Media and Assimilation: Evidence from the Golden Age of Radio*

Gianluca Russo†

April 7, 2024

Abstract

The advent of the radio networks in the early twentieth century marked the inception of the first national media platform in US history. This development coincided with the conclusion of the Age of Mass Migration, a historical period characterized by record-high diversity within the US society. How did the sudden rise of national radio networks affect the assimilation of millions of immigrants in the US? I answer this question by building quasi-random variation in the initial expansion of radio networks coverage. Through the application of a difference-in-difference design and leveraging evidence from diverse samples, I document that higher exposure to radio networks facilitated immigrant assimilation across various outcomes, including higher rates of intermarriage with natives, increased naturalization, and a greater propensity to adopt names that sounded less foreign. Additionally, I find that exposure to radio networks led to enhanced English proficiency among recently arrived immigrants. I argue that a key channel of assimilation was through the impact of role models on the media. Using data on baseball, I show that immigrants perceived baseball players as media role models of assimilation. Immigrant households were more likely to name their children after baseball players who shared similar ancestral origins to them or who achieved notable success during a season.

Keywords: Media, Immigrant Assimilation, Role Models

JEL Classification: J15, N32, Z10

*I would like to thank Bob Margo, James Feigenbaum and Daniele Paserman who provided excellent guidance during the PhD. In addition, I would like to thank Ran Abramitzky, Felipe Valencia Caicedo, Ruben Durante, Claudia Goldin, Leander Heldring, Martin Fiszbein, Vicky Fouka, Luigi Pascali, Maria Petrova, Silvia Vannutelli, Marco Tabellini, Nico Voigtlander, Joachim Voth, Tianyi Wang for helpful comments. I would also like to thank seminar participants at UdeSA-EHA conference, Harvard econ history lunch seminar, Stanford econ history lunch seminar, NBER, EHS 2019 annual meetings, EHA 2019 annual meeting, SSHA 2019 annual meeting, UPF, UAB, Northwestern econ history lunch seminar, ASSA meetings and FCC. A previous version of this paper circulated as "Mass Media and Cultural Homogenization: Broadcasting the American Dream on the Radio." I am grateful to Caterina Mauri for sharing valuable data and to Daniil Sherstnev for providing excellent research assistance.

†Barcelona School of Economics, email:russo.gianluca.eco@gmail.com

1 Introduction

The past three decades have witnessed a nearly doubling of the number of international migrants, with the United States standing out as one of the largest recipients globally.[1] As of 2021, the foreign-born population in the U.S. has reached 14%, a figure comparable only to the peak observed during the Age of Mass Migration in the early twentieth century.[2] This rapid increase in immigrant arrivals has reignited the public debate on whether immigrants assimilate over time, and if so, what policies are conducive to successful assimilation.[3] The topic of immigrant assimilation is not new in the American public debate. About a century ago, on the eve of quota restrictions that halted unrestricted immigration from Europe, concerns about the integration of culturally distant immigrants were also at the forefront of political debate (Higham 2002).

Motivated by this debate and exploiting the context of the Age of the Mass Migration, recent studies have attempted at providing systematic evidence on the dynamic of immigrant assimilation. One side of this literature has indeed concluded that immigrants did assimilate over time (Abramitzky et al. 2020, 2014). Another branch of research has instead focused on understanding what fostered immigrant assimilation. However, despite a growing body of work and a rising interest in the topic, we still have scant evidence on how societies can actively promote the assimilation of immigrants. Offering systematic evidence on successful determinants of assimilation is crucial not only to shed light on this key historical juncture but also to inform the broader contemporary debate.

This paper fills this gap by studying the impact of radio networks, the first national media platform of American history, on immigrant assimilation. After the start of radio broadcasting in the early 1920s, several radio stations came together under two major radio networks in 1929: the Columbia Broadcasting System (CBS) and the National Broadcasting Company (NBC). Exploiting economies of scales and wider sponsorship, these networks could produce and distribute high quality content uniformly to their affiliates nationwide (Sterling and Kittross 2001), resulting in a notable decrease in content diversity across the airwaves and fostering novel avenues for national connectivity (Lippmann 2008). The hypothesis of this paper is that the steep rise in the homogeneity of the media content introduced by the

[1]The last international migration report from the United Nations before the pandemic reported a rise to 278 millions migrants worldwide. In 1990 the same number was 153 millions. During the same period the world population only increased by 50%. The United States is home to more international migrants than any other country in the world, about 50 millions (United Nations and Social Affairs 2019).

[2]For example, see the data elaborated by the Migration Policy Initiative here

[3]For a discussion of the public debate on immigration see for example Kirchick (2019) or Kaya (2023). Hainmueller and Hiscox (2010) and Citrin et al. (1997) analyze survey data and discuss how cultural and economics considerations regarding integration of immigrants are equally important.

networks facilitated immigrant assimilation through the provision of common cultural norms and similar references across geographic areas, immigrant groups and nativity status.

There are two key advantages of investigating the impact of homogeneous media on assimilation in this historical setting. First, the rise of national radio networks occurred within a relatively simpler media environment compared to today. Studying the introduction of a new media outlet in this context is less likely to be confounded by other media sources with similar reach, facilitating clearer insights into its effects. Second, the advent of radio networks followed the end of the Age of the Mass Migration, one of the most transformative periods in American history and the only one with comparable levels of immigration to today. Between 1850 and 1920, approximately 30 million immigrants arrived in the US from Europe, generating an extraordinary level of diversity in the US society (Abramitzky and Boustan 2017). How did the sudden rise of the first national media platform affect the assimilation of immigrants in the aftermath of the Age of the Mass Migration?

I answer this question by linking repeated cross-sections of immigrants from the US full count census to county-level variation in the initial expansion of radio networks coverage. To construct county-level exposure to the networks, I collect data on network-affiliated radio stations active in 1929, the first year of full network operation. I then feed the technical characteristics of the stations to a radio propagation model suitable to predict AM radio frequencies, the only technology available at the time. The output of the radio propagation model is a continuous variable that I standardize and use as the main independent variable of the paper (Enikolopov et al. 2011; Yanagizawa-Drott 2014).

Comparing the evolution of assimilation outcomes in counties differentially exposed to radio networks signal before and after the rise of the networks is prone to several concerns. Most notably, urbanization and agglomeration forces are key potential confounding factors, since they not only attracted the location of the most powerful antennas but could also simultaneously affect assimilation trends over time. To circumvent this threat, I implement an empirical strategy that adapts recent advancements in the media literature to the AM radio setting (Olken 2009; Durante et al. 2019; Wang 2021). Using this strategy, I isolate variation in radio networks coverage driven by the idiosyncratic mix of soil conditions affecting the propagation of the radio wave through space.

