yale University Press.

Lieberson, S., & Bell, E. O. (1992). Children's first names: An empirical study of social taste. *American Journal of Sociology*, *98*(3), 511–554.

Lipscomb, M., Mobarak, A. M., & Barham, T. (2013). Development effects of electrification: Evidence from the topographic placement of hydropower plants in brazil. *American Economic Journal: Applied Economics*, *5*(2), 200–231.

Locke, E. (2025). *Ethnic-occupational niches: Evidence from the age of mass migration*, Working Paper.

Logan, T. D., & Parman, J. M. (2017). Segregation and homeownership in the early twentieth century. *American Economic Review*, *107*(5), 410–414.

Long, J., Medici, C., Qian, N., & Tabellini, M. (2024). *The impact of the chinese exclusion act on the economic development of the western us*, NBER Working Paper No. 33019.

Marcus, M., & Sant'Anna, P. H. (2021). The role of parallel trends in event study settings: An application to environmental economics. *Journal of the Association of Environmental and Resource Economists*, *8*(2), 235–275.

Margo, R. A. (1990). The competitive dynamics of racial exclusion: Employment segregation in the south, 1900-1950. In R. A. Margo (Ed.), *Race and schooling in the south, 1880-1950: An economic history* (pp. 87–108). University of Chicago Press.

Martin, G. J., & Yurukoglu, A. (2017). Bias in cable news: Persuasion and polarization. *American Economic Review*, *107*(9), 2565–2599.

Marx, K., & Engels, F. (1848). *Communist manifesto* (S. Moore, Trans.) [Translated work published in 1888]. Charles H. Kerr & Company.

Medici, C. (2025). *Closing ranks: Organized labor and immigration*, CESifo Working Paper No. 11437.

Mokyr, J. (1992). *The lever of riches: Technological creativity and economic progress*. Oxford University Press.

Mokyr, J. (2005). Long-term economic growth and the history of technology. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of economic growth* (pp. 1113–1180, Vol. 1). Elsevier.

Mokyr, J. (2001). The rise and fall of the factory system: Technology, firms, and households since the industrial revolution. *Carnegie-Rochester Conference Series on Public Policy*, *55*(1), 1–45.

Mokyr, J., & Strotz, R. H. (1998). The second industrial revolution, 1870-1914. *Storia dell'economia Mondiale*, *21945*(1), 219–245.

Mowery, D. C., & Rosenberg, N. (1999). *Paths of innovation: Technological change in 20th-century america*. Cambridge University Press.

Nuwer, M. (1988). From batch to flow: Production technology and work-force skills in the steel industry, 1880–1920. *Technology and Culture*, *29*(4), 808–838.

Nye, D. E. (2013). *America's assembly line*. Mit Press.

Nye, D. E. (1992). *Electrifying america: Social meanings of a new technology*. Mit Press.

Olken, B. A. (2009). Do television and radio destroy social capital? evidence from indonesian villages. *American Economic Journal: Applied Economics*, *1*(4), 1–33.

Ottaviano, G. I., & Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, *10*(1), 152–197.

Pagnini, D. L., & Morgan, S. P. (1990). Intermarriage and social distance among us immigrants at the turn of the century. *American journal of sociology*, *96*(2), 405–432.

Polanyi, K. (1944). *The great transformation*. Farrar Rinehart.

Posch, M., & Raz, I. T. (2025). *Doux commerce: Markets, culture, and cooperation in 1850-1920 u.s.*, Working Paper.

Posch, M., Schulz, J., & Henrich, J. (2024). *How cultural diversity drives innovation: Surnames and patents in u.s. history*, SSRN Working Paper 4531209.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of american community*. Simon & Schuster.

Putnam, R. D. (1995). Tuning in, tuning out: The strange disappearance of social capital in america. *PS: Political Science & Politics*, *28*(4), 664–683.

Rosenberg, N. (1998). The role of electricity in industrial development. *The Energy Journal*, *19*(2), 7–24.

Ross, S. G. (1985). A perspective on international tax policy. *Tax Notes*, *26*, 701–706.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, *4*(3), 305–322.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, *235*(2), 2218–2244.

Rubinfield, D. L., Shapiro, P., & Roberts, J. (1987). Tiebout bias and the demand for local public schooling. *Review of Economics and Statistics*, 426–437.

Ruggles, S., Flood, S., Sobek, M., Backman, D., Chen, A., Cooper, G., Richards, S., Rodgers, R., & Schouweiler, M. (2024). Ipums ancestry full count data: Version 4.0 [dataset]. *Minneapolis, MN, IPUMS.*

Rumsey, D., & Williams, M. (2002). Historical maps in gis. In A. K. Knowles (Ed.), *Past time, past place: Gis for history* (pp. 1–18). ESRI Press.

Russo, G. (2024). *Media and assimilation: Evidence from the golden age of radio*, SSRN Working Paper 4813824.

Schiefer, D., & van der Noll, J. (2017). The essentials of social cohesion: A literature review. *Social Indicators Research*, *132*(2), 579–603.

Sequeira, S., Nunn, N., & Qian, N. (2020). Immigrants and the making of america. *Review of Economic Studies*, *87*(1), 382–419.

Severnini, E. (2023). The power of hydroelectric dams: Historical evidence from the united states over the twentieth century. *The Economic Journal*, *133*(649), 420–459.

Sokoloff, K. L. (1984). Was the transition from the artisanal shop to the nonmechanized factory associated with gains in efficiency?: Evidence from the us manufacturing censuses of 1820 and 1850. *Explorations in Economic History*, *21*(4), 351–382.

Sonenblum, S. (1990). Electrificationon and productivityity growth in manufacturing. In S. H. Schurr, C. C. Burwell, W. S. Devine, & S. Sonenblum (Eds.), *Electricity in the american economy: Agent of technological progress* (pp. 277–324). Greenwood Press.

Sue, C. A., & Telles, E. E. (2007). Assimilation and gender in naming. *American Journal of Sociology*, *112*(5), 1383–1415.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, *49*(2), 326–365.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *Review of Economic Studies*, *87*(1), 454–486.

Thompson, E. P. (1967). Time, work-discipline, and industrial capitalism. *Past Present*, *38*(1), 56–97.

Tomaskovic-Devey, D., Zimmer, C., Stainback, K., Robinson, C., Taylor, T., & McTague, T. (2006). Documenting desegregation: Segregation in american workplaces by race, ethnicity, and sex, 1966–2003. *American Sociological Review*, *71*(4), 565–588.

Tönnies, F. (1887). *Community and society* (C. P. Loomies, Trans.) [Translated work published in 1957]. Michigan State University Press.

Trevor, J. B. (1925). *Japanese exclusion: A study of the policy and the law.* Government Printing Office.

Vidart, D. (2024). Human capital, female employment, and electricity: Evidence from the early 20th-century united states. *Review of Economic Studies*, *91*(1), 560–594.

Vidart, D. (2025). *Revisiting the link between electrification and fertility: Evidence from the early 20th century united states*, SSRN Working Paper 5397596.

Waldinger, R. (1987). Beyond nostalgia: The old neighborhood revisited. *New York Affairs*, *10*(1), 1–12.

Wang, T. (2021). Media, pulpit, and populist persuasion: Evidence from father coughlin. *American Economic Review*, *111*(9), 3064–3092.

Waters, M. C., & Jiménez, T. R. (2005). Assessing immigrant assimilation: New empirical and theoretical challenges. *Annual Review of Sociology*, *31*, 105–125.

Weber, M. (1922). *Economy and society. an outline of interpretive sociology* (G. Roth & C. Wittich, Trans.) [Translated work published in 1968]. Bedminster Press.

Wildsmith, E., Gutmann, M. P., & Gratton, B. (2003). Assimilation and intermarriage for us immigrant groups, 1880–1990. *The History of the Family*, *8*(4), 563–584.

Williams Jr, R. M. (1947). The reduction of intergroup tensions: A survey of research on problems of ethnic, racial, and religious group relations. *Social Science Research Council Bulletin.*

Xu, D. (2020). The effects of immigration restriction laws on immigrant segregation in the early twentieth century us. *Journal of Comparative Economics*, *48*(2), 422–447.

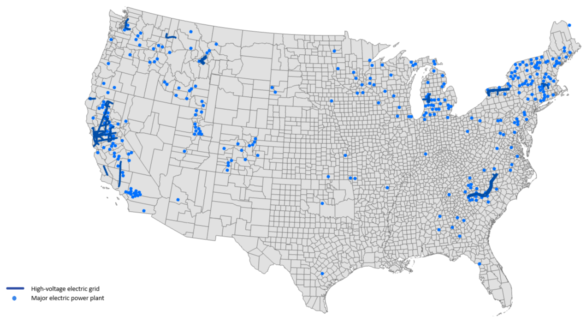
Xu, D., & Zhang, Y. (2022). Identifying ethnic occupational segregation. *Journal of Population Economics*, *35*(3), 1261–1296.

Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the rwandan genocide. *Quarterly Journal of Economics*, *129*(4), 1947–1994.

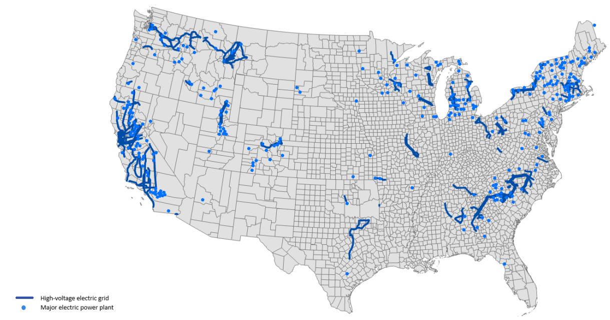
Figures

Figure 1: Electrification expansion in the United States, 1910–1940

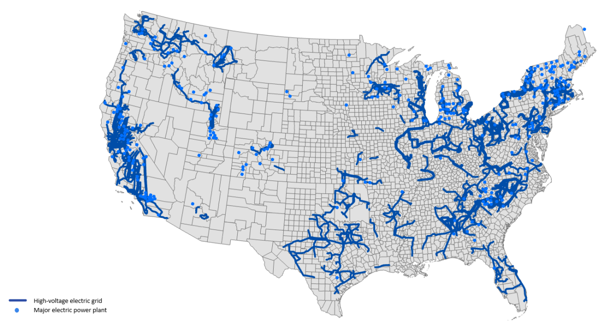
(a) 1910



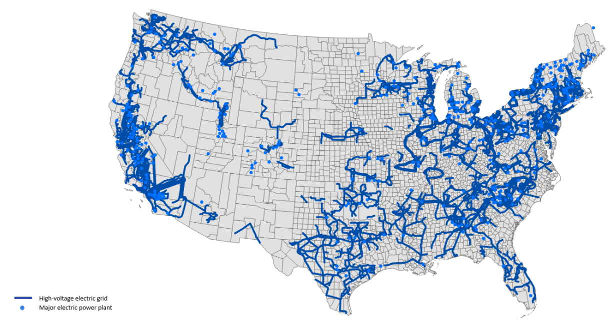
(b) 1920



(c) 1930



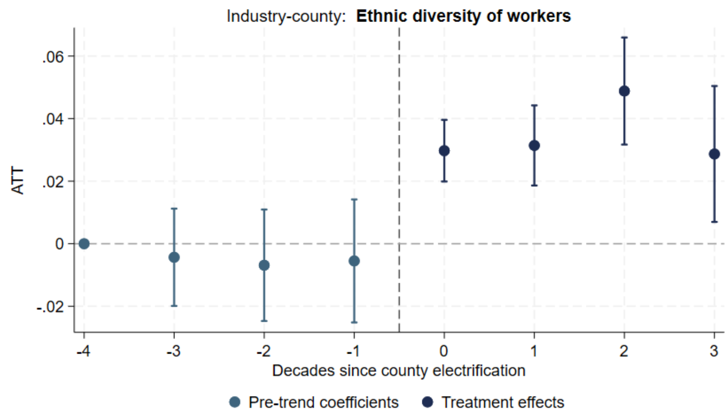
(d) 1940



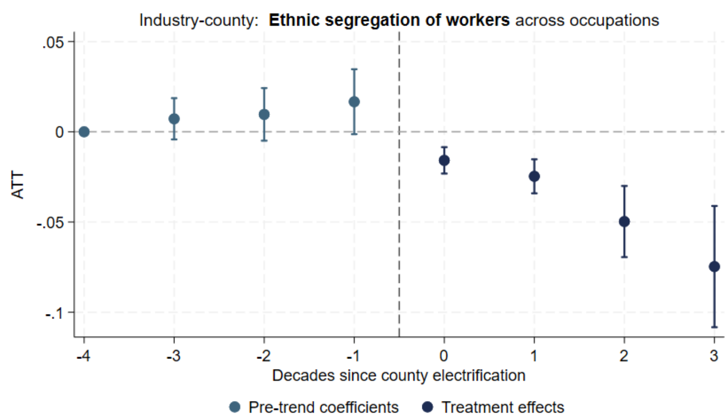
Note: This figure illustrates the decennial expansion of electrification between 1910 and 1940. The dark-blue lines represent the high-voltage electric grid, while the light-blue dots indicate the location of a major electric power plant. For 1910, Figure 1a shows the 1908 electric grid map and the 1912 power plant data. For 1920, Figure 1b shows the 1918 electric grid map and the 1912 power plants. For 1930, Figure 1c shows the 1928 electric grid map and the 1912 power plants. For 1940, Figure 1d shows the 1940 electric grid map, along with the 1935 and 1912 power plants. The maps of the high-voltage grid are digitized from the Edison Electrical Institute (1962); the map of the 1912 power plants is digitized from the U.S. Department of Commerce (1912); the map of the 1935 power plants is digitized from the U.S. Federal Power Commission (1935). A county c is considered as "treated" in decade t if at least one of the following conditions is met: (i) the county lies within a 50 km (ca. 30 miles) radius of a central power station, or (ii) the county is intersected by a 5 km (ca. 3 miles) buffer around a high-voltage transmission line. Appendix Figure B.4 shows the set of electrified counties by decade.

Figure 2: Event study results: Electrification and ethnic integration of the labor force in manufacturing industries

(a) Ethnic diversity of workers within manufacturing industry



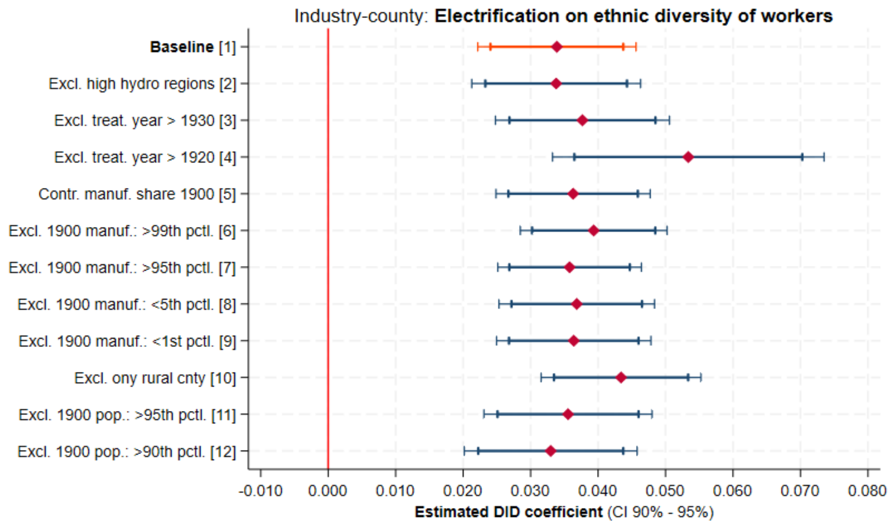
(b) Ethnic segregation of workers across occupations within manufacturing industry



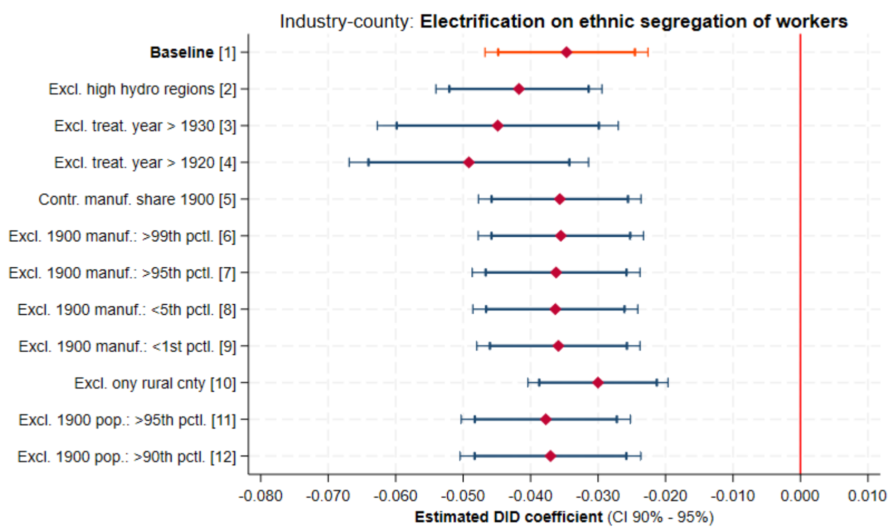
Note: This figure shows the event study coefficients estimated with the imputation method by Borusyak et al. (2024). Figure 2a shows the results for the ethnic diversity of workers within the manufacturing industry (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"). Figure 2b reports the results for the ethnic segregation of workers across occupations within the manufacturing industry (from 0 ="no segregation" to 1 = "complete segregation"). The unit of analysis is a manufacturing industry q in county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The y-axis reports the estimated coefficient and the 95% confidence interval of the difference between treated and control units (i.e., the average treatment effect on the treated, ATT). The horizontal dashed line indicates 0, i.e., no difference between the treated and control units. The vertical, dashed line indicates the occurrence of treatment (i.e., electrification). A county c "treated" in decade t if it is electrified. All industries located in an electrified county are treated. The x-axis shows the number of decades since treatment (event study periods): electrification happens in period 0, negative values indicate pre-treatment and positive values indicate post-treatment periods. The estimation method sets the estimated coefficient of the pre-treatment period farthest from the treatment to 0. All estimation models include: (i) industry $\times$ county FEs, (ii) state $\times$ decade FEs, (iii) industry $\times$ decade FEs, and (iv) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution. Estimates are weighted by the industry share of employment. Standard errors are clustered at the county level. Estimated coefficients are reported in Appendix Table **??** and the outcomes' trends in the data is shown in Appendix Figure C.5.