I incorporate the identifying variation in a standard difference-in-difference model which compares decadal changes in immigrant assimilation before and after the rise of the networks. The baseline model includes county, decade and country of origin fixed effects. The inclusion of county fixed effects allows me to absorb time-invariant characteristics, including the own impact of geography on the receiving county. Additionally, I directly control for a wide range of controls and in particular geographic characteristics interacted with decade fixed effects.

This allows me to further bolster the interpretation that the variation in networks signal strength I reconstruct is driven by the mix of soil characteristics encountered by the signal on its way to the receiver and not by local geographic characteristics.

Using this empirical strategy and focusing on male immigrants, I document that radio networks spurred immigrant assimilation across four key dimensions.[4] Firstly, exposure to the radio networks increased the likelihood of intermarriage between immigrants and natives: a standard deviation surge in networks coverage boosted intermarriage probability by 6 p.p. This result suggests that radio networks led immigrants to one of the final steps of assimilation (Gordon 1964), one that involved both the willingness of immigrant to assimilate and the desire of natives to accept them. Exposure to the networks also boosted assimilation outcomes depending solely on immigrants' effort such as naturalization and naming decisions. A standard deviation rise in networks coverage translated into a higher likelihood of naturalization by 2.5 p.p. and a reduction in the adoption of distinctively foreign names by about 5 p.p. Focusing on immigrants of recent arrivals, I also document that radio networks improved English proficiency by 8 p.p. The magnitude of the estimates are comparable across outcomes. In terms of pre-treatment standard deviations, one standard deviation increase in network coverage resulted in a 12% boost in intermarriage, a 5% rise in naturalization, a 12% decrease in the foreignness of immigrant children's names and a 16 % improvement in English proficiency.

I corroborate the casual nature of my results in several ways. I start by showing that the identifying variation I reconstruct is orthogonal to the levels of all outcome variables at baseline year (1920). Then, I demonstrate that the same variation is uncorrelated to several county-level socioeconomic characteristics, including urbanization, density or the size of the foreign-born community at baseline year. Third, I employ an event study design and document the absence of pre-trends in all four outcome variables. I then move on to perform a host of robustness checks to corroborate the stability of the estimated coefficients. Most notably, I provide additional evidence using naming decisions in a sample of pseudo birth cohorts extracted from the 1940 US census wave. Using yearly variation in naming patterns within the same households before and after the rise of the networks, I find results that paint a similar picture to the one observed in the repeated cross-section: exposure to the networks led immigrant households to adopt names that were less distinctively foreign. I discuss in detail these and other robustness checks in section 5.3, after the main results section.

[4] I focus on male immigrants in the main analysis because some of the questions in the US census used to build my outcome variables cover less census decades for females. In addition, results for female immigrants are qualitatively similar to the male sample. The estimates are commented in the text and reported in the Appendix.

In the second part of the paper, I provide evidence that a key channel behind the main results of the paper is the direct impact of the personalities who appeared on the networks and acted as successful models of assimilation. I focus on this specific channel for two reasons: first, media role models have been robustly identified as key drivers of behavioral change in society.[5] Second, historical accounts suggest a natural candidate of the media role models for immigrants during the interwar period: baseball players. Historians discuss how immigrant groups who were at the time plagued by stereotypes were able to finally change the way the US society looked at them through their success in the *national pastime* (Baldassaro 2011; Levine 1993).

I test whether baseball players acted as role models on the networks through the adoption of their first names in society (La Ferrara et al. 2012). I do so by constructing two key characteristics of the players that social psychologists have associated with successful role models: identification with the audience and success in reaching their objectives (Lockwood and Kunda 1997; Bandura and Walters 1977; Bandura et al. 1999). Exploiting within household variation in naming patterns, I investigate whether radio networks promoted the adoption of players' names active during the year of a child's birth, depending on their ancestral origin (identification) and their performance (success).

I provide evidence consistent with players acting as role model through the networks: both audience identification and players' success were pivotal factors influencing the adoption of players' names by immigrants. More specifically, I demonstrate that immigrant households exposed to radio networks exhibited a higher propensity to adopt baseball players' names if the player shared a similar ancestral origin or if the player achieved notable success during that season.

According to the historical context, the impact of baseball players should be concentrated on the players who, through their fame in baseball, were able to update the social image associated with immigrants from a similar origin. In line with this hypothesis, my findings reveal that the influence of players is primarily driven by baseball players with ancestral origins from Eastern Europe, Southern Europe, Russia, or Germany, countries whose immigrants in the United States faced significant discrimination during the early 1900s (Tabellini 2019; Fouka 2019b). Consistently, I also document that immigrants from these same countries drive the impact of radio networks on name adoption.

To contextualize the evidence from the main analysis within the results of the role model channel, I show that also the main results display a similar heterogeneous pattern along immigrant origin. More specifically, I document that the impact of radio networks on intermarriage, naturalization, name-foreignness and English proficiency is driven by immigrants

[5] For a review of the evidence see DellaVigna and La Ferrara (2015).

from Eastern Europe, Russia and Germany, who experienced discrimination and negative stereotypes at the start of the twentieth century.

One might be concerned that the adoption of successful player's names might just be a byproduct of the higher appearance of more successful players on the radio networks. I argue that this is not the case by showing specular responses on the negative side. While immigrants were more likely to pick names of successful players by 1.3 p.p. they were about equally less likely to choose a name that underperformed in a given season. These result shows convincing evidence that immigrants paid attention to the content of the broadcasts and the adoption of names was characterized by the success of the players on the networks.

Successful role models could impact immigrant assimilation also by reducing negative stereotypes from natives. In the last exercise of the paper, I explore this possibility by examining whether successful players also impacted naming choices among native households. By replicating the performance analysis on native households, I find that radio networks similarly increased the probability of native households naming their children after successful baseball players active during the same year of a child's birth. This result suggests that successful second-generation immigrant players featured on the networks might have played a role in alleviating negative stereotypes held by natives towards immigrants.

This paper makes several contributions. First, my results contribute to the growing literature on immigrant social assimilation (Abramitzky et al. 2014; Lleras-Muney and Shertzer 2015; Bandiera et al. 2019; Abramitzky et al. 2020; Gagliarducci and Tabellini 2022; Algan et al. 2022), intermarriage and cultural integration (Bisin and Verdier 2000; Bisin et al. 2004; Bisin and Tura 2019; Adda et al. 2020). I add to this literature novel evidence of a successful driver of social assimilation. Specifically, I complement a recent set of studies, some of which highlight the adverse effects of imposed assimilation policies, while others argue that reconciliation and assimilation occurred at the cost of increased discrimination towards other minority groups (Ang 2023; Esposito et al. 2023; Fouka et al. 2021). Differently from these studies, my findings shed light on an inclusive determinant of immigrant assimilation, which hinges on the voluntary consumption of media embedding shared cultural references through positive role models.

This paper is also related to the literature studying the impact of media on social and political outcomes (Stromberg 2004; Gentzkow 2006; Gentzkow et al. 2011; Enikolopov et al. 2011; DellaVigna et al. 2014; Adena et al. 2015; Durante et al. 2019) and to the studies linking media to nation building and conflict (Blouin and Mukand 2019; Armand et al. 2020; Yanagizawa-Drott 2014). In this strand of work, my paper is closely related to two other studies also examining the impact of radio in a similar historical context. In a recent publication, Wang (2021) shows that famous personalities on radio were able to disseminate racist and antisemitic values during the interwar period. In another study, Armand et al. (2023) argue that progressive radio programming had a positive impact on progressive attitudes towards civil rights, racial integration while reducing discriminatory values and practices. Exploiting a similar background, my results add to this literature evidence that media can be harnessed to promote the social assimilation of immigrants, particularly during a pivotal historical period when immigrant assimilation was synonymous with nation-building in the US. In line with these studies, my paper underscores the significance of media personalities, providing new insights into how their portrayal can enhance the impact of media or undermine it.

A key contribution of this paper is to the large literature in economics testing the efficacy of *edutainment* media programs through the emulation of positive (Jensen and Oster 2009; La Ferrara et al. 2012; Riley 2022) or negative (Farré and Fasani 2013) media role models.[6] My results contribute to this body of work by explicitly linking the characteristics of role models highlighted by social psychologists to their efficacy in promoting emulation (Lockwood and Kunda 1997; Bandura and Walters 1977; Bandura et al. 1999). Further, while most of the literature investigates the impact of positive and negative role models across different contexts, I provide first evidence of the variation in efficacy that is dependent on these characteristics within the same setting.

2 Radio Networks and the Golden Age of Radio

The Golden Age of Radio defined the first half of the twentieth century as much as the automobile did (Gordon 2017). In this section, I first briefly discuss the events surrounding the arrival of radio, the creation of radio networks and their social impact. I then describe the relationship between baseball and immigrants, how radio networks changed the way the country experienced baseball and how this fed back into the representation of second generation immigrants.

2.1 From Early Broadcasting to the Rise the Networks

Until the end of World War I, radio broadcasting was just a fad for amateurs. In 1919, Frank Conrad, an engineer at Westinghouse, started broadcasting music and covering baseball games regularly from Pittsburgh. Soon after, the popularity of the broadcasts led to the establishment of KDKA, the first official radio station in the United States (Scott 2008).

Early stations ran a basic schedule, and their equipment was quite rudimentary. The

[6] For a further discussion see La Ferrara (2016).

average station would broadcast a great deal of music and a mix of lectures, news and experimental broadcasts; interference and hours of dead air were not uncommon. By 1923, there were over five hundred active stations in the US but many were not powerful enough to reach farther than their immediate surroundings. The industry was characterized by high turnover rates; stations bloomed and quickly died out as owners grew concerned with their cost. In this context, investing in higher quality radio programs was not deemed profitable. Hence, the quality of the show was low and the schedule precarious (Lippmann 2008).

Despite the issues with radio schedule and interference, US households were instantly fascinated by "the voice from the air" (Sterling and Kittross 2001). Two aspects of radio contributed to its quick diffusion. First, assembling a rudimentary radio set was easy and it did not require a battery or electricity to work. Second, manufacturers exploited broadcasting to stimulate the sale of sets. To match demand between 1922 and 1925, 4.1 million radio sets had already been produced in the country (Sterling and Kittross 2001). The fast growth in popularity of the new medium made the US one of the largest radio audiences in the world (Lenthall 2008). Appendix Figure A.1 shows how fast radio ownership spread across the US between 1930 and 1940 despite the Great Depression.

The years between 1926 and 1933 marked the transition of radio as a fad for amateurs to a proper industry. In these years few key events led to the creation of the radio landscape that will dominate the air in the decades to follow: the modernization of federal regulation, the rise of the networks and the success of advertising. These elements led to a spectacular increase of the radio audience and the start of the *Golden Age of Radio* in the mid 1930s.

With the large number of radio stations, technical problems were systematically present. The US government responded to these pressing issues by promulgating the Radio Act in 1927. The new regulation established the Federal Radio Commission (FRC) with the goal of overseeing broadcasting stations. The Radio Act settled that radio waves were public property, entry was free of charge, but stations had to be licensed by the government on the basis of being "in the public interest, convenience, or necessity" (Sterling and Kittross 2001). The Act had two major outcomes: it helped alleviating interference between stations and facilitated the organization of the radio spectrum and the creation of networks.

The national networks as we know them developed in the late 1920s from the temporary and experimental networks. The first proper national network to arise was the National Broadcasting Network (NBC). In 1926 Radio Corporation of America (RCA) bought from AT&T the New York radio station WEAF and its experimental network of 15 stations. Separately, RCA purchased the Washington-based station WCAP. On November 1st 1926, RCA formally established the first permanent network through the foundation of the NBC corporation. Despite the formal inauguration in 1926, and following some internal restructuring, on December 23rd 1928 NBC began its full-time coast-to-coast national radio network (Sterling and Kittross 2001).

The second nation-wide radio network to arise was the Columbia Broadcasting System. Its foundation started from the merging of the United Independent Broadcasters (UIB) corporation and the Columbia Phonograph Corporation, a phonograph company. The two entities created a separate corporation in 1927 with the goal of securing a well-functioning radio network and boosting sales of phonographs. The first year of operation seemed to lead to the end of the new network after losing $100,000 only in the first month. In 1928, William S. Paley bought CBS and in just one year was able to turn the tide and make CBS a profitable permanent radio network. CBS achieved permanent status by September 1929 (Sterling and Kittross 2001).