Figure 3: Summary of robustness checks: Electrification and ethnic integration of the labor force in manufacturing industries

(a) Ethnic diversity

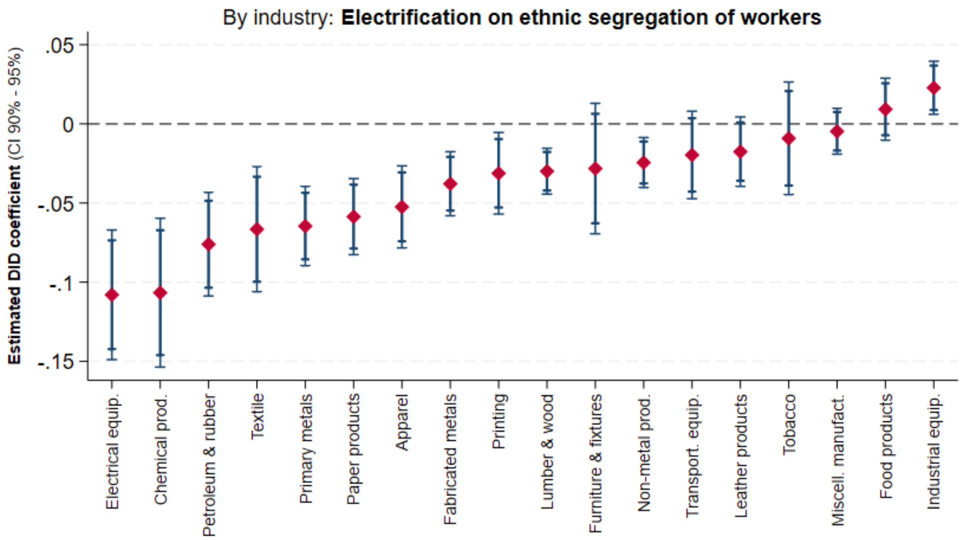


(b) Ethnic segregation



Note: This figure shows a summary of robustness checks for the industry-county-decade results regarding the effect of electrification on the ethnic integration of the labor force in manufacturing industries. The unit of analysis is a manufacturing industry q in county c and decade t. Counties are consistent over time, fixed at the 1900 borders. Figure 3a shows the results for the ethnic diversity of workers within the manufacturing industry (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"). Figure 3b reports the results for the ethnic segregation of workers across occupations within the manufacturing industry (from 0 = "no segregation" to 1 = "complete segregation"). Each reported coefficient is obtained by a separate DID estimation, performed with the imputation method by Borusyak et al. (2024). The y-axis specifies the performed robustness check. The x-axis reports the estimated DID coefficient and the 90% and 95% confidence interval of the difference between treated and control units (i.e., the average treatment effect on the treated, ATT), corresponding to the robustness check reported on the y-axis. The red vertical line indicates 0, i.e., no difference between the treated and control units. A county c "treated" in decade t if it is electrified. All industries located in an electrified county are treated. All estimation models include: (i) industry $\times$ county FEs, (ii) state $\times$ decade FEs, (iii) industry $\times$ decade FEs, and (iv) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution. Estimates are weighted by the industry share of employment. Standard errors are clustered at the county level. The corresponding estimated coefficients are reported in Appendix C.2, together with additional robustness checks.

Figure 4: Heterogenity by industry: Electrification and ethnic integration of the labor force in manufacturing industries



Note: This figure shows the effect of elctrification on the ethnic segregation of workrers across occupations within industry, separated by 2-digit SIC industries. The unit of analysis is a manufacturing industry q in county c and decade t. Each manufacturing industry q belongs to a broader industrial category defined by 2-digit SIC codes. Counties are consistent over time, fixed at the 1900 borders. Each reported coefficient is obtained by the DID estimation performed with the imputation method by Borusyak et al. (2024), allowing for heterogeneity across 2-digit SIC industries. The x-axis specifies the industry. The y-axis reports the estimated DID coefficient and the 90% and 95% confidence interval of the difference between treated and control units (i.e., the average treatment effect on the treated, ATT), corresponding to the industry reported on the x-axis. A county c "treated" in decade t if it is electrified. All industries located in an electrified county are treated. The estimation model include: (i) industry $\times$ county FEs, (ii) state $\times$ decade FEs, (iii) industry $\times$ decade FEs, and (iv) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution. Estimates are weighted by the industry share of employment. Standard errors are clustered at the county level.

Tables

Table 1: Electrified vs. Non-electrified: Check of counties characteristics

	County level analysis: Comparison of electrified in t vs. electrified in $t+1$							
	Log Popul. (1)	*Perc. Urban* (2)	*Perc. Liter.* (3)	*Perc. US-White* (4)	*Perc. Immigr.* (5)	*Perc. US-Black* (6)	*Ethn. Divers.* (7)	*Perc. Manuf.* (8)
Panel A: Level (standardized outcome variables)								
Electrified	0.190***	0.164***	0.020**	-0.023*	0.043***	0.009	0.065***	0.081***
	(0.015)	(0.016)	(0.010)	(0.014)	(0.009)	(0.012)	(0.013)	(0.012)
Observation N	19,283	19,283	19,283	19,283	19,283	19,283	19,283	19,283
Panel B: Growth Rate (standardized outcome variables)								
Electrified	-0.037***	-0.030	-0.010	0.010	0.008	-0.017	0.007	-0.030***
	(0.012)	(0.019)	(0.008)	(0.007)	(0.006)	(0.011)	(0.005)	(0.011)
Observation N	15,397	7,913	15,397	15,397	15,342	14,467	15,387	15,397
State × Decade FE	✓	✓	✓	✓	✓	✓	✓	✓
Decade-Pair FE	✓	✓	✓	✓	✓	✓	✓	✓
Cluster SE: County$_{1900}$	✓	✓	✓	✓	✓	✓	✓	✓

Note: This table reports the estimated relationship between electrification and various county characteristics. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. The goal is to **compare counties that became electrified in a given decade with those that became electrified in the following decade**. I create a panel dataset where each observation is a county–decade pair. I consider two groups of counties: those electrified in decade t and those electrified in decade $t+1$. I assign treatment = 1 to the counties electrifies in decade t. I repeat this procedure for each pair of consecutive decades: 1910–1920, 1920–1930, 1930–1940, and 1940–non-electrified, to assign treatment. I obtain 4 datasets and then stack all these datasets together. The regression analyis estimates the relationship between a series of county characteristics and the treatment indicator. Each coefficient in the table is the result of a separte regression. In Panel A, the outcome variables are reported in levels; each variable is standardized to have mean = 0 and standard deviation = 1. In Panel B, the outcome variables are the growth rate (percentage change) between $t-1$ and t; each variable is standardized to have mean = 0 and standard deviation = 1. All estimation models include (i) state × decade fixed effects, and (ii) period fixed effects (which indicate the consecutive pairs). The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Standard errors are clustered at the county level and are reported in parentheses below the coefficients. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Summary Statistics

	Num.	Mean	St. Dev.	Pctl. 1	Pctl. 5	Pctl. 95	Pctl. 99
Panel A: Industry-county-decade							
Employment Number$_{qct}$	356,442	87.96	907.37	0.017	1	256	1,394
Employment Share$_{qct}$	356,442	0.004	0.014	0.000	0.0001	0.014	0.055
Immigrant Share$_{qct}$	356,442	0.104	0.210	0	0	0.529	1
Native White Share$_{qct}$	356,442	0.812	0.276	0	0.027	1	1
Native Black Share$_{qct}$	356,442	0.084	0.216	0	0	0.625	1
Ethnic Diversity$_{qct}$	309,020	0.147	0.214	0	0	0.597	0.740
Ethnic Segregation$_{qct}$	178,453	0.202	0.228	0	0	0.680	0.905
Panel B: County-decade							
Total Population ('000)$_{ct}$	14,097	37.470	122.51	1.364	3.692	96.246	388.32
Immigrant Popul. Share$_{ct}$	14,097	0.068	0.085	0.0001	0.0005	0.245	0.353
Immigrant Groups Number$_{ct}$	14,097	27.222	9.884	1	3	36	40
Native Black Share$_{ct}$	14,097	0.120	0.196	0	0	0.582	0.767
Manufact. Workers Number$_{ct}$	14,097	2,224.1	12,258.5	3	16	7,675	35,763
Manufact. Sector Share$_{ct}$	14,097	0.094	0.100	0.003	0.008	0.312	0.440
Manufact. Industries Number$_{ct}$	14,097	25.285	13.225	2	6	49	68
Residential Segregation$_{ct}$	11,766	0.096	0.126	0.006	0.014	0.296	0.612
Foreignness Children Names$_{ct}$	8,614	45.394	12.698	7.195	22.545	64.723	77.118
Native-Foreign Intermarr. Rate$_{ct}$	13,768	0.060	0.119	0	0	0.301	0.500
Num. Public Service Occ. (/1,000)$_{ct}$	14,097	9.655	3.760	2.865	4.109	16.129	20.219
Num. Teachers (/1,000)$_{ct}$	14,097	7.615	3.202	1.852	2.879	13.201	16.611
Num. Doctors (/1,000)$_{ct}$	14,097	1.041	0.530	0.178	0.386	1.973	2.521
Num. Policemen (/1,000)$_{ct}$	14,097	0.438	0.441	0	0	1.241	2.047
Num. Firefighters (/1,000)$_{ct}$	14,097	0.104	0.225	0	0	0.610	1.034
Num. Public Admin. (/1,000)$_{ct}$	14,097	0.457	0.826	0	0	1.550	3.110

Note: This table shows the summary statistics of the key variables at the industry-county- decade level (Panel A) and at the county-decade level (Panel B).

Table 3: DID results: Electrification and ethnic composition of the labor force in manufacturing industries

	Industry-county-decade level analysis		
	Share Immigrant Workers (1)	*Share Native White Workers* (2)	*Share Native Black Workers* (3)
Electrified	0.016***	-0.016***	-0.0002
	(0.005)	(0.006)	(0.003)
Observation N	276,245	276,245	276,245
Cluster N	2,799	2,799	2,799
Outcome Mean_{1900}	0.188	0.742	0.070
Outcome $\text{Mean}_{\text{Non-treated}}$	0.110	0.797	0.093
Industry × County FE	✓	✓	✓
State × Decade FE	✓	✓	✓
Industry × Decade FE	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓
Pre-trend Test: F-stat	0.149	0.679	1.282
$\beta_{h<0} = 0$ [p-value]	[0.930]	[0.565]	[0.279]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in column (1), the share of immigrant (foreign-born) workers in the industry; in column (2), the share of White U.S.-born workers in the industry; and in column (3), the share of Black U.S.-born workers in the industry. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: DID results: Electrification and ethnic integration of the labor force in manufacturing industries

	Industry-county-decade level analysis					
	Ethnic Diversity		*Ethnic Segregation*			
	(1)	(2)	(3)	(4)	(5)	(6)
Electrified	0.034***	0.021***	-0.035***	-0.035***	-0.026***	-0.026***
	(0.006)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Observation N	235,062	235,062	122,264	122,264	89,246	89,246
Cluster N	2,799	2,799	2,710	2,710	2,554	2,554
Outcome Mean_{1900}	0.201	0.201	0.114	0.114	0.114	0.114
Outcome $\text{Mean}_{\text{Non-treated}}$	0.132	0.132	0.184	0.184	0.184	0.184
Industry × County FE	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓
Add. Controls: *Share White Natives*	✗	✓	✗	✓	✗	✓
Add. Controls: *Ethnic Diversity*			✗	✗	✓	✓
Pre-trend Test: F-stat	0.239	0.701	1.300	1.327	1.076	0.993
$\beta_{h<0} = 0$ [p-value]	[0.869]	[0.551]	[0.273]	[0.327]	[0.358]	[0.395]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in columns (1)-(2), the index of ethnic diversity of workers within the industry (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"); and in columns (3)-(6), the index of ethnic segregation of workers across occupations within the industry (from 0 = "no segregation" and 1 = "complete segregation"). The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. Columns (2) and (4)-(6) show the results for the same model as in column (1) and (3), respectively, with the addition of industry-county-decade controls, as specified in the table. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Heterogeneity by industry characteristics – DID Results: Electrification and ethnic integration of the labor force in manufacturing industries

	Industry-county-decade level analysis							
	Ethnic Segregation of Workers							
	Energy Use Intensity						Scope for Reorganization	
	Energy Expense$_{1900}$		Energy Use$_{1930}$		Electricity Use$_{1963}$		Establishment Size$_{1900}$	
	$\geq$ Med. (1)	$<$ Med. (2)	$\geq$ Med. (3)	$<$ Med. (4)	$\geq$ Med. (5)	$<$ Med. (6)	$\geq$ Med. (7)	$<$ Med. (8)
Electrified	-0.043***	-0.026***	-0.036***	-0.030***	-0.041***	-0.027***	-0.043***	-0.015**
	(0.008)	(0.006)	(0.007)	(0.007)	(0.008)	(0.006)	(0.007)	(0.007)
Observation N	66,635	55,629	82,112	40,152	71,579	50,685	60,836	61,428
Cluster N	2,576	2,665	2,672	2,572	2,588	2,660	2,625	2,628
Outcome Mean$_{1900}$	0.110	0.117	0.114	0.113	0.112	0.116	0.125	0.104
Outcome Mean$_{\text{Non-treat.}}$	0.197	0.167	0.187	0.177	0.198	0.164	0.185	0.183
Industry $\times$ County FE	✓	✓	✓	✓	✓	✓	✓	✓
State $\times$ Decade FE	✓	✓	✓	✓	✓	✓	✓	✓
Industry $\times$ Decade FE	✓	✓	✓	✓	✓	✓	✓	✓
County Popul.$_{1900}$ $\times$ $\mathbb{I}_t$	✓	✓	✓	✓	✓	✓	✓	✓
Clust. SE: County$_{1900}$	✓	✓	✓	✓	✓	✓	✓	✓
Pre-trend Test: F-stat	1.300	1.300	1.300	1.300	1.300	1.300	1.300	1.300
$\beta_{h<0} = 0$ [p-value]	[0.273]	[0.273]	[0.273]	[0.273]	[0.273]	[0.273]	[0.273]	[0.273]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the index of ethnic segregation of workers across occupations within the industry (from 0 = "no segregation" and 1 = "complete segregation"). The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. Results are reported by **industry characteristics**: fuel and energy expenses over wage bill in 1900 in columns 1–2 (from census of manufacturers); total horsepower per worker in 1930 in columns 3–4 (from census of manufacturers); share of electricity use in manufacturing in 1963 in columns 5–6 (from input-output tables); average establishment size (number of workers) in 1900 in columns 7–8 (from census of manufacturers). High/low is defined as above/below median, at the national and industry level in the year of reference. All estimation models include: (i) industry $\times$ county FEs, (ii) state $\times$ decade FEs, (iii) industry $\times$ decade FEs, and (iv) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: DID results: Electrification and residential segregation of the manufacturing labor force