Despite some difference in their internal organization, NBC and CBS had the same business model. Both networks were formed by one flagship station and a number of affiliated local stations across the country. In 1929, the first year of full-blown operation, NBC had 56 affiliated stations whereas CBS 27. Often content was produced in the same city where the flagship station was located (usually New York City or Chicago) and then broadcasted live by network affiliates. Radio networks had obvious economies of scale in producing radio programs while enjoying major sponsoring from the growing advertising industry (Scott 2008). As advertising on radio became more systematic, networks were in the ideal position to provide advertisers with an unprecedented audience size. In turn, the increase in revenues from advertising made investing in high quality shows and programs very profitable, enhancing the quality of radio network schedules and the rise of the *Golden Age of Radio*.

From the early 1930s, radio networks became the most important producers and distributors of mass culture in the US characterized by an increase in the homogenization of the programs consumed by Americans. A distinctive feature of the networks was the steep decline in diversity and rise of homogenization on the air. This led to controversial opinions from contemporary on the potential benefits and costs of the networks. On the one hand, industry and government leaders saw radio networks as a cultural project that could help homogenize and modernize a country that due to the high rates of immigration looked very fragmented in the early 1920s (Sterling and Kittross 2001). While in the earliest years of radio, stations were small and their orientation was distinctively local, the arrival of the networks completely changed this picture. Broadcasting became a corporate business with a national focus. It gave citizens from different backgrounds, origins or social status a common cultural experience making them part of a larger national "imagined community" (Cohen 1990). Unlike newspapers that until then reinforced local parochialism, networks turned America into a land of listeners, entertaining and educating with no distinction, assimilating all despite age, race and class into a common culture (Gordon 2017; Hilmes 1997). By World War II, broadcasting had grown into a national medium for culture, information and entertainment that helped contribute to the "nationalization" of American culture (Lippmann 2008).

On the other hand, the rise of networks was accompanied by criticism from contemporaries regarding the disappearance of local identities. Some critics feared that concentrating the power of modern, network-based broadcasting in the hand of few corporations might have represented a threat to enliven US democracy (Lenthall 2008). Others worried that local uniqueness would be lost through the increasing imposition of the urban, North-Eastern mass culture across the radio waves into small towns and American homes (Lippmann 2008).

2.2 Baseball and Second Generation Players on the Networks

Early in the 1900s, baseball was already considered the *national pastime*. Contemporaries saw it as the perfect incarnation of the American values of competition, fair play and opportunity (Levine 1993). Before the end of the Age of the Mass Migration, when immigration from Europe was virtually unrestricted, baseball was also seen as a way to teach children of immigrants how to become proper citizens. In 1923, Fried Lieb, president of the Baseball Writers' Association of America, concluded that "there has been no greater agency in bringing our different races together than our national game, baseball. Baseball is our real melting pot" (Baldassaro 2011, as cited in p. XXXVI).

The 1920s was a revolutionary decade for baseball fans as radio changed the way games could be experienced. Stations began to broadcast baseball games live, allowing fans who could not attend the games in person to follow their favorite teams and players. Despite the increased excitement about the game, initial audiences were still predominantly local. For example, back in 1921 only about 2,000 listeners tuned in to follow the Red Sox playing and they all resided in the greater Boston area (Halper 2017).

The arrival of NBC and CBS marked instead a key discontinuity in two dimensions: access and experience of the games. In terms of access, the rise of the networks killed the distance between the location of the games and the audience. By some estimates, more than twenty million people tuned in for the World Series games, listening via the network-affiliated stations that carried the games. Increased access from coast to coast also meant that some of the local baseball stars became national personalities to the rank of proper radio stars (Halper 2017).

In addition to increase access, networks also enhanced the way baseball fans experienced the games. Nationally renowned network announcers like Graham McNamee, Major J.

Andrew White, or Edward "Ted" Husing were professionals that could describe the scene at the ballpark so realistically that listeners felt as if they were personally witnessing the plays of their favorite players (Halper 2017).

The arrival of the networks also coincided with a larger presence of second generation immigrants playing baseball professionally. Despite derogatory stereotypes prevalent in American society against immigrants of recent arrivals, baseball managed to partially overcome these barriers, enabling the inclusion of several players descendent of immigrants at the professional level. The iconic sports magazine Sporting News declared "[...] 'nationality' is never a matter of moment if he can pitch, hit or field" (Levine 1993, as cited in p. 3). However, discrimination persisted and the life of early baseball players was not easy. According to Baldassaro (2011), immigrant baseball in the early 1900s were "unwitting pioneers." Nonetheless, the relatively lower discrimination in baseball provided an avenue for young second-generation immigrants from recent immigrant groups to commence playing and actively participate in larger numbers throughout the interwar period.

As more second generation immigrants played baseball professionally, they also started to appear more systematically on the networks. Thanks to the wider audience that radio networks provided to baseball, players of immigrant origin were able to slowly change the perception that society had of them. While initially praise mingled often with stereotypes, over time Italian, Jewish or Eastern European second generation immigrants in baseball became part of the public consciousness as something other than their respective negative stereotypes (Baldassaro 2011).

The broadcasting of players of foreign origin in baseball has been associated with two key changes in the perceptions of society. Firstly, their higher presence on the media had a ripple effect within the immigrant communities of similar origins. Fellow immigrants from new sending countries started paying more attention to the game forming a new, deeper connection with it. They finally saw one of their own at the center of the stage of the national pastime, making them both proud of their heritage and giving them some sense of belonging in the country (Baldassaro 2011).

Secondly, through their success, endeavors and their demeanors in the American National Game, players of immigrant origin began to break down stereotypes that plagued immigrants, especially those pertaining to the second wave of immigration. Even tho slowly over time and while stereotypes were still present, these players were sending a subtle message to the public that regardless of ingrained biases, these children of immigrants were succeeding in the *national pastime*, and becoming more and more American in the process. These players created a generation of players of immigrant origin which arose to the status of heroes for the immigrant communities. For instance, Joe DiMaggio embedded the perfect example of the immigrant who made it and was at the same time ethnic hero and American icon (Baldassaro 2011). Similarly, amidst rising anti-semitism in the US and especially in Europe, the triumphs of Jewish sons of immigrants carried special symbolic importance for American Jews concerned with American acceptance and Jewish survivals (Levine 1993).

3 Data

This section provides details on data sources and the construction of the variables used in the analysis. To investigate the effect of radio networks on immigrant assimilation, I link variation in radio networks coverage to measures of assimilation constructed from the US full count census for the period 1870-1940 (Ruggles et al. 2021). I reconstruct radio networks coverage and consistently crosswalk all county-level variables to the 1930 county boundaries (Ferrara et al. 2021).[7] Accordingly, I crosswalk all individual-level data from the US census waves to the 1930 county boundaries (Berkes et al. 2023).[8]

3.1 Measures of Assimilation

Leveraging the US full count census, I follow repeated cross-sections of immigrants in the US before and after the expansion of radio networks in 1929 and construct four measures of assimilation: intermarriage with natives, naturalization, naming patterns and English proficiency. For naming decisions, I construct an additional sample of birth cohorts entirely from the 1940 census waves, which allows me to exploit yearly within households variation.

Intermarriage. The decision to marry a local is considered one of the strongest forms of assimilation by social scientists (Gordon 1964; Fouka et al. 2021). It does not only capture the desire of the immigrant to integrate but also involves the acceptance of native families. Using information on marital status and nativity from the census, I identify repeated cross-sections of married immigrants between 1880 and 1940. I define intermarriage as such if an immigrant married a US-born spouse and focus on immigrants in wedding age (between 15 and 35 years old) who are likely to have married recently.

The ideal definition of exogamous intermarriage would require to condition also on the parental country of origin. Unfortunately, in the 1940 census wave parental nativity is only available for a small sub-sample composed by individuals who were sample lines or that

[7]Ferrara et al. (2021) provide several weights to crosswalk county-level data; I use the weights based on population which the authors show to be the most accurate.

[8]The match rate between the samples I construct and the crosswalks provided in the Census Place Project ranges between .95 and .99 depending on the data I link.

were still residing with their parents. This stronger measure of exogamy is only well defined for 5.4% of my 1940 intermarriage sample.[9] Given the small sample size and the sample selection introduced by those individuals who resided with their parents even after marriage, I focus on the measure of intermarriage solely based on spousal nativity status.

Naturalization. A natural outlet of immigrant assimilation is becoming a citizen. From the early twentieth century, the Bureau of Immigration and Naturalization standardized the pathway to citizenship for immigrants, predominantly processing naturalization cases through federal courts (Bolger 2013). Upon arrival or shortly thereafter, immigrants typically submitted a Declaration of Intention, often referred to as "first papers." After a period of at least two years, they became eligible to file a petition for naturalization, known as "second papers," marking the final step before the court's formal approval of the naturalization process. Requirements for naturalization were essentially a five-year residency condition and the provision of two witnesses. Rejection rates were particularly low, hence naturalization was more likely to reflect the willingness of the immigrant to become US citizen rather than courts' willingness to grant the citizenship (Fouka et al. 2021).

By 1922, the US congress further facilitated naturalization for aliens who had been member of the US army and for married women (Bolger 2013). The new procedure of naturalization got rid of the need for a Declaration of Intention and allowed immigrants to directly apply for naturalization. This resulted in a significantly faster process. In addition, the five-year residency was eliminated for men who had served in the army and it was shortened to one year for married women.

Starting from the 1900 wave, the US census began asking to adult immigrants whether they had been naturalized US citizens. Using this information, I construct an indicator equal to one if an immigrant formally became a US citizen. While this information is available for all male immigrants aged at least twenty years old, it is only available for female immigrants starting 1920. Accordingly, I construct a sample of repeated cross-sections of male immigrants between 1900 and 1940 and female immigrants between 1920 and 1940.

Naming Patterns. First names have an important cultural component and can help identify parental ethnic, racial or cultural preferences (Lieberson and Bell 1992; Gureckis and Goldstone 2011). I use the degree of foreignness of a first name as a proxy for the parental taste towards assimilation into the US society. Exploiting the US full count census,

[9] Before 1940, on average every three intermarriages based on spouse nativity is also an intermarriage as defined by parental nativity. This ratio drops to 2% in 1940 due to the restricted coverage of parental birthplace in the 1940 census.

I construct the *foreignness name index* (FNI) which computes how distinctively foreign a name is (Fryer and Levitt 2004; Fouka 2019a; Abramitzky et al. 2020; Fouka et al. 2021).

The FNI is the standardized relative probability a name is found among foreigners over natives. To calculate the FNI of a name from a given birth cohort, I use names of individuals born in the previous twenty birth cohorts (Abramitzky et al. 2020). The index ranges between 0 and 100, with names only found among foreigners having a value of 100. I take the inverse hyperbolic sine transformation of the FNI to reduce the weight of outliers (Fouka 2019a).[10] The scale of a variable matters when transforming it with IHS. The average value of the raw FNI is large enough (55) to ensure that the scale is not an issue in this context (Bellemare and Wichman 2020). I also show in the Appendix that I obtain similar results with the raw FNI or using logs.

To study naming patterns I construct two complementary datasets. The first one mirrors the rest of the outcome variables and relies on decennial variation from the US census. First names are easily accessible starting from 1870. Thus, I compute the FNI for all cross-sections of newborns from immigrant households born on census year from 1870 to 1940. I define an immigrant household as one where both parents were foreign-born. Table A.5 gives an overview of the most and least distinctively foreign names based on the 1930 birth cohort by sex.

The second dataset I employ to study the impact of radio networks on immigrants' naming decisions leverages pseudo-birth cohorts entirely extracted from the 1940 census. I identify all children born in the US to immigrant parents who continue to reside in the parental household. This process allows me to construct birth cohorts of second-generation immigrants born between 1922 and 1940 who were still alive in 1940. By employing this approach, I am able to track naming decisions within the same household on an annual basis, both before and after the advent of radio networks.

English Proficiency. Learning the local language has economic returns in the labor market and can facilitate social interaction with natives (Bleakley and Chin 2004; Ward 2019). In decades between 1900 and 1930, the US census asked immigrants at least ten years old whether they were able to speak English. I use this question in the census to track repeated cross-sections of immigrants older than ten, from non-English speaking countries and study

[10] The exact formula to compute the relative probability of a name being foreign is as follows:

$$R_{name} = \frac{\frac{\#Native_{name}}{total\#Native}}{\frac{\#Foreign_{name}}{total\#Foreign}};$$

Then the FNI is standardized to be within 0 and 100: $FNI_{name} = 100 \cdot \frac{R_{name}}{1+R_{name}}$.

the impact of radio networks on their self-declared English proficiency.

3.2 Data on Baseball

I employ data on baseball to study whether immigrants, through radio networks, used baseball players names to inform their naming decision. I recover data on baseball from Friendly et al. (2019).[11] For this exercise, I rely on the sample of immigrant newborns I construct from the 1940 census wave which allows me to observe yearly naming patterns. Accordingly, I gather names of baseball players active between 1922 and 1940 and construct an indicator equal to one if an immigrant child born in a given year has the same name of a player active in that year.