	County-decade level analysis					
	Residential Segregation of Manufacturing Workers					
	Baseline Definition of Ethnicity			Incl. Second Generat. Immigr.		
	(1)	(2)	(3)	(4)	(5)	(6)
Electrified	-0.010***	-0.012***	-0.010***	-0.008**	-0.009***	-0.008**
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Observation N	11,036	11,036	11,036	11,036	11,036	11,036
Cluster N	2,574	2,574	2,574	2,574	2,574	2,574
Outcome Mean_{1900}	0.090	0.090	0.090	0.091	0.091	0.091
Outcome $\text{Mean}_{\text{Non-treated}}$	0.107	0.107	0.107	0.109	0.109	0.109
County FE	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Manufacture $\text{Share}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓
Add. Cntr: *Share White Natives*	✗	✓	✗	✗	✓	✗
Add. Cntr: *Share 2nd-Gener. Immigr.*				✗	✓	✗
Add. Cntr: *Ethnic Diversity*	✗	✗	✓	✗	✗	✓
Pre-trend Test: F-stat	0.588	0.524	0.579	0.331	0.218	0.348
$\beta_{h<0} = 0$ [p-value]	[0.623]	[0.666]	[0.629]	[0.803]	[0.884]	[0.791]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a county c in a decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the index of residential segregation of manufacturing workers across enumeration districts within the county (from 0 = "no segregation" and 1 = "complete segregation"). In columns (1)–(3), residential segregation is computed using the baseline definition of ethnic groups: U.S.-born Whites, U.S.-born Blacks, and foreign-born individuals classified by their country of birth. In columns (4)–(6), the measure is recalculated to include second-generation immigrants –U.S.-born individuals with foreign-born parents– grouped together with the foreign-born. Second-generation individuals are assigned to an ethnic group based on the father's country of birth. In this case, White and Black U.S.-natives are defined as individuals born in the U.S. to U.S.-born parents. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All estimation models include: (i) county FEs, (ii) state × decade FEs, (iii) the county population in 1900 × decade dummies as control, and (iv) the county share of manufacturing employment in 1900 × decade dummies as control. Columns (2) and (3) show the results of the same model as in column (1), additionally controlling for the county share of White U.S.-native and the county ethnic diversity, respectively. Columns (5) and (6) show the results of the same model as in column (4), additionally controlling for the county share of White U.S.-native plus share of U.S.-native with foreign-born parents and the county ethnic diversity, respectively. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the manufacturing share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Heterogeneity by county manufacturing characteristics – DID results: Electrification and residential segregation of the manufacturing labor force

	County-decade level analysis			
	Residential Segregation of Manufacturing Workers			
	Baseline Definition of Ethnicity		Incl. Second Generat. Immigr.	
	Manuf. Energy Intensity$_{1900}$		Manuf. Energy Intensity$_{1900}$	
	$\geq$ Median (1)	$<$ Median (2)	$\geq$ Median (3)	$<$ Median (4)
Electrified	-0.014***	-0.003	-0.012***	-0.001
	(0.004)	(0.005)	(0.004)	(0.005)
Observation N	5,803	5,233	5,803	5,233
Cluster N	1,306	1,268	1,306	1,268
Outcome Mean$_{1900}$	0.085	0.097	0.087	0.095
Outcome Mean$_{\text{Non-treated}}$	0.103	0.110	0.107	0.109
County FE	✓	✓	✓	✓
State $\times$ Decade FE	✓	✓	✓	✓
County Population$_{1900}$ $\times$ $\mathbb{I}_t$	✓	✓	✓	✓
Manufacture Share$_{1900}$ $\times$ $\mathbb{I}_t$	✓	✓	✓	✓
Clustered SE: County$_{1900}$	✓	✓	✓	✓
Pre-trend Test: F-stat	0.588	0.588	0.331	0.331
$\beta_{h<0} = 0$ [p-value]	[0.623]	[0.623]	[0.803]	[0.803]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a county c in a decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the index of residential segregation of manufacturing workers across enumeration districts within the county (from 0 = "no segregation" and 1 = "complete segregation"). In columns (1)–(3), residential segregation is computed using the baseline definition of ethnic groups: U.S.-born Whites, U.S.-born Blacks, and foreign-born individuals classified by their country of birth. In columns (4)–(6), the measure is recalculated to include second-generation immigrants –U.S.-born individuals with foreign-born parents– grouped together with the foreign-born. Second-generation individuals are assigned to an ethnic group based on the father's country of birth. In this case, White and Black U.S.-natives are defined as individuals born in the U.S. to U.S.-born parents. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. Results are reported by a **characteristic of the manufacturing sector at the county level**: the average energy intensity of the manufacturing industries in the county in 1900, calculated using the average energy expenses on wage bill at the industry level (from census of manufacturers). High/low is defined as above/below median, considering all counties in the year of reference. All estimation models include: (i) county FEs, (ii) state $\times$ decade FEs, (iii) the county population in 1900 $\times$ decade dummies as control, and (iv) the county share of manufacturing employment in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the manufacturing share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: OLS results: Electrification and provision of local public services

	County-decade level analysis				
	Workers in All Occupations Related to Public Service Provision				
		Manuf. Employment Share$_{1900}$		Manuf. Energy Intensity$_{1900}$	
	All Counties	$\geq$ Median	$<$ Median	$\geq$ Median	$<$ Median
	(1)	(2)	(3)	(4)	(5)
Electrified	-0.213***	-0.298***	0.080	-0.084	-0.302***
	(0.071)	(0.090)	(0.105)	(0.094)	(0.107)
Imm. Share	-9.987***	-12.061***	-1.580	-12.112***	-7.238***
	(1.198)	(1.373)	(1.710)	(1.236)	(2.040)
Imm. Share $\times$ *Electr.*	1.845***	2.709***	-3.106**	1.678**	1.878*
	(0.630)	(0.686)	(1.389)	(0.708)	(1.090)
Observation N	13,797	6,825	6,962	6,842	6,944
Cluster N	2,763	1,366	1,395	1,370	1,391
Outcome Mean$_{1900}$	7.766	7.990	7.546	8.358	7.184
Outcome Mean$_{\text{Non-treated}}$	9.033	8.636	9.311	9.734	8.403
County FE	✓	✓	✓	✓	✓
State $\times$ Decade FE	✓	✓	✓	✓	✓
County Popul.$_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓
Manuf. Share$_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓
Clust. SE: County$_{1900}$	✓	✓	✓	✓	✓

Note: This table reports the OLS coefficients of the regression model with an interaction term between the county share of immigrant population (*Perc. Immigrant*) and the indicator variable with value of 1 for the period after electrification (*Electrified*). The unit of analysis is a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the total number of workers employed in occupations related to the provision of local public services, per 1,000 inhabitants. The occupations are teachers, doctors, policemen, firefighters, and public administrators. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. The estimated of *Perc. Immigrant* indicates the relationship between the county share of immigrant population and employment in occupations related to the provision of public services, in non-electrified counties. The estimated coefficient of the interaction term *Perc. Immigrant* $\times$ *Electrified* indicates how this relationship differs in electrified counties relative to non-electrified counties. In column 1, the regression is run on the full sample of counties. In columns 2–5, results are reported by **county characteristics**: the share of employment in manufacturing in 1900 (calculated from individual full-count census) in columns 2–3; the energy intensity of the manufacturing industries in the county in 1900 (calculated from census of manufacturers). High/low is defined as above/below median, considering all counties in the year of reference. All estimation models include: (i) county FEs, (ii) state $\times$ decade FEs, (iii) the county population in 1900 $\times$ decade dummies as control, and (iv) the county share of employment in manufacturing in 1900 $\times$ decade dummies as control. Estimates are weighted by the manufacturing share of employment in the county. In all estimation models, the sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Standard errors are clustered at the county level and are reported in parentheses below the coefficients. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: DID results: Electrification and cultural assimilation of immigrant manufacturing workers

	County-decade level analysis					
	Households with at least one spouse in manufacturing					
	Native + foreign spouses			Both foreign-born spouses		
	Foreign-Native Intermarriage Rate			*Foreignness Children Names*		
		Manuf. Energy Intensity$_{1900}$			Manuf. Energy Intensity$_{1900}$	
	All Counties	$\geq$ Median	$<$ Median	All Counties	$\geq$ Median	$<$ Median
	(1)	(2)	(3)	(4)	(5)	(6)
Electrified	0.012*	0.013*	0.010	0.004	0.012	-0.012
	(0.007)	(0.006)	(0.010)	(0.007	(0.009)	(0.009)
Observation N	13,128	6,448	6,680	7,663	4,466	3,197
Cluster N	2,752	1,363	1,389	2,453	1,148	960
Outcome Mean$_{1900}$	0.107	0.123	0.090	0.460	0.462	0.458
Outcome Mean$_{\text{Non-treated}}$	0.061	0.078	0.047	0.456	0.455	0.458
County FE	✓	✓	✓	✓	✓	✓
State $\times$ Decade FE	✓	✓	✓	✓	✓	✓
County Popul.$_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Manuf. Share$_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clust. SE: County$_{1900}$	✓	✓	✓	✓	✓	✓
Pre-trend Test: F-stat	0.830	0.830	0.830	0.987	0.987	0.987
$\beta_{h<0} = 0$ [p-value]	[0.477]	[0.477]	[0.477]	[0.398]	[0.398]	[0.398]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a county c in a decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in Column (1), the foreign-native intermarriage rates, as the fraction of marriages with at least one U.S.-born spouse celebrated in the last decade; in Column (2), the average foreignness index of names given to children born in the U.S. in the last decade to foreign-born parents. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. In columns 1 and 4, the regression is run on the full sample of counties. In columns 2–3 and 5–6, results are reported by a **characteristic of the manufacturing sector at the county level**: the average energy intensity of the manufacturing industries in the county in 1900, calculated using the average energy expenses on wage bill at the industry level (from census of manufacturers). High/low is defined as above/below median, considering all counties in the year of reference. All estimation models include: (i) county FEs, (ii) state $\times$ decade FEs, (iii) the county population in 1900 $\times$ decade dummies as control, and (iv) the county share of manufacturing employment in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the manufacturing share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix

A. Background Appendix

A.1 Historical migration to the US

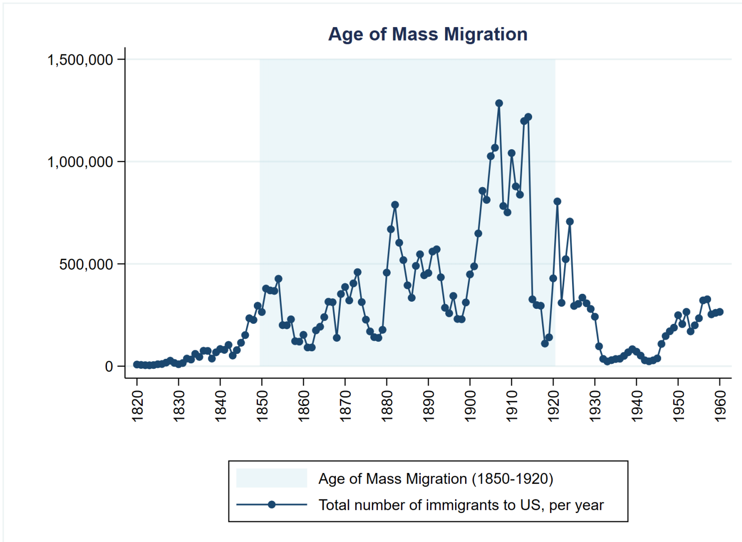
Table A.1: Foreign-born population in the United States, 1900–1940

Year	Total Population	Foreign-born Population	Share foreign Population	Total Workforce	Foreign-born Workforce	Share foreign Workforce
1900	75,944,591	10,526,240	13.86%	26,003,049	5,238,331	20.15%
1910	92,082,573	13,674,116	14.85%	33,162,519	6,794,446	20.49%
1920	105,663,051	14,053,239	13.30%	34,044,286	6,382,305	18.75%
1930	122,667,091	14,410,414	11.75%	41,310,260	6,632,089	16.05%
1940	131,851,589	11,824,993	8.97%	53,109,472	6,153,237	11.59%

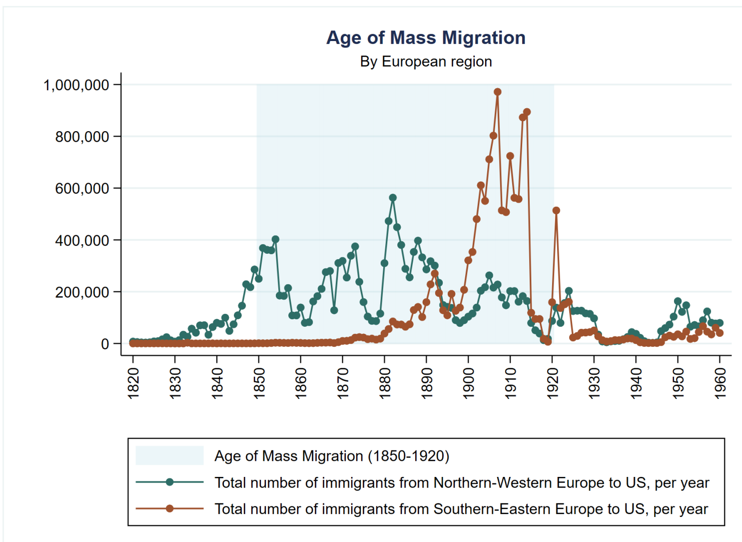
Note: This table reports the total number of U.S. population, foreing-born population, and the share of foreign-born in the population, between 1900 and 1940 (columns 2–4). It also reports the total number of U.S. workforce, foreing-born workforce, and the share of foreign-born in the workforce, between 1900 and 1940 (columns 5–7). Data are calculated from the individual U.S. full-count census (Ruggles et al., 2024).

Figure A.1: Number of immigrants to the United States, 1820–1960

(a) Total number of immigrants per year



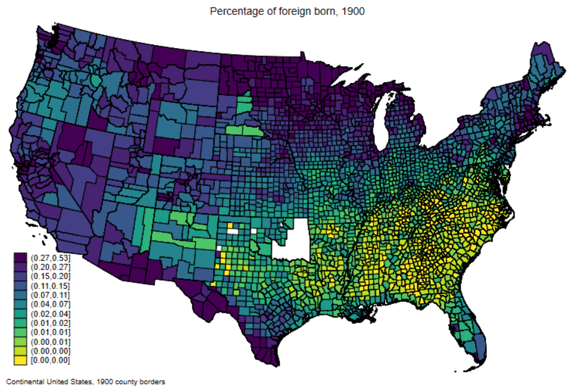
(b) Number of European immigrant per year



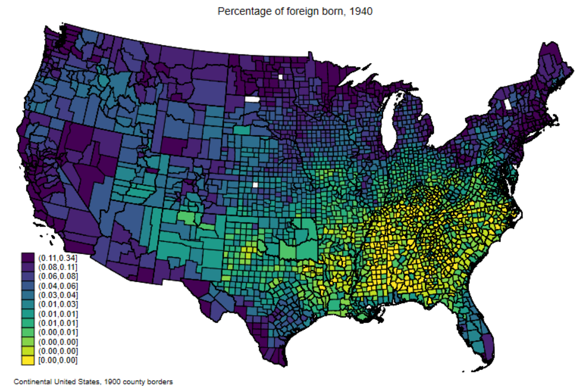
Note: Figures elaborated from the Historical Statistics of the United States, Millennial Edition. Figure A.1a reports the total number of immigrants in the US, per year, between 1820 and 1960. It is obtained by aggregating data from Table Series Ad106-120, Ad136-148, Ad162-172, Ad191-195, and Ad206-213. During the Age of Mass Migration (1850-1920), the average number of immigrants entering the US per year was 441,660.44. Figure A.1b reports the number of immigrants from Europe in the US, per year, between 1820 and 1960. It is obtained from Table Series Ad106-120. Immigrants from Northern and Western Europe are the sum of immigrants from Great Britain, Ireland, Scandinavia, Germany, and other North-Western European countries. Immigrants from Southern and Eastern Europe are the sum of immigrants from Poland, Russia (Soviet Union and Baltic states), Greece, Italy, Portugal, Spain, and other Central, Eastern and Southern European countries.