In addition to their given names, I characterize players along two dimensions: their yearly performance and their perceived ancestry. Although not all players born in the U.S. were second-generation immigrants, the presence of a last name associated with the immigrant's country of origin could have spurred the perception that they were part of the second-generation immigrant community. In the data, out of 2762 players, only 55 were foreigner. For players born in the U.S., I proxy their ancestry by considering the frequency of their last name in the immigrant population in the 1920 full count census.

To determine baseball players's performance during a season, I rely on two types of measures. First, I measure outstanding performance in a season by looking at whether the player received any award related to his performance.[12] Second, I try to disentangle both positive and negative performance in a year through the *wins above replacement* (WAR) statistic. WAR is a statistic that evaluates the overall contribution of a player to his team. A WAR value is supposed to measure a player's value in all facets of the game by determining how many more wins he or she's worth over a replacement-level player (a player who may be added to the team for minimal cost and effort) at her or his same position (Major League Baseball 2024). Differently from prizes, WAR is a continuous variable that allows me to detect also negative performance. I collect data on historical WAR statistics for all players in my sample (Baseball Reference 2024), link it to each player and compute for every season they were active indicators of positive and negative performance.

Finally, I collapse the data at the name-by-year level to link it to my sample of newborn immigrants born between 1922 and 1940. I collapse players performance at the name-by-year level weighting the performance of a given name by the share of games played by the players

[11] The Lahman baseball data is accessible here.

[12] Awards I can observe in the Lahman data are whether the player was featured as part of the Baseball Magazine All Star team, or the TSN All Star team, the TSN MVP or the overall number of awards received in a year.

with that name in that year. I do so to account for the differential game time and proxy for the differential presence on the networks. I drop players who did not play and for which a WAR value is not defined.

Differently from performance, I cannot weight perceived ancestry by the number of games played in a season. Hence, I assign to a name in a year the country of origin of the player who played the largest number of games within that year. That way, I recover a name-by-year dataset with information on what country of origin most likely is associated to each player name in the data.

3.3 Control Variables

I use the full count US census and other data sources to construct a host of control variables at the county-level to complement the immigrant individual-level samples. I measure these controls as of 1920, my baseline year and the last census decade before the rise of radio networks.

From the full count census (Ruggles et al. 2021) I obtain several socio-economic characteristics measuring demographic features (eg share black or share foreign-born) as well as variables gauging the local economic structure (eg share in agriculture, share unemployed); I include further variables not included in the census on the local economy, in particular data on manufacturing, from Haines et al. (2010). I complement socio-economic characteristics by adding data on 1920 electoral outcomes, i.e. the republican vote share and the total electoral turnout (Clubb et al. 2006).

To perform robustness checks I source data on the Great Depression and New Deal spending from Fishback and Shawn (2018), while data on exposure to immigration quotas comes from Ager et al. (2023). The paper also leverages a wide range of geographic controls that I borrow from Bazzi et al. (2020).

In Table A.1, I present county-level descriptive statistics of my baseline county sample, detailing the main socioeconomic characteristics alongside the four measures of assimilation as of 1920. The sample is limited to U.S. counties with complete data on all outcomes.

Appendix Figure A.3 illustrates the temporal evolution of the outcome variables. The graphs depict a consistent trend across decades, wherein social assimilation exhibited a decline until the 1910s and 1920s, followed by a subsequent increase somewhere between the 1910s and the 1930s.

Table A.2 provides descriptive statistics for the pseudo birth cohorts sample constructed from the 1940 census wave, while Table A.3 outlines the characteristics of the baseball dataset.

3.4 Radio Networks Coverage

I reconstruct the rise of US radio networks by collecting detailed data on radio antennas operating in 1929 and feeding it to a model of radio signal propagation appropriate for AM radio, the only radio technology available at the time.

For each transmitter, I digitize data on its location (geo-referenced at the city centroid), height, power of antenna and frequency of operation. The source of the data is the historical magazine *White's Radio Log*. I digitize the January 1929 edition of such magazine which provides information on all active radio stations by network affiliation vetted by the Federal Radio Commissioners. Some stations have no data on height, in such cases I impute it using the average antenna height in the data. In addition to technical specifications, I recover information on whether the radio station was affiliated to one of the two radio networks active at the time: CBS or NBC. Figure A.2 shows the location of network affiliated radio stations across the US in 1929, the first year of radio networks full-blown operation. Not surprisingly, the most powerful antennas surged in the Northeast and the Midwest.

During the initial expansion of the radio networks, radio stations broadcasted exclusively on AM frequencies. Similarly to the FM radio stations, the quality of the AM radio transmission is determined by the characteristics of the antenna such as power and frequency and by the distance between transmitter and the receiver. However, differently from FM, the ground waves produced by AM radio travel farther than line of sight and are less affected by the presence of physical obstacles between transmitter and receiver. The main determinant of AM radio coverage is instead the degree of soil conductivity, that is the ability of the ground to conduct electromagnetic waves (Kirby et al. 1954).

Ground conductivity is the electrical conductivity of the subsurface of the earth and it is an extremely important factor in determining the propagation of AM ground radio signal (Kirby et al. 1954). Higher levels of ground conductivity allow radio waves to travel farther. Saline water is the most conductive surface, wetlands and agricultural lands have generally good levels of conductivity whereas desert land has typically the lowest conductivity observed in nature. I recover ground conductivity from the Federal Communication Commission (FCC). Figure B.8a shows the original map and the same map digitized and averaged at the county-level.[13]

[13] While the source of ground conductivity is not contemporaneous to the rise of radio networks, FCC considers ground conductivity to be constant over time to a first approximation and provides users a map from 1975 to predict AM radio coverage today. The original map and information provided by the FCC can be found at this link.

I use the radio stations data with a rigorous AM radio signal propagation model to compute radio networks coverage at the county-level. To do so, I borrow the LFMF radio propagation model from the International Telecommunication Union (ITU). The LFMF model works differently from the standard model of propagation used in the media literature so far because it is specifically designed to predict ground radio waves for frequencies between 10 kHz and 30 MHz, the range of frequencies used by AM radio.[14] In online appendix B, I give further details on AM radio and how I employ the LFMF model to construct radio networks signal coverage.[15]

The output of the radio propagation model is a continuous variable expressing signal strength value (field strength) measured in decibels referenced to 1 microvolt per meter (dBμV/m). I compute county-level radio networks coverage by taking the maximum signal strength value over all network-affiliated radio stations reaching a county. I also compute non-network radio coverage similarly.