Figure A.2: County-level immigrant population, 1900–1940

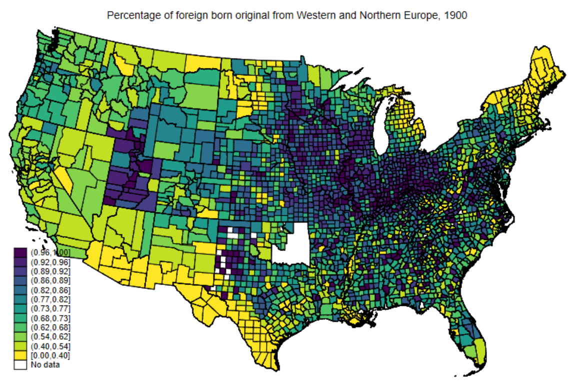
(a) Immigrants as percentage of county population, 1900



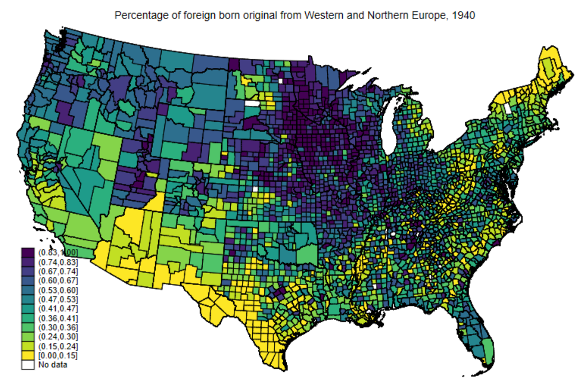
(b) Immigrants as percentage of county population, 1940



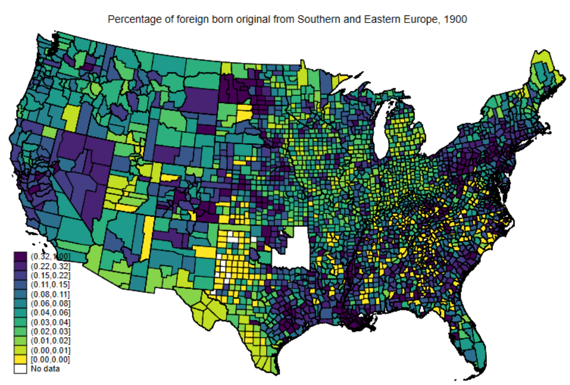
(c) Percentage of "Old" stock immigrants (Northern and Western Europe), 1900



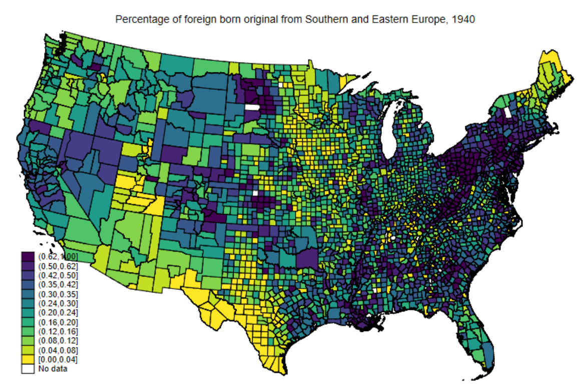
(d) Percentage of "Old" stock immigrants (Northern and Western Europe), 1940



(e) Percentage of "New" stock immigrants (Southern and Eastern Europe), 1900



(f) Percentage of "New" stock immigrants (Southern and Eastern Europe), 1940

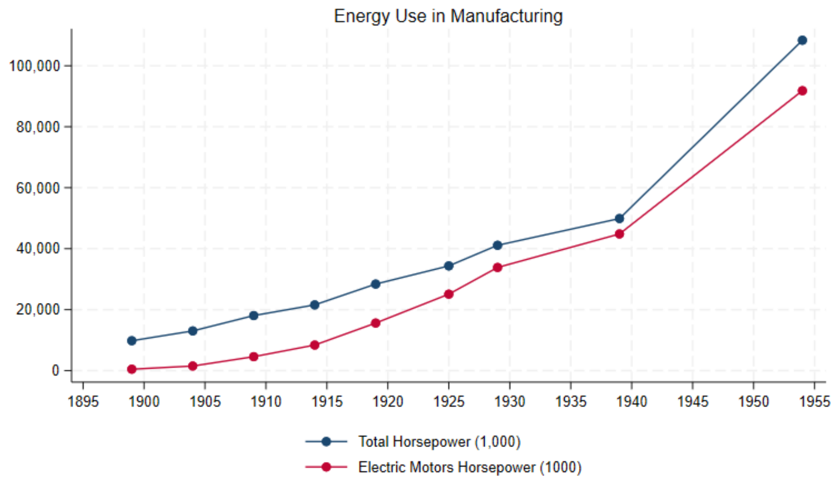


Note: Figures A.2a and A.2b display the share of foreign-born individuals in the total county population in 1900 and 1940, respectively. Figures A.2c and A.2d show the share of "old stock immigrants" (from Northern and Western Europe) among all immigrants at the county level in 1900 and 1940. Figures A.2e and A.2f present the corresponding share of "new stock immigrants" (from Southern and Eastern Europe) among all immigrants at the county level in 1900 and 1940.

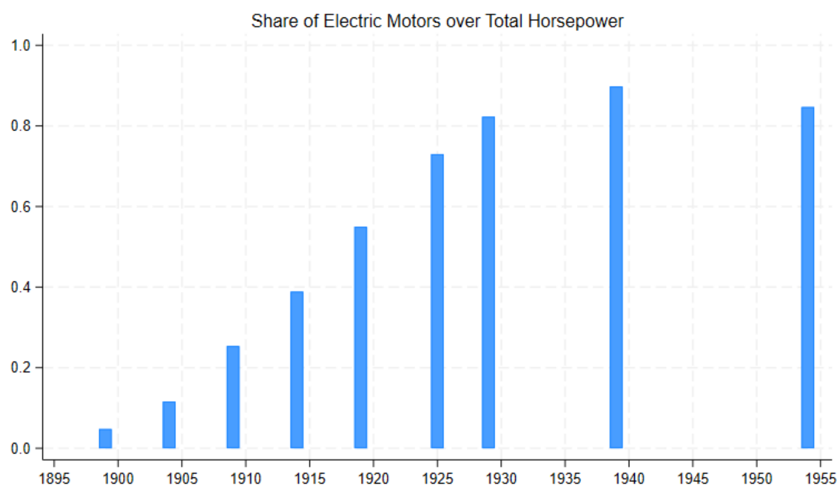
A.2 Electrification of U.S. manufacturing

Figure A.3: Electricity adoption in manufacturing (1899–1954)

(a) Use of energy in manufacturing sector



(b) Share of electric motors over total horsepower in manufacturing

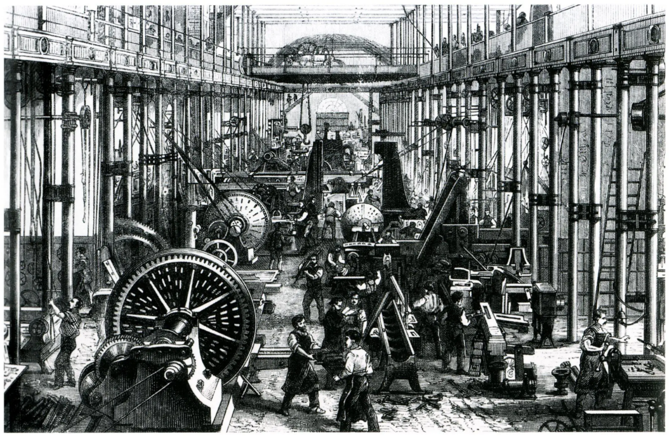


Data Source: U.S. Bureau of the Census, *U.S. Census of Manufactures, 1954*, vol. I; Washington, D.C., 1957.

A.3 Historical pictures of manufacturing industries

Figure A.4: Factory outline under different sources of power

(a) Steam-powered factory outline (ca. 1870)



(b) Electricity-powered factory outline (ca. 1920)

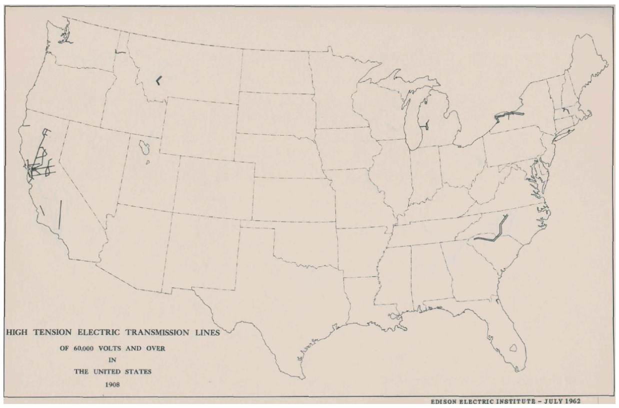
Note: Figure A.4a shows the typical "shafts-and-belts" outline of a manufacturing factory powered by an in-plant steam engine (Source: Wikipedia, Hartmann Maschinenhalle). Production machines were connected by a direct mechanical link to the power source that drove them, usually a single centrally located prime mover (water wheel or steam engine), which turned metal line shafts suspended from the ceiling and covering the entire length of each factory floor. The line of shafts turned, via pulleys and belts, shorter shafts parallel to the line shaft, to which production machinery was belted. Production machinery was necessarily arranged based on the shaft structure, following the logic of placing the machines as close as possible to the source of power, to limit power losses due to line friction. The entire network of shafts rotates continuously, regardless of the actual number of machines being used (Devine Jr, 1983). Figure A.4b shows the typical outline of an "electric unit drive" outline of a manufacturing factory powered by electricity sourced externally from the central power plant (Source: Missouri History Museum, Photographs and Prints Collections, Emerson Electric Manufacturing Company). As central power plants and the electrical distribution network expanded, manufacturing facilities no longer needed to operate an in-house prime mover. This allowed for a gradual change in the organization of the production process and the factory outline, first adopting electric group drive (a series of smaller engines, each connected to a group of production machinery) and later switching to the electric unit drive, with a single motor mounted right on the machine being driven. The new system eliminated power losses due to friction in shafts and belts; each machine could be operated independently of the others, optimizing energy intensity and use according to the specific production process. Factory buildings could be constructed more lightly and at lower cost, as they no longer needed to support the heavy systems of metal lines and shafts. Space on the factory floor could be optimized by placing machines closer together, since the need for workers to move between them was reduced. Additionally, machinery could be arranged according to the logic of the production process, rather than based on proximity to the power source (Hounshell, 1984).

B. Data Appendix

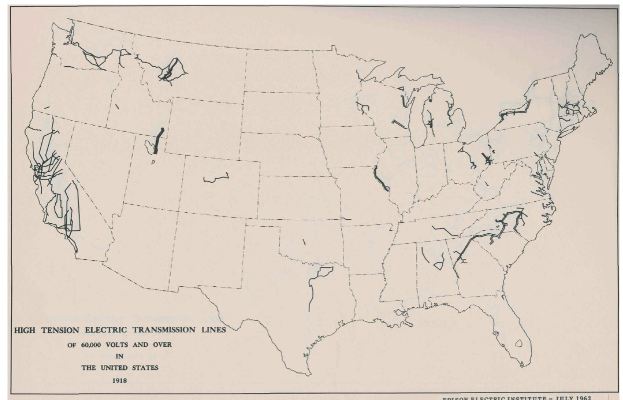
B.1 U.S. Electrification, 1900–1940

Figure B.1: Original maps: Electric grid and location of major electricity plants (1908–1946)

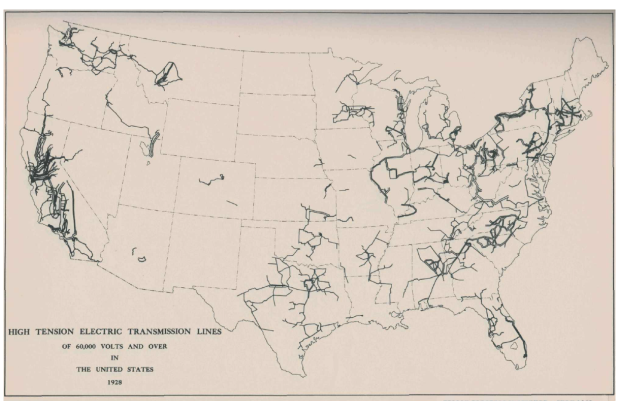
(a) High-voltage grid, 1908



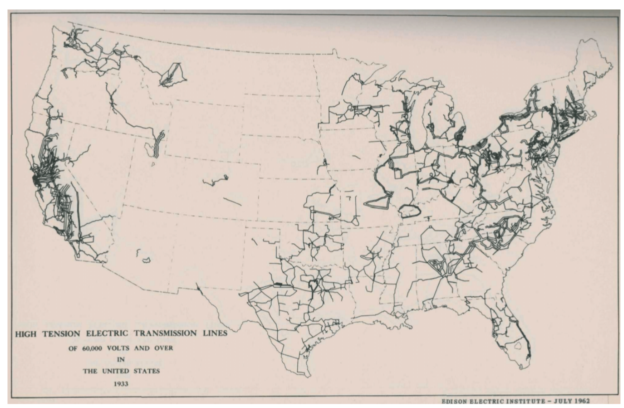
(b) High-voltage grid, 1918



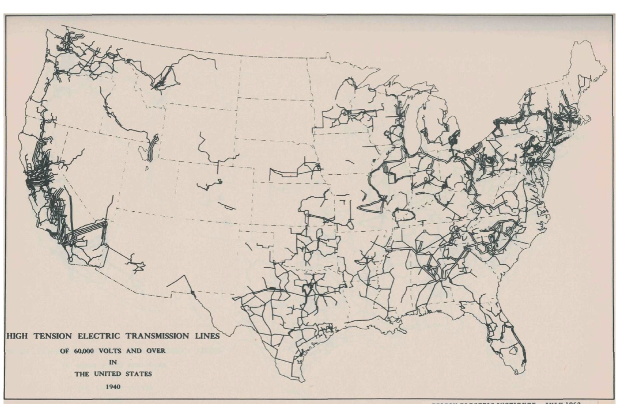
(c) High-voltage grid, 1928



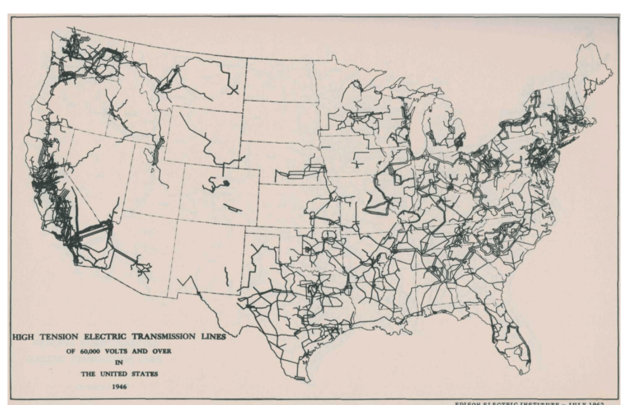
(d) High-voltage grid, 1933



(e) High-voltage grid, 1940



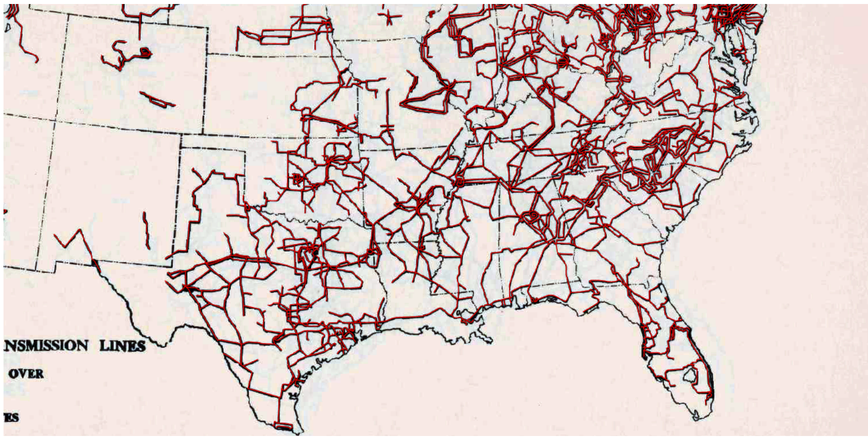
(f) High-voltage grid, 1946



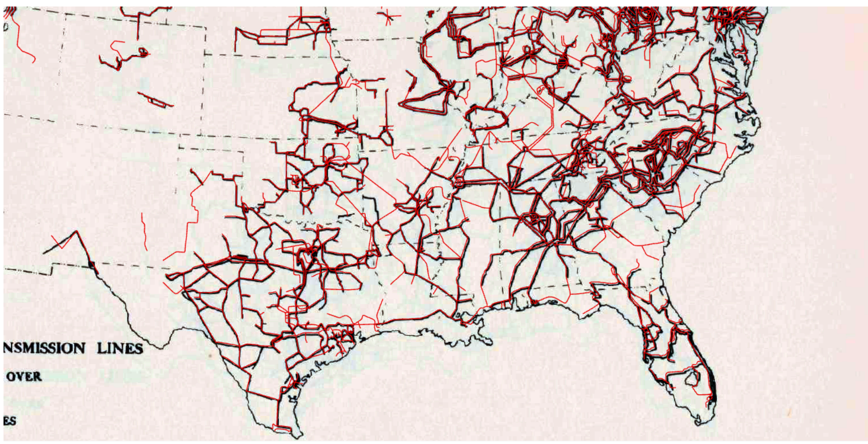
Note: This figure reports the original maps that display the expansion of the high-voltage grid in the United States between 1908 and 1946. Source: Edison Electrical Institute, 1962.