The exact relationship between signal strength and quality of reception is hard to define as it depends on many factors which are difficult to measure systematically, and especially so in the historical context at hand. For example, urban households typically needed higher signal strength than rural households to achieve the same quality of broadcast (Federal Communication Commission 1947).[16] Hence, I rely on continuous variation in signal strength throughout the analysis (Enikolopov et al. 2011; Yanagizawa-Drott 2014; Durante et al. 2019; Wang 2021). To simplify interpretation of the results, all variables measuring signal strength are expressed in standard deviations.

4 Empirical Strategy

In this section, I present the empirical relationship of interest and introduce the source of variation used to estimate it. I then provide evidence substantiating a causal interpretation of the results.

The primary focus of this paper is to uncover the impact of radio networks on immigrant assimilation. Estimating such relationship in a naive OLS framework is prone to important

[14]The usual radio signal propagation model, first used in economics by Olken (2009), is the Longley-Rice Irregular Terrain Model (ITM) and it is a model for propagation in higher frequencies, those typically employed with FM radio or TV broadcasts. A notable exception is represented by Armand et al. (2023) who outsource the computation of the AM radio signal to an engineering firm.

[15]The original software from the ITU is available at this link.

[16]The most important local characteristics determining the minimum level of usable field strength are the quality of the radio sets and the environment surrounding the receiver, that is urban status. Especially historically, the quality of radio sets was not as homogenous as today, further complicating the task of pinning down the exact threshold for good quality reception. Typical sources of interference in the 1920s were the overlap of multiple radio signals, electric street cars or the presence of telephone ringers.

concerns. Indeed, radio networks signal strength might have been correlated with other factors that also influenced changes in immigrant assimilation. One of the most notable examples is the role of urbanization, which in the interwar period attracted the largest and most powerful stations. Agglomeration and urbanizations might have also affected assimilation trends independently of the impact of radio networks, acting as a confounding factor. Thus, to estimate the causal effect of radio networks on immigrant assimilation, I implement an identification strategy based on recovering quasi-exogenous variation in radio networks signal coverage.

To build a source of identifying variation, I leverage local topography as a quasi-random determinant of signal strength and use it to derive the causal impact of radio networks on immigrant assimilation (Olken 2009). I implement this strategy in three steps, following recent development in the media literature that I carefully adapt to the AM radio setting (Durante et al. 2019; Wang 2021). First, I use the LFMF model described in the previous section to reconstruct geographic variation in radio networks coverage. Second, I simulate an hypothetical radio networks coverage with ideal conductivity. Following the media literature, I denote such variation as the *free* signal strength variation. When constructing the hypothetical network coverage, I hold the rest of the parameters identical to the first step and only change the conductivity value. Finally, I regress assimilation outcomes on the actual radio networks variation ($Network$) conditional on the hypothetical signal strength in the absence of topographical obstacles ($Free$). Controlling for the hypothetical network signal, variation in actual signal strength is driven by differences in the ground conductivity mix *between* transmitting antennas and county centroids. This is the key identifying variation exploited throughout the paper.

To strengthen further the identification strategy, I control for county fixed effects and local geographic characteristics interacted with decade fixed effects. This allows me to absorb the direct impact of local geographical features, including ground conductivity, ensuring that the residual variation I exploit is driven by the soil conductivity mix the signal crosses on its way to the receivers and not the conductivity in the receiving location.

Figure 1 presents the key sources of variation of the paper in three panels. Panel A shows the actual radio networks signal strength, panel B the residualized signal strength after partialing out hypothetical network signal and geographic controls and panel C the underlying soil conductivity. Panel A of Figure 1 suggests that radio networks coverage was highest around large urban centers, especially in the Northeast and the Midwest where the most powerful stations were located. Panel B and Panel C instead show a completely different picture. A quick visual inspection of the two maps suggests that the residualized radio networks variation is less dependent on the location of urban areas and it appears to be driven by the underlying ground conductivity.

[Figure 1 about here.]

Equation 4 formalizes the linear regression model I employ to study the impact of radio networks on immigrant assimilation. I incorporate the conditional variation in network signal strength shown in panel B of Figure 1 in a standard difference-in-difference design. The model compares assimilation outcomes of immigrant cross-sections observed before and after the rise of radio networks in 1929, and between counties with different intensity of conditional network signal strength. The main coefficient of interest in the equation is β and it captures the increase in a dimension of immigrant assimilation given a one standard deviation increase in radio networks coverage, everything else equal.[17]

$$Y_{ict} = \beta \; Network_c Post_t + \beta_f Free_c \alpha_t + \gamma X_c \alpha_t + \alpha_c + \alpha_o + \alpha_t + \varepsilon_{ict} \tag{1}$$

Y_{ict} is one of the outcome variables measuring assimilation for immigrant i in county c in decade t (eg an indicator equal to one if the immigrant is naturalized); $Network_c$ and $Free_c$ are the actual signal strength and the *free* signal strength in county c; X_c is a set of time invariant county-level controls that can include geographic controls, population, other socioeconomic characteristics and political returns as of 1920; α_c, α_o and α_t are county, country of origin and decade fixed effects respectively. ε_{ict} is the error term. More saturated specifications include the age of the immigrant and fixed effects interactions (eg region by census decade). I cluster standard errors at the county-level but also present results with standard errors clustered at the state-level and with standard errors accounting for spatial correlation in the data.

The standard identification assumption to assign a causal interpretation to the estimate of β is the lack of differences in the outcome pre-trends between treated and control group. While in the next section I present evidence consistent with this assumption based on an event study analysis, I start by providing evidence of exogeneity showing that the radio networks variation I reconstruct is orthogonal to the *levels* of the outcome variables in the pre-period.

Table 1 shows the null impact of the conditional variation in radio networks signal on the main assimilation outcomes as of 1920, the last decade before the onset of radio networks. In this exercise, I regress county-level variation in assimilation outcomes in 1920 on the radio

[17] A key point for the interpretation of the estimate of β is that it recovers an intention to treat parameter (Durante et al. 2019). This is a common caveat in the media literature since fine level variation in listenership is typically not available. Also in this case I am not aware of data on radio networks take up during the interwar period.

networks variation I reconstruct. The county-level networks variation is conditional on the same baseline machinery I use in the main analysis: *free* radio networks signal, geographic controls and population. To absorb some of the vast geographic differences in assimilation outcomes, I include state level fixed effects. All columns present estimated coefficients that are small and statistically non-distinguishable from zero. The result provides evidence that the same identifying variation I use in the analysis is orthogonal to levels of assimilation practices before the rise of radio networks.