Figure B.2: Example of the digitalization process of the electric grid

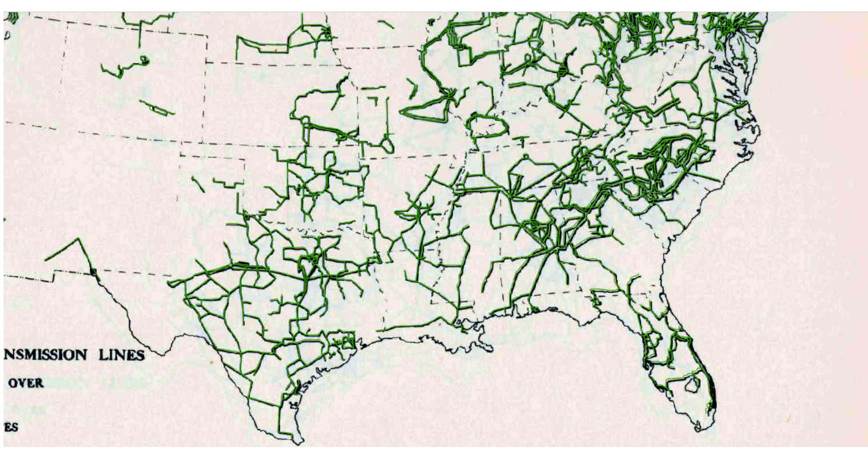
(a) 1946 electric grid shapefile overlaps 1946 grid map



(b) 1946 electric grid shapefile overlaps 1940 grid map



(c) 1940 electric grid shapefile overlaps 1940 grid map

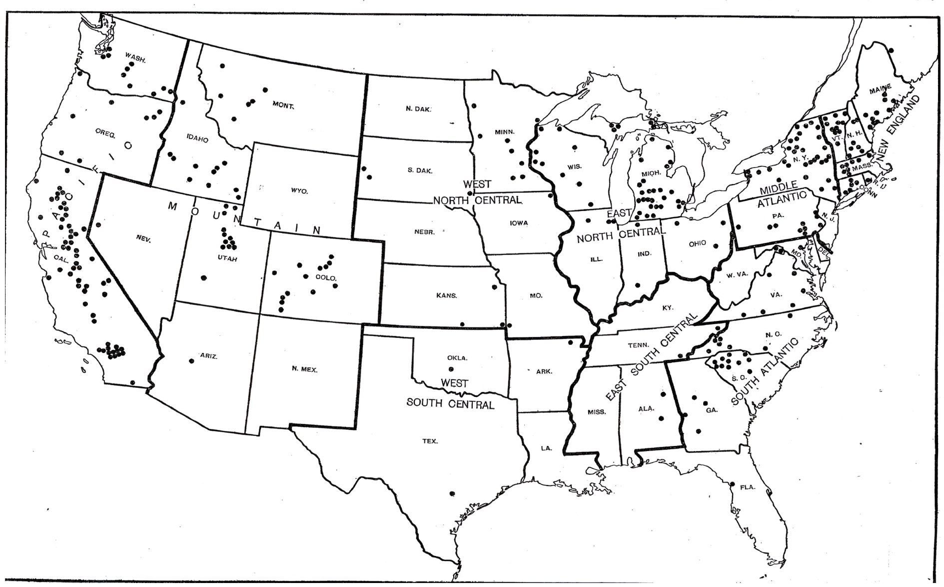


Note: Figure B.2a shows the shapefile of the high-voltage electric grid in 1946 (red line), digitized from the 1946 map by the Edison Electrical Institute (EEI, 1962), overlapped on the 1946 georeferenced map. This is the starting point of the mapping process. Figure B.2b shows the shapefile of the high-voltage electric grid in 1946 (red line), overlapped on the 1940 georeferenced map from the EEI Report. I manually delete the grid segments that are part of the 1946 grid shapefile but do not appear on the 1940 grid map, to obtain the shapefile of the high-voltage electric grid in 1940. Figure B.2c shows the shapefile of the high-voltage electric grid in 1940 (green line), obtained from the manual deletion of the segments, overlapped on the 1940 georeferenced map from the EEI Report.

Figure B.3: Original maps: Location of major electricity plants (1912, 1935)

(a) Major electricity plants, 1912

MAP 2.—LOCATION OF HYDROELECTRIC CENTRAL STATIONS REPORTING WATER POWER OF 1,000 HORSEPOWER OR MORE, BY GEOGRAPHIC DIVISIONS: 1912.



(b) Major electricity plants, 1935

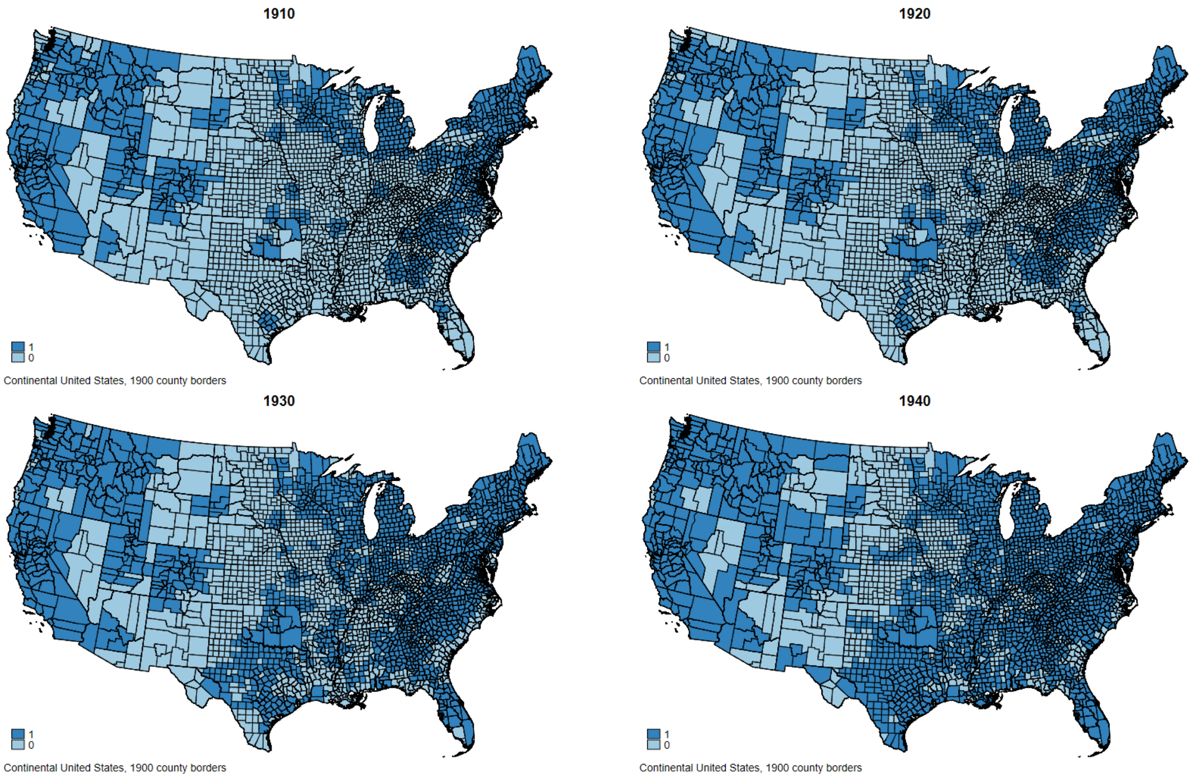
LOCATION OF ELECTRIC GENERATING STATIONS
WITH
ANNUAL OUTPUT OF MORE THAN 100 MILLION KILOWATT HOURS
(PLANTS INDICATED PRODUCE 73% OF THE ENERGY REQUIREMENTS OF THE NATION)

NATIONAL POWER SURVEY

FEDERAL POWER COMMISSION

Note: Figure B.3a shows the location of hydroelectric central stations reporting waterpower of at least 1,000 horsepower in 1912. Source: U.S. Department of Commerce (1912, p. 37). Figure B.3b shows the location of electric generating stations with annual output exceeding 100 million kilowatt hours in 1935. Source: .S. Federal Power Commission (1935, p. 5).

Figure B.4: Electrified counties, 1910–1940



Note: This figure reports in dark blue counties that are electrified in each decade. A county is considered as electrified if at least one of these two conditions occurs: (i) the county lies within a 50 km (ca. 30 miles) radius of a central power station, or (ii) the county is intersected by a 5 km (ca. 3 miles) buffer around a high-voltage transmission line. Figure 1 shows the decennial expansion of the high-voltage electric grid and the location of major electric power plants.

B.2 Categorization of variables from the full-count individual census

Country of Birth. My analysis covers the period 1900–1940, a time of substantial geopolitical change in Europe, with shifting national borders due to the two World Wars and related events. To ensure consistency over time, I apply standard consolidations to the country-of-birth information reported in the census, following conventions established in the literature. I closely follow the classification logic embedded in the IPUMS "birthplace" (BPL) variable, which reflects the original strings recorded on historical census schedules and is coded by IPUMS researchers.

Several consolidations are applied. England, Scotland, and Wales are grouped under the United Kingdom, while Northern Ireland is merged with Ireland. Prussia, when specified, is included under Germany, along with the Länder that later formed East and West Germany. Austria and Hungary are treated as separate entities, but entries for Austria-Hungary, when present, are assigned to Austria. Following IPUMS, I define Poland based on the BPL category 45500, which includes territories such as Austrian Poland, German Poland, East and West Prussia, and Russian Poland.[1] Estonia, Latvia, and Lithuania are grouped as the Baltic States, while Balkan countries (excluding Greece) are consolidated under Yugoslavia. Russia is defined using the IPUMS BPL category 46500, encompassing both the Russian Empire and later the U.S.S.R.[2] Smaller countries (e.g., Albania, Iceland, Luxembourg, and European microstates), as well as generic regional entries (e.g., "Western Europe"), are grouped under a residual "Other European" category.

While the analysis includes all countries of origin reported in the census, most immigrants during this period came from Europe. For non-European regions, I use broader groupings to avoid overly fragmented categories with limited representation. In the Americas, Canada and Mexico are treated as individual countries of birth due to their relative prominence, while the Caribbean, Central America, and South America are grouped into regional categories. In Asia, Japan and China are considered individually, given their importance in U.S. migration patterns and policies during and after the Age of Mass Migration.[3] The remaining Asian regions are grouped into Eastern, South-Eastern, and Middle Eastern Asia. For the Pacific, I separate Australia and New Zealand from a broader category for Pacific Islands. Finally, Africa is divided into Northern and Sub-Saharan regions. All individuals with unspecified or unclassified birthplaces outside Europe are grouped into a residual "Other" category.

Industry. The IPUMS variable IND1950 includes 148 categories covering all economic activities (from agriculture to public administration), as well as 14 categories for individuals outside the labor force (e.g., housework at home, students, retirees, and non-industrial responses). I exclude the latter from the analysis, focusing only on labor force participants.

For comparability and alignment with standardized classifications, I construct a crosswalk between the IND1950 codes and the U.S. Department of Labor's Standard Industrial Classification (SIC) Manual. The crosswalk aims to retain as much detail from the IPUMS classification as possible while linking it to the appropriate SIC codes. This results in variation in the granularity of the SIC codes used: 84 IPUMS categories (56.8%) map to 3-digit SIC codes, 56 (37.8%) to 2-digit codes, and 3 to 4-digit codes (all in retail trade). Five categories related to public administration

[1] In particular, the IPUMS BPL 45500 category includes Poland, Austrian Poland, Galicia, German Poland, East Prussia, Pomerania, Posen, Prussian Poland, Silesia, West Prussia, and Russian Poland.

[2] In particular, the IPUMS BPL 46500 category includes Russia, USSR, Byelorussia, Moldavia, Bessarabia, Ukraine, Armenia, Azerbaijan, Republic of Georgia, Kazakhstan, Kirghizia, Tadzhikistan, Turkmenistan, Uzbekistan, and Siberia.

[3] For example, see Trevor (1925), E. Lee (2003), and Long et al. (2024).

are excluded from SIC matching, as this sector is not covered by the SIC Manual.[4]

In some cases, IPUMS categories are broader than their SIC counterparts. For example, the IPUMS category "Apparel and Accessories" (IND1950 448) encompasses several 3-digit SIC codes ranging from 231 to 238.[5] By contrast, the SIC code 239 ("Miscellaneous Fabricated Textile Products") corresponds directly to a separate IPUMS category (IND1950 449).

In other instances, the classification logic between the two systems differs, requiring aggregation to establish correspondence. For example, IPUMS classification has three categories for fabricated metal products, distinguished by type of metal (IND1950 346 "Fabricated Steel Products", 347 "Fabricated Nonferrous Metal Products", and 348 "Not Specified Metal Industries"). These three IPUMS categories can all be linked to SIC code 34 ("Fabricated Metal Products, Except Machinery and Transportation Equipment"), but not to more detailed SIC subcategories, since SIC organizes these industries by product type.[6]

This harmonization process yields a final classification of 135 industry categories, including 49 in the manufacturing sector.

Occupation. The IPUMS variable OCC1950 includes 269 occupational categories covering jobs and professions across all economic activities, along with 14 categories for individuals outside the labor force (e.g., housekeeping at home, students, retirees, the unemployed, and inmates). I exclude the latter from the analysis, focusing only on individuals in the labor force.

To harmonize occupational classifications, I compare OCC1950 with the 1958 International Standard Classification of Occupations (ISCO-58) developed by the International Labour Organization.[7] I construct a crosswalk between the two systems to preserve the detail of the IPUMS occupational classification while aligning with the ISCO principle of minimizing dependence on industry-specific context.

In some cases, a one-to-one correspondence exists. For example, "architects" are recorded in both OCC1950 (category 3) and ISCO-58 (unit group 1, under minor group 0 "Architects, Engineers and Surveyors"). In most cases, however, multiple OCC1950 categories map to a single ISCO category. For instance, OCC1950 categories 42–49 identify types of engineers by industry (e.g., chemical, civil, metallurgical),[8] while ISCO-58 assigns all to a single category (unit group 2, under minor group 0). In such cases, I follow ISCO's logic to maintain the occupation classification as separate as possible from the industry, and group the corresponding OCC1950 categories into a consolidated category (such as, "engineers").

This harmonization results in 108 occupation categories. Some are specific to particular manufacturing industries (e.g., spinner in textiles, pressman in printing), while others are common across sectors (e.g., bookkeeper, engineer).

---

[4]For robustness, I also map each IPUMS category to its corresponding 2-digit SIC code.

[5]IPUMS categories: 231 "Men's And Boys' Suits, Coats, And Overcoats"; 232 "Men's And Boys' Furnishings, Work Clothing, And Allied Garments"; 233 " Women's, Misses', And Juniors' Outerwear"; 234 "Women's, Misses', Children's, And Infants'", 235 "Hats, Caps, And Millinery", 236 "Girls', Children's, And Infants' Outerwear"; 237 "Fur Goods"; 238 "Miscellaneous Apparel And Accessories".

[6]Examples of 3-digit SIC codes within group 34 include: 341 "Metal Cans and Shipping Containers," 342 "Cutlery, Handtools, and General Hardware," 343 "Heating Equipment," 344 "Structural Metal Products," and others.

[7]When no direct correspondence exists with ISCO-58, I consult more recent ISCO versions (1968, 1988, and 2008) available from the same source.

[8]OCC1950 categories: 42 "Engineers, chemical"; 43 "Engineers, civil"; 44 "Engineers, electrical"; 45 "Engineers, industrial"; 46 "Engineers, mechanical"; 47 "Engineers, metallurgical"; 48 "Engineers, minning"; 49 "Engineers, not elsewhere classified".

B.3 Summary statistics

Table B.1: Number of counties by electrification decade

	Total:			Decade of electrification:			
	All	Electr.	Not electr.	1910	1920	1930	1940
Panel A: Continental United States							
Number of Counties	2,843	2,092	751	1,011	114	542	425
Percentage		73.58%	26.42%	35.56%	4.01%	19.06%	14.95%
Panel B: New England							
Number of Counties	67	64	3	64			
Percentage		95.52%	4.48%	95.52%			
Panel C: Middle Atlantic							
Number of Counties	149	144	5	107	3	22	12
Percentage		96.64%	3.36%	71.81%	2.01%	14.77%	8.05%
Panel D: South Atlantic							
Number of Counties	519	418	101	246	35	92	45
Percentage		80.54%	19.46%	47.4%	6.74%	17.73%	8.67%
Panel E: East North Central							
Number of Counties	435	399	36	201	24	128	46
Percentage		91.72%	8.28%	46.21%	5.52%	29.43%	10.58%
Panel F: East South Central							
Number of Counties	356	260	96	39	16	96	109
Percentage		73.03%	26.97%	10.96%	4.49%	26.97%	30.62%
Panel G: West North Central							
Number of Counties	587	243	344	112	10	37	84
Percentage		41.40%	58.60%	19.08%	1.70%	6.30%	14.31%
Panel H: West South Central							
Number of Counties	411	314	97	30	20	154	110
Percentage		76.40%	23.60%	7.30%	4.87%	37.47%	26.76%
Panel I: Mountain							
Number of Counties	189	130	59	110	2	4	14
Percentage		68.78%	31.22%	58.20%	1.06%	2.12%	7.41%
Panel J: Pacific							
Number of Counties	126	120	6	102	4	9	5
Percentage		95.24%	4.76%	80.95%	3.17%	7.14%	3.97%

Note: This table reports the number of counties that get electrified between 1910 and 1940 and the corresponding percentage. Panel A reports the total numbers for Continental US. Panels B–J report the numbers by U.S. Census Division (source: U.S. Census). **New England**: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont. **Middle Atlantic**: New Jersey, New York, Pennsylvania. **East North Central**: Illinois, Indiana, Michigan, Ohio, Wisconsin. **West North Central**: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota. **South Atlantic**: Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia, District of Columbia. **East South Central**: Alabama, Kentucky, Mississippi, Tennessee. **West South Central**: Arkansas, Louisiana, Oklahoma, Texas. **Mountain**: Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming. **Pacific**: California, Oregon, Washington.

Table B.2: Correlation Matrix: Industry-County-Decade Variables

		(1)	(2)	(3)	(4)	(5)	(6)	
(1)	$\textit{Employment Number}_{qct}$	1						
(2)	$\textit{Employment Share}_{qct}$	0.347	1					
(3)	$\textit{Immigrant Share}_{qct}$	0.027	-0.063	1				
(4)	$\textit{Native White Share}_{qct}$	-0.030	0.003	-0.827	1			
(5)	$\textit{Native Black Share}_{qct}$	0.008	0.098	-0.211	-0.376	1		
(6)	$\textit{Ethnic Diversity}_{qct}$	0.069	-0.004	0.762	-0.816	0.162	1	
(7)	$\textit{Ethnic Segregation}_{qct}$	-0.089	-0.142	-0.058	0.035	0.034	-0.088	1

Note: This table shows the correlation coefficient between the key variables at the industry-county-decade level.

Table B.3: Correlation Matrix: County-Decade Variables

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1)	*Total Population ('000)$_{ct}$*	1							
(2)	*Immigrant Popul. Share$_{ct}$*	0.168	1						
(3)	*Immigrant Groups Number$_{ct}$*	0.347	0.334	1					
(4)	*Native Black Share$_{ct}$*	-0.013	-0.407	-0.221	1				
(5)	*Manufact. Workers Number$_{ct}$*	0.927	0.165	0.316	-0.026	1			
(6)	*Manufact. Sector Share$_{ct}$*	0.245	0.089	0.391	-0.019	0.326	1		
(7)	*Manufact. Industries Number$_{ct}$*	0.271	-0.105	0.449	-0.027	0.214	0.351	1	
(8)	*Residential Segregation$_{ct}$*	-0.061	0.037	0.004	-0.072	-0.100	-0.329	0.098	1
(9)	*Foreignness Children Names$_{ct}$*	0.073	0.068	0.117	0.027	0.059	0.082	0.006	-0.013
(10)	*Native-Foreign Intermarr. Rate$_{ct}$*	0.102	0.712	0.173	-0.259	0.097	0.068	-0.096	0.028
(11)	*Num. Public Service Occ. (/1,000)$_{ct}$*	0.033	-0.017	0.239	-0.370	0.046	-0.021	0.230	0.019
(12)	*Num. Teachers (/1,000)$_{ct}$*	-0.084	-0.003	0.098	-0.384	-0.068	-0.124	0.092	0.115
(13)	*Num. Doctors (/1,000)$_{ct}$*	0.144	-0.018	0.110	-0.127	0.115	-0.031	0.058	-0.131
(14)	*Num. Policemen (/1,000)$_{ct}$*	0.395	0.087	0.492	-0.051	0.387	0.330	0.389	-0.170
(15)	*Num. Firefighters (/1,000)$_{ct}$*	0.357	0.021	0.535	0.028	0.356	0.422	0.410	-0.215
(16)	*Num. Public Admin. (/1,000)$_{ct}$*	-0.008	-0.109	0.114	-0.073	0.012	0.044	0.242	-0.076
		(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(9)	*Foreignness Children Names$_{ct}$*	1							
(10)	*Native-Foreign Intermarr. Rate$_{ct}$*	0.061	1						
(11)	*Num. Public Good Occ. (/1,000)$_{ct}$*	-0.150	-0.074	1					
(12)	*Num. Teachers (/1,000)$_{ct}$*	-0.168	-0.059	0.914	1				
(13)	*Num. Doctors (/1,000)$_{ct}$*	0.047	0.007	0.111	-0.117	1			
(14)	*Num. Policemen (/1,000)$_{ct}$*	0.022	0.021	0.421	0.144	0.140	1		
(15)	*Num. Firefighters (/1,000)$_{ct}$*	0.052	-0.020	0.286	0.022	0.203	0.630	1	
(16)	*Num. Public Admin. (/1,000)$_{ct}$*	-0.092	-0.112	0.520	0.241	0.061	0.374	0.286	1

Note: This table shows the correlation coefficient between the key variables at the county-decade level.

Table B.4: Correlation Matrix: Industry Characteristics (2-digit SIC codes)

Panel A: Variable Levels		(1)	(2)	(3)	(4)
(1)	*Energy Expense*$_{1900}$	1			
(2)	*Energy Use*$_{1930}$	0.647	1		
(3)	*Electricity Use*$_{1963}$	0.889	0.612	1	
(4)	*Establishment Size*$_{1900}$	0.759	0.521	0.681	1

Panel B: Variable Dummies		(1)	(2)	(3)	(4)
(1)	*High Energy Expense*$_{1900}$	1			
(2)	*High Energy Use*$_{1930}$	0.478	1		
(3)	*High Electricity Use*$_{1963}$	0.789	0.478	1	
(4)	*High Establishment Size*$_{1900}$	0.578	0.689	0.578	1

Note: This table shows the correlation coefficient between the key industry characteristics used in the heterogeneity analysis of the effect of electrification on the ethnic segregation of workers within industries. Panel A shows the correlation between the variables expressed in levels. Panel B shows the correlation between the dummy variables that equal one if the industry level is above the respective median and zero if it below.

C. Result Appendix

C.1 Additional Results: County Level Checks

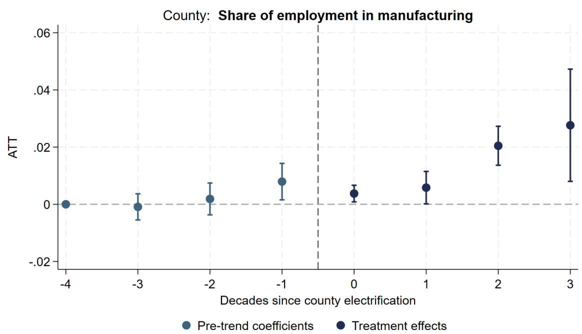
Table C.1: DID Results: Electrification and county share of manufacturing employment

	County-decade level analysis			
	Share of Employment in Manufacturing			
	Baseline		Treatment anticipated one decade	
	(1)	(2)	(3)	(4)
Electrified	0.011***	0.011***	0.011***	0.011***
	(0.003)	(0.003)	(0.002)	(0.002)
Observation N	13,576	13,484	8,927	8,835
Cluster N	2,799	2,773	1,811	1,785
Outcome Mean_{1900}	0.062	0.062	0.062	0.062
Outcome $\text{Mean}_{\text{Non-treated}}$	0.064	0.064	0.064	0.064
County FE	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓
Manufacture $\text{Share}_{1900} \times \mathbb{I}_t$	✓	✗	✓	✗
Clustered SE: County_{1900}	✓	✓	✓	✓
Pre-trend Test: F-stat	6.232	5.737	1.049	1.007
$\beta_{h<0} = 0$ [p-value]	[0.000]	[0.001]	[0.350]	[0.366]

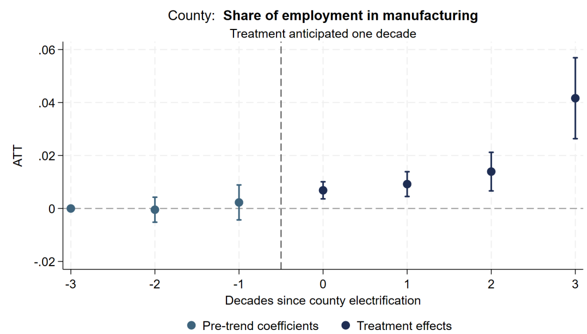
Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (ATT). The unit of analysis is a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the share of county employment in the manufacturing sector. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Columns (1)-(2) report the estimates using the first decade of electrification as treatment (**baseline**). Columns (3)-(4) report the estimates using the **decade before electrification** as treatment. All estimation models include: (i) county FEs, (ii) state × decade FEs, and (iii) the county population in 1900 × decade dummies as control. In columns (2) and (4), the model also includes the county manufacturing share of employment in 1900 × decade dummies as control. In all estimation models, the sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, with $h = 0$ start of treatment), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation (F-statistics and p-value in square brackets). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure C.1: Event-Study Results: Electrification and county share of manufacturing employment

(a) Event study



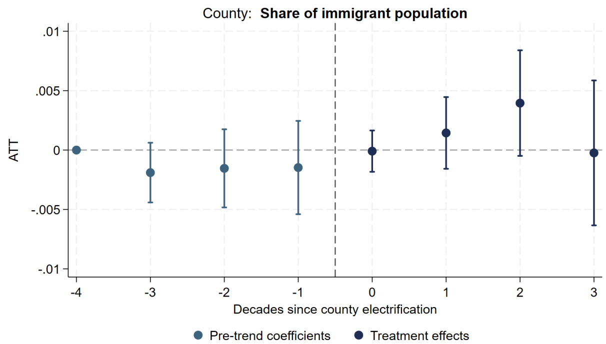
(b) Event study with treatment anticipated one decade



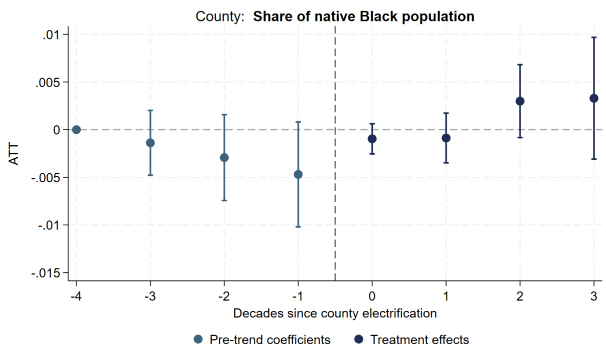
This figure plots the event-study coefficients estimated with the imputation method by Borusyak et al. (2024). The dependent variable is the county share of employment in the manufacturing sector. The y-axis reports the estimated coefficient and the 95% confidence interval of the difference between treated and control units (i.e., the average treatment effect on the treated, ATT). The horizontal dashed line indicates 0, i.e., no difference between the treated and control units. The vertical, dashed line indicates the occurrence of treatment, i.e., electrification. County c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. The estimation model includes: (i) county FEs, (ii) state $\times$ decade FEs, and (iii) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Standard errors are clustered at the county level.

Figure C.2: Event-Study Results: Electrification and ethnic composition of the county population

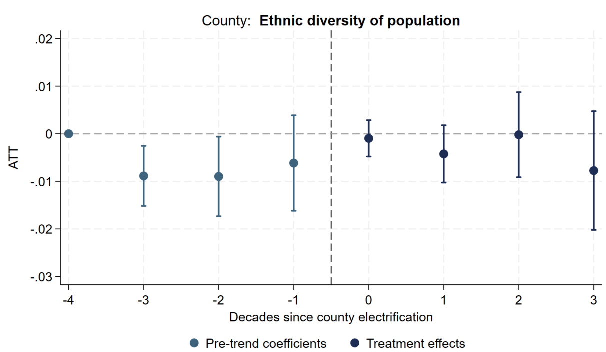
(a) Share of Foreign-born Population



(b) Share of Black U.S.-native Population



(c) Ethnic Diversity of Population



This figure plots the event-study coefficients estimated with the imputation method by Borusyak et al. (2024). The dependent variable is reported at the top of each figure: at the top, the share of foreign-born population in the county; in the middle, the share of Black U.S.-native population in the county; and at the bottom, the ethnic diversity of the county population. The y-axis reports the estimated coefficient and the 95% confidence interval of the difference between treated and control units (i.e., the average treatment effect on the treated, ATT). The horizontal dashed line indicates 0, i.e., no difference between the treated and control units. The vertical, dashed line indicates the occurrence of treatment, i.e., electrification. County c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. The estimation model includes: (i) county FEs, (ii) state $\times$ decade FEs, and (iii) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Standard errors are clustered at the county level.

Table C.2: DID Results: Electrification and ethnic composition of the county population

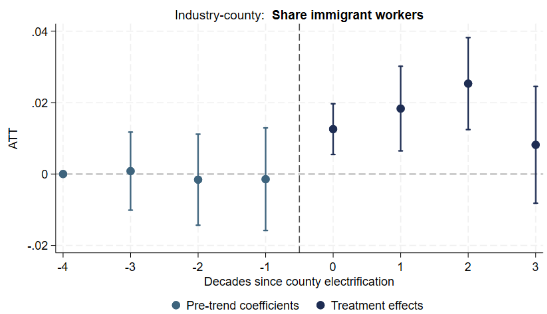
	County-decade level analysis		
	(1) *Share Foreign-born Population*	(2) *Share Black U.S.-born Popul.*	(3) *Ethnic Diveristy of Population*
Electrified	0.001	0.000	-0.003
	(0.001)	(0.001)	(0.003)
Observation N	13,576	13,576	13,576
Cluster N	2,799	2,799	2,799
Outcome mean$_{1900}$	0.091	0.134	0.357
Outcome mean$_{\text{Non-treated}}$	0.065	0.129	0.299
County FE	✓	✓	✓
State-decade FE	✓	✓	✓
County Population$_{1900} \times \mathbb{I}_t$	✓	✓	✓
Clustered SE: County$_{1900}$	✓	✓	✓
Pre-trend test: F-stat	0.888	1.112	3.391
$\beta_{h<0} = 0$ [p-value]	[0.446]	[0.343]	[0.017]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (ATT). The unit of analysis is a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in column (1), the share of immigrant (foreign-born) individuals in the county; in column (2), the share of Black U.S.-native individuals in the county; and in columns (3), the index of ethnic diversity of the county population (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"). The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. All estimation models include: (i) county FEs, (ii) state × decade FEs, and (iii) the county population in 1900 × decade dummies as control. In all estimation models, the sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, with $h = 0$ start of treatment), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation (F-statistics and p-value in square brackets). Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

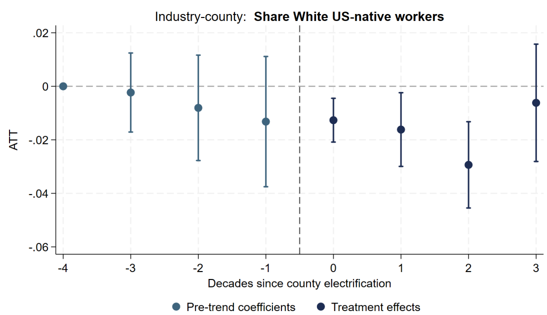
C.2 Additional Results: Employment Consequences of Electrification

Figure C.3: Event-Study Results: Electrification and ethnic composition of the labor force in manufacturing industries

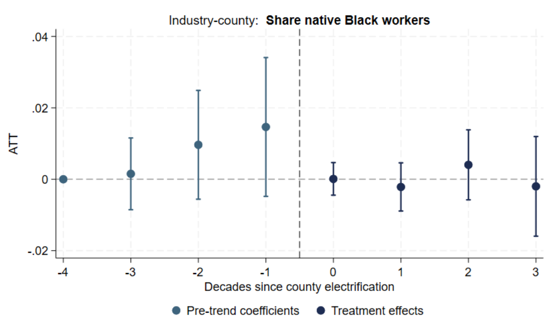
(a) Share of Immigrant Workers



(b) Share of White U.S.-Native Workers



(c) Share of Black U.S.-Native Workers



This figure plots the event-study coefficients estimated with the imputation method by Borusyak et al. (2024). The dependent variable is reported at the top of each figure: at the top, the share of immigrant (foreign-born) workers in the industry; in the middle, the share of U.S.-born workers in the industry (regardless of race); at the bottom, the share of Black U.S.-born workers among U.S.-native workers in the industry. The y-axis reports the estimated coefficient and the 95% confidence interval of the difference between treated and control units (i.e., the average treatment effect on the treated, ATT). The horizontal dashed line indicates 0, i.e., no difference between the treated and control units. The vertical, dashed line indicates the occurrence of treatment (i.e., electrification). A county c "treated" in decade t if it is electrified. All industries located in an electrified county are treated. The x-axis shows the number of decades since treatment (event study periods): electrification happens in period 0, negative values indicate pre-treatment and positive values indicate post-treatment periods. The estimation method sets the estimated coefficient of the pre-treatment period farthest from the treatment. The estimation model includes: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level.

Table C.3: DID Results: Electrification and ethnic composition of the labor force in manufacturing industries

	Industry-county-decade level analysis		
	(1) *Share Immigrant Workers*	(2) *Share Native Workers*	(3) *Share Black on Native Workers*
Electrified	0.016***	-0.016***	0.003
	(0.005)	(0.005)	(0.004)
Observation N	276,245	276,245	276,173
Cluster N	2,799	2,799	2,799
Outcome mean_{1900}	0.188	0.812	0.073
Outcome $\text{mean}_{\text{Non-treated}}$	0.110	0.890	0.096
Industry × County FE	✓	✓	✓
State × Decade FE	✓	✓	✓
Industry × Decade FE	✓	✓	✓
County $\text{population}_{1900} \times \mathbb{I}_t$	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓
Pre-trend test: F-stat	0.149	0.149	1.254
$\beta_{h<0} = 0$ [p-value]	[0.930]	[0.930]	[0.289]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in column (1), the share of immigrant (foreign-born) workers in the industry; in column (2), the share of U.S.-born workers in the industry (regardless of race); and in column (3), the share of Black U.S.-born workers among U.S.-native workers in the industry. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.4: DID Results: Electrification and ratio of occupational score by ethnic categories in manufacturing industries

	Industry-county-decade level analysis		
	Occupational Score Ratio		
	(1) *Immigrant / Native White*	(2) *Immigrant / Native Black*	(3) *Native Black / Native White*
Electrified	0.211	0.531	-0.019
	(0.254)	(1.444)	(0.147)
Observation N	93,815	20,094	50,951
Cluster N	2,799	2,799	2,799
Outcome mean_{1900}	1.017	1.155	0.940
Outcome $\text{mean}_{\text{Non-treated}}$	1.622	4.217	1.859
Industry $\times$ County FE	✓	✓	✓
State $\times$ Decade FE	✓	✓	✓
Industry $\times$ Decade FE	✓	✓	✓
County $\text{population}_{1900} \times \mathbb{I}_t$	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓
Pre-trend test: F-stat	1.848	0.936	1.292
$\beta_{h<0} = 0$ [p-value]	[0.136]	[0.410]	[0.276]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in column (1), the occupational score ratio between immigrant and White U.S.-native workers in the industry; in column (2), the occupational score ratio between immigrant and Black U.S.-native workers in the industry; and in column (3), the occupational score ratio between Black and White U.S.-native workers in the industry. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry $\times$ county FEs, (ii) state $\times$ decade FEs, (iii) industry $\times$ decade FEs, and (iv) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

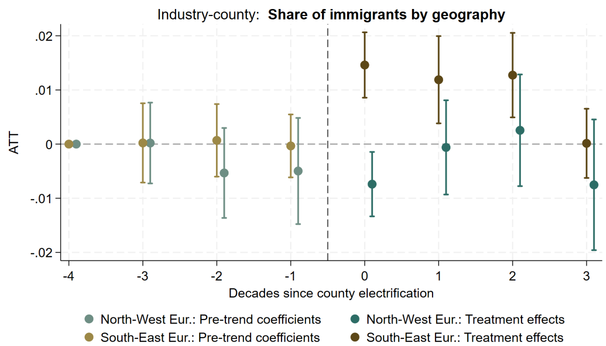
Table C.5: DID results: Electrification and share of immigrant workers by their characteristics

	Industry-county-decade level DID analysis					
	Share of immigrant workers by characteristics:					
	Geography		*Language*		*Religion*	
	(1) North-West Europe	(2) South-East Europe	(3) English Europe	(4) Non-English Europe	(5) Protestant Europe	(6) Non-Protestant Europe
Electrified	-0.003	0.011***	-0.004*	0.011***	-0.004	0.012***
	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)
Observation N	276,245	276,245	276,245	276,245	276,245	276,245
Cluster N	2,799	2,799	2,799	2,799	2,799	2,799
Outcome mean$_{1900}$	0.140	0.017	0.039	0.118	0.124	0.032
Outcome mean$_{\text{Non-treated}}$	0.075	0.017	0.019	0.073	0.068	0.024
Industry-county FE	✓	✓	✓	✓	✓	✓
State-decade FE	✓	✓	✓	✓	✓	✓
Industry-decade FE	✓	✓	✓	✓	✓	✓
County Population$_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clustered SE: County$_{1900}$	✓	✓	✓	✓	✓	✓
Pre-trend test: F-stat	2.733	0.150	1.726	0.583	3.148	0.347
$\beta_{h<0} = 0$ [p-value]	[0.042]	[0.930]	[0.160]	[0.626]	[0.024]	[0.792]

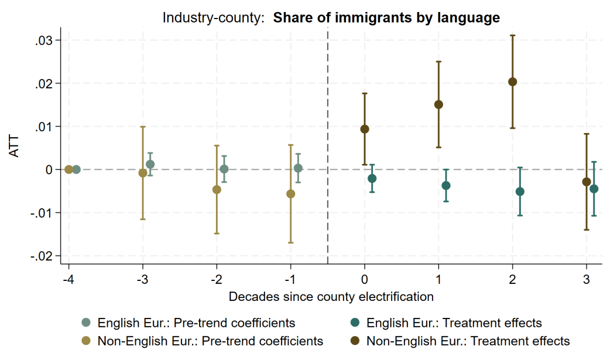
Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: (1) share of immigrant workers from Northern and Western Europe; (2) share of immigrant workers from Southern and Eastern Europe; (3) share of immigrant workers from English-speaking Europe; (4) share of immigrant workers from non-English-speaking Europe; (5) share of immigrant workers from majority-protestant Europe; (6) share of immigrant workers from majority-non-protestnat Europe. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry $\times$ county FEs, (ii) state $\times$ decade FEs, (iii) industry $\times$ decade FEs, and (iv) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure C.4: Event-Study Results: Electrification and share of immigrant workers by their characteristics

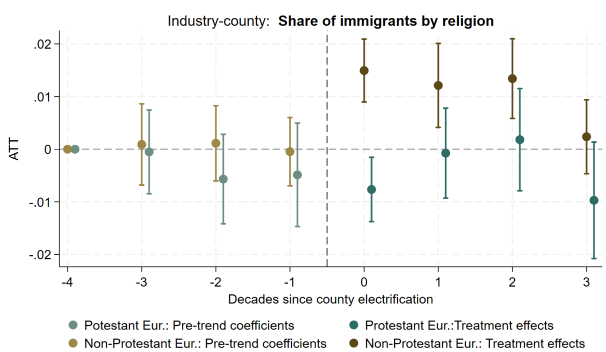
(a) Share of Immigrant Workers, by Geography of Origin



(b) Share of Immigrant Workers, by Language of Origin



(c) Share of Black U.S.-Native Workers



This figure plots the event-study coefficients estimated with the imputation method by Borusyak et al. (2024). The dependent variable is the share of immigrant workers by their characteristics, as reported at the top of each figure: at the top, by geographical origin (Northern-Western/Southern-Eastern); in the middle, by language of origin (English/non-English); at the bottom, by religion of origin (Protestant/non-Protestant). The y-axis reports the estimated coefficient and the 95% confidence interval of the difference between treated and control units (i.e., the average treatment effect on the treated, ATT). The horizontal dashed line indicates 0, i.e., no difference between the treated and control units. The vertical, dashed line indicates the occurrence of treatment, i.e., electrification. County c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. The estimation model includes: (i) industry $\times$ county FEs, (ii) state $\times$ decade FEs, (iii) industry $\times$ decade FEs, and (iv) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level.

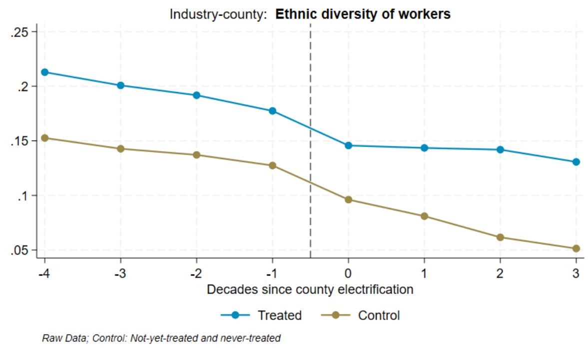
Table C.6: Event-study results: Electrification and ethnic integration of the labor force in manufacturing industries

	Industry-county-decade level analysis	
	(1) *Ethnic Diversity*	(2) *Ethnic Segregation*
$Post\text{-}treatment_3$	0.029***	-0.075***
	(0.011)	(0.017)
$Post\text{-}treatment_2$	0.049***	-0.050***
	(0.009)	(0.010)
$Post\text{-}treatment_1$	0.031***	-0.025***
	(0.007)	(0.005)
$Post\text{-}treatment_0$	0.030***	-0.016***
	(0.005)	(0.004)
$Pre\text{-}treatment_1$	-0.006	0.017*
	(0.010)	(0.009)
$Pre\text{-}treatment_2$	-0.007	0.010
	(0.009)	(0.007)
$Pre\text{-}treatment_3$	-0.004	0.007
	(0.008)	(0.006)
$Pre\text{-}treatment_4$	0.000	0.000
	(.)	(.)
Observation N	235,062	122,264
Cluster N	2,799	2,710
Outcome mean_{1900}	0.201	0.114
Outcome $\text{mean}_{\text{Non-treated}}$	0.132	0.184
Industry × County FE	✓	✓
State × Decade FE	✓	✓
Industry × Decade FE	✓	✓
County $\text{population}_{1900} \times \mathbb{I}_t$	✓	✓
Clustered SE: County_{1900}	✓	✓
Pre-trend test: F-stat	0.239	1.300
$\beta_{h<0} = 0$ [p-value]	[0.869]	[0.273]

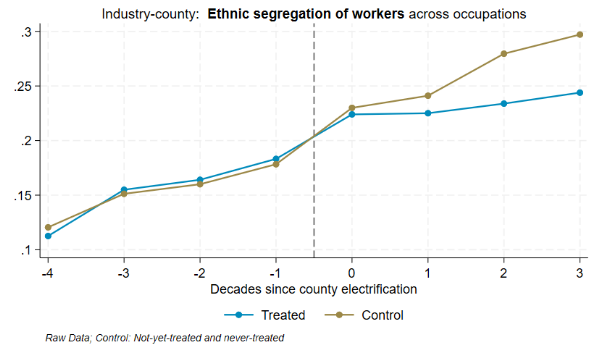
Note: This table reports the event-study coefficients estimated with the imputation method by Borusyak et al. (2024). Each estimated coefficient is the difference between treated and control units in the corresponding event-study period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in column (1), the index of ethnic diversity of workers within the industry (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"); and in column (2), the index of ethnic segregation of workers across occupations within the industry (from 0 = "no segregation" and 1 = "complete segregation"). The variables *Post-treatment* indicate the post-treatment period ($h \geq 0$); the variables *Pre-treatment* indicate the pre-treatment period ($h < 0$). A county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. In all estimation models, the sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. These estimates correspond to the coefficients plotted in Figure 2.

Figure C.5: Trend in the data: Electrification and the ethnic integration of the labor force in manufacturing industries

(a) Ethnic diversity of workers within industry



(b) Ethnic segregation of workers across occupations within industry



Note: This figure shows the trends in the data, built from the raw variables.The light-blue line represents treated units and the ochre line represents control units. To construct the outcome trends, I analyze each decade of electrification separately. Units that become electrified in a given decade are classified as treated, and their event time is defined based on that decade. Units not yet electrified in that decade serve as controls and are included until they eventually receive treatment. For each decade, I assign control units the same event-time periods as their corresponding treated units. I repeat this process for all decades, stack the resulting data, and then compute the average outcomes for treated and control groups by event time. Figure C.5a reports the trend for the index of ethnic diversity of workers in manufacturing industries (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"). Figure C.5b reports the trend for the index of ethnic segregation of workers across occupations within manufacturing industries (from 0 = "no segregation" to 1 = "complete segregation"). The unit of analysis is a manufacturing industry q in county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The y-axis reports the mean value of the respective outcome. The vertical, dashed line indicates the occurrence of treatment (i.e., electrification). A county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. These figures on the outomes' trends in the data complement the event study results shown in Figure 2.

Table C.7: Robustness – DID Results: Electrification and share of immigrant workers in manufacturing industries, excluding geographical regions

	Industry-county-decade level analysis			
	Share Immigrant Workers			
	(1) Excl. NY	(2) Excl. CA	(3) Excl. New England	(4) Excl. South
Electrified	0.013**	0.016***	0.018***	0.022***
	(0.005)	(0.005)	(0.005)	(0.008)
Observation N	268,326	269,992	273,068	166,377
Cluster N	2,742	2,743	2,739	1,532
Outcome mean_{1900}	0.187	0.184	0.186	0.248
Outcome $\text{mean}_{\text{Non-treated}}$	0.108	0.108	0.109	0.156
Industry × County FE	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓
County $\text{population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓
Pre-trend test: F-stat	0.150	0.149	0.147	0.324
$\beta_{h<0} = 0$ [p-value]	[0.930]	[0.930]	[0.932]	[0.808]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the share of immigrant (foreign-born) workers in the industry. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. In column (1), the sample **excludes counties in the state of New York**. In column (2), the sample **excludes counties in the state of California**. In column (3), the sample **excludes counties in New England** (ME, VT, NH, MA, CT, RI). In column (4), the sample **excludes counties in the South** (AL, AR, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV). Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.8: Robustness – DID Results: Electrification and ethnic diversity of workers in manufacturing industries, excluding geographical regions

	Industry-county-decade level analysis			
	Ethnic Diversity of Workers			
	(1) Excl. NY	(2) Excl. CA	(3) Excl. New England	(4) Excl. South
Electrified	0.033***	0.033***	0.036***	0.039***
	(0.006)	(0.006)	(0.006)	(0.007)
Observation N	227,217	228,843	231,903	161,848
Cluster N	2,742	2,743	2,739	1,532
Outcome mean_{1900}	0.198	0.198	0.197	0.240
Outcome $\text{mean}_{\text{Non-treated}}$	0.129	0.130	0.130	0.157
Industry × County FE	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓
County $\text{population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓
Pre-trend test: F-stat	0.244	0.239	0.238	0.411
$\beta_{h<0} = 0$ [p-value]	[0.865]	[0.869]	[0.870]	[0.745]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the ethnic diversity of workers in manufacturing industries (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"). The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. In column (1), the sample **excludes counties in the state of New York**. In column (2), the sample **excludes counties in the state of California**. In column (3), the sample **excludes counties in New England** (ME, VT, NH, MA, CT, RI). In column (4), the sample **excludes counties in the South** (AL, AR, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV). Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.9: Robustness – DID Results: Electrification and ethnic segregation of workers in manufacturing industries, excluding geographical regions

	Industry-county-decade level analysis			
	Ethnic Segregation of Workers			
	(1) Excl. NY	(2) Excl. CA	(3) Excl. New England	(4) Excl. South
Electrified	-0.036***	-0.042***	-0.033***	-0.015**
	(0.007)	(0.006)	(0.006)	(0.006)
Observation N	116,911	118,532	119,943	75,834
Cluster N	2,653	2,655	2,650	1,505
Outcome mean_{1900}	0.116	0.114	0.115	0.097
Outcome $\text{mean}_{\text{Non-treated}}$	0.186	0.184	0.185	0.151
Industry × County FE	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓
County $\text{population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓
Pre-trend test: F-stat	1.266	1.300	1.310	2.266
$\beta_{h<0} = 0$ [p-value]	[0.285]	[0.273]	[0.270]	[0.080]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the ethnic segregation of workers across occupations within manufacturing industries (from 0 = "no segregation" to 1 = "complete segregation"). The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. In column (1), the sample **excludes counties in the state of New York**. In column (2), the sample **excludes counties in the state of California**. In column (3), the sample **excludes counties in New England** (ME, VT, NH, MA, CT, RI). In column (4), the sample **excludes counties in the South** (AL, AR, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV). Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.10: Robustness – DID Results: Electrification and ethnic composition and integration of the labor force in manufacturing industries, excluding treatment in 1940 and 1930

	Industry-county-decade level analysis					
	Share Immigrants		*Ethnic Diveristy*		*Ethnic Segregation*	
	(1)	(2)	(3)	(4)	(5)	(6)
Electrified	0.020***	0.029***	0.038***	0.053***	-0.045***	-0.049***
	(0.005)	(0.008)	(0.007)	(0.010)	(0.009)	(0.009)
Observation N	228,547	166,509	195,329	143,618	101,949	73,734
Cluster N	2,358	1,828	2,358	1,828	2,295	1,777
Outcome mean$_{1900}$	0.192	0.199	0.205	0.208	0.112	0.113
Outcome mean$_{\text{Non-treated}}$	0.113	0.112	0.133	0.128	0.180	0.188
Industry × County FE	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓	✓	✓
County population$_{1900}$ × $\mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clustered SE: County$_{1900}$	✓	✓	✓	✓	✓	✓
Excluding electrified after 1930	✓		✓		✓	
Excluding electrified after 1920		✓		✓		✓
Pre-trend test: F-stat	0.676	0.093	0.576	0.619	2.865	0.465
$\beta_{h<0} = 0$ [p-value]	[0.509]	[0.760]	[0.562]	[0.432]	[0.057]	[0.495]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in columns (1)-(2), the share of immigrant (foreign-born) workers in the industry; in columns (3)-(4), the ethnic diversity of workers in manufacturing industries (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"); and in columns (5)-(6), the thnic segregation of workers across occupations within manufacturing industries (from 0 = "no segregation" to 1 = "complete segregation"). The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. In columns (1), (3), and (5), the sample **excludes counties that get electrified between 1930 and 1940.** In columns (2), (4), and (6), the sample **excludes counties that get electrified between 1920 and 1940.** Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.11: Robustness – DID Results: Electrification and share of immigrant workers in manufacturing industries, including controls for county manufacturing

	Industry-county-decade level analysis						
	Share Immigrant Workers						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Electrified	0.017***	0.013**	0.019***	0.015***	0.017***	0.017***	0.020***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Observation N	275,408	276,245	272,742	259,505	273,846	267,930	211,276
Cluster N	2,773	2,799	2,748	2,649	2,745	2,633	1,837
Outcome Mean_{1900}	0.188	0.188	0.186	0.182	0.188	0.190	0.195
Outcome $\text{Mean}_{\text{Non-treated}}$	0.110	0.110	0.110	0.106	0.111	0.113	0.121
Industry × County FE	✓	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓	✓
Add. Ctrl.: Cnty Manuf. $\text{Share}_{1900} \times \mathbb{I}_t$	✓						
Add. Ctrl.: Cnty Manuf. Share_t		✓					
Excl. Cnty Manuf. $\text{Share}_{1900} > 99^{\text{th}}$pctl			✓				
Excl. Cnty Manuf. $\text{Share}_{1900} > 95^{\text{th}}$pctl				✓			
Excl. Cnty Manuf. $\text{Share}_{1900} < 1^{\text{st}}$pctl					✓		
Excl. Cnty Manuf. $\text{Share}_{1900} < 5^{\text{th}}$pctl						✓	
Excl. Cnty Fully $\text{Rural}_{1900-1940}$							✓
Pre-trend test: F-stat	0.140	0.235	1.049	1.181	0.139	0.103	0.299
$\beta_{h<0} = 0$ [p-value]	[0.936]	[0.872]	[0.370]	[0.316]	[0.937]	[0.958]	[0.826]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the share of immigrant (foreign-born) workers in the industry. Each column **controls for some aspects related to the county orientation towards manufacturing**. Details are reported on the table for each column. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.12: Robustness – DID Results: Electrification and share of ethnic diversity of workers in manufacturing industries, including controls for county manufacturing

	Industry-county-decade level analysis						
	Ethnic Diversity of Workers						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Electrified	0.036***	0.027***	0.039***	0.036***	0.036***	0.037***	0.043***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
Observation N	234,285	235,062	231,812	219,634	233,085	228,449	178,468
Cluster N	2,773	2,799	2,748	2,649	2,745	2,633	1,837
Outcome Mean_{1900}	0.201	0.201	0.199	0.192	0.202	0.204	0.223
Outcome $\text{Mean}_{\text{Non-treated}}$	0.132	0.132	0.131	0.126	0.133	0.135	0.157
Industry × County FE	✓	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓	✓
Add. Ctrl.: Cnty Manuf. $\text{Share}_{1900} \times \mathbb{I}_t$	✓						
Add. Ctrl.: Cnty Manuf. Share_t		✓					
Excl. Cnty Manuf. $\text{Share}_{1900} > 99^{\text{th}}$pctl			✓				
Excl. Cnty Manuf. $\text{Share}_{1900} > 95^{\text{th}}$pctl				✓			
Excl. Cnty Manuf. $\text{Share}_{1900} < 1^{\text{st}}$pctl					✓		
Excl. Cnty Manuf. $\text{Share}_{1900} < 5^{\text{th}}$pctl						✓	
Excl. Cnty Fully $\text{Rural}_{1900-1940}$							✓
Pre-trend test: F-stat	0.170	0.434	1.398	2.040	0.166	0.164	0.604
$\beta_{h<0} = 0$ [p-value]	[0.916]	[0.729]	[0.242]	[0.106]	[0.919]	[0.921]	0.612]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the index of ethnic diversity of workers within the industry (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"). Each column **controls for some aspects related to the county orientation towards manufacturing**. Details are reported on the table for each column. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.13: Robustness – DID Results: Electrification and share of ethnic segregation of workers in manufacturing industries, including controls for county manufacturing

	Industry-county-decade level analysis						
	Ethnic Segregation of Workers						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Electrified	-0.036***	-0.029***	-0.036***	-0.036***	-0.036***	-0.036***	-0.030***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Observation N	122,185	122,264	120,291	111,295	121,736	120,133	104,510
Cluster N	2,692	2,710	2,667	2,568	2,668	2,578	1,825
Outcome Mean_{1900}	0.114	0.114	0.114	0.116	0.114	0.113	0.112
Outcome $\text{Mean}_{\text{Non-treated}}$	0.184	0.184	0.184	0.187	0.183	0.182	0.180
Industry × County FE	✓	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓	✓
Add. Ctrl.: Cnty Mnf. $\text{Sh.}_{1900} \times \mathbb{I}_t$	✓						
Add. Ctrl.: Cnty Mnf. Sh._t		✓					
Excl. Cnty Mnf. $\text{Sh.}_{1900} > 99^{\text{th}}$pctl			✓				
Excl. Cnty Mnf. $\text{Sh.}_{1900} > 95^{\text{th}}$pctl				✓			
Excl. Cnty Mnf. $\text{Sh.}_{1900} < 1^{\text{st}}$pctl					✓		
Excl. Cnty Mnf. $\text{Sh.}_{1900} < 5^{\text{th}}$pctl						✓	
Excl. Cnty Fully $\text{Rural}_{1900-1940}$							✓
Pre-trend test: F-stat	1.355	2.018	1.473	2.285	1.356	1.362	1.131
$\beta_{h<0} = 0$ [p-value]	[0.255]	[0.109]	[0.220]	[0.077]	[0.255]	[0.253]	[0.336]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the index of ethnic segregation of workers across occupations within the industry (from 0 = "no segregation" and 1 = "complete segregation"). Each column **controls for some aspects related to the county orientation towards manufacturing**. Details are reported on the table for each column. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.14: Robustness – DID Results: Electrification and ethnic composition and integration of the labor force in manufacturing industries, with treatment anticipated one decade

	Industry-county-decade level analysis		
	(1) *Share Immigrants*	(2) *Ethnic Diveristy*	(3) *Ethnic Segregation*
Electrified	-0.004	0.016**	-0.026***
	(0.004)	(0.006)	(0.007)
Observation N	160,406	134,742	69,260
Cluster N	1,536	1,536	1,723
Outcome mean_{1900}	0.188	0.201	0.114
Outcome $\text{mean}_{\text{Non-treated}}$	0.110	0.132	0.184
Industry × County FE	✓	✓	✓
State × Decade FE	✓	✓	✓
Industry × Decade FE	✓	✓	✓
County $\text{population}_{1900} \times \mathbb{I}_t$	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓
Pre-trend test: F-stat	0.728	0.165	0.551
$\beta_{h<0} = 0$ [p-value]	[0.483]	[0.848]	[0.577]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in column (1), the share of immigrant (foreign-born) workers in the industry; in column (2), the the ethnic diversity of workers in manufacturing industries (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"); and in column (3), the ethnic segregation of workers across occupations within manufacturing industries (from 0 = "no segregation" to 1 = "complete segregation"). The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. The table reports the estimates using the **decade before electrification** as treatment. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.15: Robustness – DID Results: Electrification and share of immigrant workers in manufacturing industries, including controls for ethnic characteristics of county population

	Industry-county-decade level analysis					
	Share Immigrant Workers					
	(1)	(2)	(3)	(4)	(5)	(6)
Electrified	0.015***	0.017***	0.014***	0.010***	0.016***	0.014***
	(0.004)	(0.005)	(0.005)	(0.003)	(0.005)	(0.004)
Observation N	276,245	276,245	275,408	276,245	276,245	276,245
Cluster N	2,799	2,799	2,773	2,799	2,799	2,799
Outcome Mean_{1900}	0.188	0.188	0.188	0.188	0.188	0.188
Outcome $\text{Mean}_{\text{Non-treated}}$	0.110	0.110	0.110	0.110	0.110	0.110
Industry × County FE	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓
Add. control: Cnty Immigr. $\text{Share}_{1900} \times \mathbb{I}_t$	✓					
Add. control: Cnty Black $\text{Share}_{1900} \times \mathbb{I}_t$		✓				
Add. control: Cnty Ethn. $\text{Divers.}_{1900} \times \mathbb{I}_t$			✓			
Add. control: Cnty Immigrant Share_t				✓		
Add. control: Cnty Black Share_t					✓	
Add. control: Cnty Ethnic Diversity_t						✓
Pre-trend test: F-stat	0.237	0.172	0.087	0.288	0.150	0.362
$\beta_{h<0} = 0$ [p-value]	[0.870]	[0.916]	[0.967]	[0.834]	[0.930]	[0.781]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the share of immigrant (foreign-born) workers in the industry. Each column **includes an ethnic characteristic of the county population**: the share of foreign-born individuals, the share of Black U.S.-native individuals, and the overall ethnic diversity of the population. Columns (1)-(3) include the value of the corresponding county variable in 1900 × decade dummies as control. Columns (4)-(6) include the decade value of the corresponding county variable as control. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.16: Robustness – DID Results: Electrification and ethnic diversity of workers in manufacturing industries, including controls for ethnic characteristics of county population

	Industry-county-decade level analysis					
	Ethnic Diversity of Workers					
	(1)	(2)	(3)	(4)	(5)	(6)
Electrified	0.034***	0.036***	0.034***	0.029***	0.034***	0.029***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)
Observation N	235,062	235,062	234,285	235,062	235,062	235,062
Cluster N	2,799	2,799	2,773	2,799	2,799	2,799
Outcome Mean_{1900}	0.201	0.201	0.201	0.201	0.201	0.201
Outcome $\text{Mean}_{\text{Non-treated}}$	0.132	0.132	0.132	0.132	0.132	0.132
Industry $\times$ County FE	✓	✓	✓	✓	✓	✓
State $\times$ Decade FE	✓	✓	✓	✓	✓	✓
Industry $\times$ Decade FE	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓
Add. control: Cnty Immigr. $\text{Share}_{1900} \times \mathbb{I}_t$	✓					
Add. control: Cnty Black $\text{Share}_{1900} \times \mathbb{I}_t$		✓				
Add. control: Cnty Ethn. $\text{Divers.}_{1900} \times \mathbb{I}_t$			✓			
Add. control: Cnty Immigrant Share_t				✓		
Add. control: Cnty Black Share_t					✓	
Add. control: Cnty Ethnic Diversity_t						✓
Pre-trend test: F-stat	0.186	0.256	0.186	0.157	0.204	0.033
$\beta_{h<0} = 0$ [p-value]	[0.906]	[0.857]	[0.906]	[0.925]	[0.894]	[0.992]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the index of ethnic diversity of workers within the industry (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"). Each column **includes an ethnic characteristic of the county population**: the share of foreign-born individuals, the share of Black U.S.-native individuals, and the overall ethnic diversity of the population. Columns (1)-(3) include the value of the corresponding county variable in 1900 $\times$ decade dummies as control. Columns (4)-(6) include the decade value of the corresponding county variable as control. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry $\times$ county FEs, (ii) state $\times$ decade FEs, (iii) industry $\times$ decade FEs, and (iv) the county population in 1900 $\times$ decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.17: Robustness – DID Results: Electrification and ethnic segregation of workers in manufacturing industries, including controls for ethnic characteristics of county population

	Industry-county-decade level analysis					
	Ethnic Segregation of Workers					
	(1)	(2)	(3)	(4)	(5)	(6)
Electrified	-0.036***	-0.035***	-0.034***	-0.034***	-0.035***	-0.034***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observation N	122,264	122,264	122,185	122,264	122,264	122,264
Cluster N	2,710	2,710	2,692	2,710	2,710	2,710
Outcome Mean_{1900}	0.114	0.114	0.114	0.114	0.114	0.114
Outcome $\text{Mean}_{\text{Non-treated}}$	0.184	0.184	0.184	0.184	0.184	0.184
Industry × County FE	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓
Add. control: Cnty Immigr. $\text{Share}_{1900} \times \mathbb{I}_t$	✓					
Add. control: Cnty Black $\text{Share}_{1900} \times \mathbb{I}_t$		✓				
Add. control: Cnty Ethn. $\text{Divers.}_{1900} \times \mathbb{I}_t$			✓			
Add. control: Cnty Immigrant Share_t				✓		
Add. control: Cnty Black Share_t					✓	
Add. control: Cnty Ethnic Diversity_t						✓
Pre-trend test: F-stat	1.382	1.344	1.345	1.329	1.457	1.346
$\beta_{h<0} = 0$ [p-value]	[0.247]	[0.258]	[0.258]	[0.263]	[0.224]	[0.258]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is the index of ethnic segregation of workers across occupations within the industry (from 0 = "no segregation" and 1 = "complete segregation"). Each column **includes an ethnic characteristic of the county population**: the share of foreign-born individuals, the share of Black U.S.-native individuals, and the overall ethnic diversity of the population. Columns (1)-(3) include the value of the corresponding county variable in 1900 × decade dummies as control. Columns (4)-(6) include the decade value of the corresponding county variable as control. The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. The sample excludes counties with population levels in the year 1900 above the 99th percentile of the distribution (29 counties that host the largest U.S. cities, which are considered as "always treated"). Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.18: Robustness – DID Results: Electrification and ethnic composition and integration of the labor force in manufacturing industries, excluding counties with large baseline population size

	Industry-county-decade level analysis					
	Share Immigrants		*Ethnic Diveristy*		*Ethnic Segregation*	
	(1)	(2)	(3)	(4)	(5)	(6)
Electrified	0.014***	0.010*	0.036***	0.033***	-0.038***	-0.037***
	(0.005)	(0.005)	(0.006)	(0.007)	(0.006)	(0.007)
Observation N	253,389	230,219	213,409	192,931	104,786	89,431
Cluster N	2,660	2,518	2,660	2,518	2,579	2,437
Outcome Mean_{1900}	0.182	0.179	0.386	0.336	0.081	0.089
Outcome $\text{Mean}_{\text{Non-treated}}$	0.104	0.100	0.390	0.328	0.104	0.122
Industry × County FE	✓	✓	✓	✓	✓	✓
State × Decade FE	✓	✓	✓	✓	✓	✓
Industry × Decade FE	✓	✓	✓	✓	✓	✓
County $\text{Population}_{1900} \times \mathbb{I}_t$	✓	✓	✓	✓	✓	✓
Clustered SE: County_{1900}	✓	✓	✓	✓	✓	✓
Excl. County $\text{Popul.}_{1900} > 95^{\text{th}}$pctl	✓		✓		✓	
Excl. County $\text{Popul.}_{1900} > 90^{\text{th}}$pctl		✓		✓		✓
Pre-trend test: F-stat	0.200	0.167	0.239	0.291	1.179	1.359
$\beta_{h<0} = 0$ [p-value]	[0.896]	[0.919]	[0.869]	[0.832]	[0.317]	[0.254]

Note: This table reports the difference-in-differences (DID) coefficients estimated with the imputation method by Borusyak et al. (2024). The estimated coefficient is the difference between treated and control units in the post-treatment period (i.e., the average treatment effect on the treated, or ATT). The unit of analysis is a manufacturing industry q in a county c and decade t. Counties are consistent over time, fixed at the 1900 borders. The dependent variable is reported at the top of the table: in columns (1)-(2), the share of immigrant (foreign-born) workers in the industry; in columns (3)-(4), the ethnic diversity of workers in manufacturing industries (from 0 = "complete homogeneity" to 1 = "complete heterogeneity"); and in columns (5)-(6), the thnic segregation of workers across occupations within manufacturing industries (from 0 = "no segregation" to 1 = "complete segregation"). The variable *Electrified* indicates the treatment: a county c is considered as "treated" in decade t if it is electrified. Once a county gets treated, it remains treated for the following periods. All manufacturing industries located in an electrified county are treated. All estimation models include: (i) industry × county FEs, (ii) state × decade FEs, (iii) industry × decade FEs, and (iv) the county population in 1900 × decade dummies as control. In columns (1), (3), and (5), the sample **excludes counties with a population in 1900 above the 95-*th* percentile** of the distribution. In columns (2), (4), and (6), the sample **excludes counties with a population in 1900 above the 95-*th* percentile** of the distribution. Estimates are weighted by the industry share of employment in the county. Standard errors are clustered at the county level and are reported in parentheses below the coefficients. The table reports the results of the direct test for the hypothesis that all pre-treatment coefficients are jointly = 0 ($\beta_{h<0} = 0$, where $h = 0$ indicates the period in which treatment starts), which corresponds to testing the DID identifying assumptions of parallel trends and no anticipation. The table reports the F-statistics and p-value (in square brackets) of the test. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.