[Table 1 about here.]

I provide further evidence that my identification strategy uncovers variation that is quasi-exogenous by showing that such variation is balanced on a wide range of socioeconomic characteristics in the pre-period. When performing such balancedness test, I pick characteristics that could be important correlates of media coverage related to the geographic distribution of the US population (eg population, population density, urbanization or the number of dwellings) or related to the local economic activity (eg literacy, unemployment and occupational score). Figure 2 shows that none of the twenty 1920 variables I regress on the conditional radio networks variation is significant at the five percent level and only one (percent manufacturing) is significant at the ten percent level.

In Table A.4 I show that these conditional estimates are statistically different from unconditional estimates of the impact of radio networks on immigrant assimilation. While the same table highlights that unconditional radio networks variation correlates with several observables, controlling for the hypothetical signal strength and geographic characteristics balances the estimates around zero. This result corroborates the assumption behind the identification strategy, that variation in conditional radio networks coverage is quasi-exogenous and that it is a fruitful source of variation to identify the impact of radio networks on immigrant assimilation.

[Figure 2 about here.]

5 Results

The advent of radio networks in the US followed the end of the Age of Mass Migration. Did the rise of the first nationally homogeneous broadcasting mass media affect the assimilation of millions of immigrants? In this section, I document that quasi-random exposure to radio networks had a first-order positive impact on immigrant assimilation. Immigrants exposed to radio networks were more likely to marry natives, become US citizens, adopt distinctively native names and speak English proficiently.

In this section I present the main result of the paper in the form of tables focusing on male immigrants. Similar results on female immigrants are shown in Appendix section A. I then provide evidence on the dynamics of the radio networks impact and the lack of pre-trends using an event study analysis. I conclude the section by presenting a wide range of robustness checks corroborating the stability of the results.

5.1 Radio Networks and Immigrant Assimilation

5.1.1 Intermarriage

[Table 2 about here.]

The starting point of the empirical analysis is the evaluation of the impact of radio networks on intermarriage between immigrants and natives. Table 2 presents the results of the analysis. Column 1 shows the most sparse specification where I only include baseline controls and baseline fixed effects. Importantly, column 1 does not include the *free* signal strength, which is a key requirement in the identification strategy. The estimated coefficient is slightly negative and imprecisely estimated.

Adding the hypothetical signal strength to the list of controls in column 2 changes significantly the results. The coefficient flips from negative to positive and its magnitude increases and becomes statistically significant at the 1% level. The point estimate indicates that one standard deviation increase in radio network coverage boosts the likelihood of intermarriage by about 6.3 percentage points (p.p.), or the equivalent of 13% of a standard deviation of the pre-treatment outcome variable.

The change between column 1 and column 2 highlights that the estimate in column 1 is biased downward. As already discussed in the empirical strategy section, naive variation in radio networks exposure is likely confounded by agglomeration and urbanization, which can have contrasting effects on immigrant assimilation. On the one hand, larger urban areas characterized by a wider menu of cultural and leisure activity might have fastened assimilation. On the other hand, the wider range of cultural options might have allowed immigrants and natives to sort into activities characterized by higher homophily, resulting in a slower assimilation process. The overall net effect is thus consistent with the close to null impact estimated in column 1.[18] I find that all four measures of assimilation exhibit similar patterns of downward bias when I do not account for the *free* network signal.

[18]Even controlling directly for share urban, which is included in the baseline controls, might not be enough to rule out the confounding effect of agglomeration. Besides concerns on the limitations of linearly controlling for urbanization, the issue persists since within urban areas there is variation in the presence of radio network antennas, and the source of this variation is hard to pin down.

Columns 3 to 6 gradually include additional controls: geographic characteristics in column 3, age of the immigrant in column 4, additional county-level socio-demographics in column 5 and electoral returns in column 6. Reassuringly, estimates remain essentially unchanged throughout, pointing to a stable positive relationship between radio network exposure and intermarriage.

I replicate the analysis on female immigrants in Appendix Table A.6. Results are qualitatively similar to the ones I document on male immigrants except coefficients are approximately half the size. The most saturated specification shows that rising radio networks exposure by a standard deviation increases intermarriage between female immigrants and natives by 3 p.p., or 7% of a standard deviation.

The analysis of intermarriage patterns relies on a sample of immigrants that are below the age of 35 to avoid conflating the data with marriages likely happened before the rise of radio networks. I exploit the fact that older immigrants at census year are considerably less likely to have gotten married after the rise of the radio networks to run a placebo analysis. I rerun the analysis on intermarriage using the sample of immigrants that are 35 or older. Table A.7 shows that for both male and female immigrants the impact of radio networks on intermarriage is zero in the placebo sample. This exercise provides further evidence that the probability to marry a native only goes up for the immigrant population who was likely to marry after the rise of the networks.

5.1.2 Immigrant Effort to Assimilate: FNI and Naturalization

[Table 3 about here.]

I now turn to the analysis of naturalization and naming patterns. Differently from intermarriage which is an equilibrium outcomes depending also on the willingness of the natives to marry a foreigner, the behaviors analyzed in this section can be interpreted as immigrants' effort to assimilate (Fouka et al. 2021). Table 3 shows the estimated effect of radio networks on the likelihood an immigrant naturalized US citizen (Panel A) and on the foreignness name index (FNI) associated to the first names of immigrants' newborns (Panel B). In panel A the sample is composed by male immigrants above the age of 20, the minimum age to apply for naturalization, and in panel B by all male newborns at census year from parents both foreign-born.

Just like in the analysis of intermarriage, column 1 does not account for the hypothetical radio network coverage without topographical obstacles. In both panels, column 1 shows a null effect of radio networks on immigrant assimilation effort, consistent with a downward bias. As soon as I add the *free* signal strength in column 2, results become larger in magnitude and consistent with higher assimilation.

In panel A, the coefficient in column 2 more than doubles relative to column 1, indicating approximately a 2 p.p. increase in the likelihood of naturalization; despite growing in size, the coefficient remains statistically not significant. Moving to column 3, the addition of geographic characteristics further enlarges the coefficient and makes it highly significant at the 1% level. In columns 4, 5 and 6 both magnitude and significance stabilize showing an increase of 2.5 p.p. in the probability of naturalization, significant at the 1% level.

Panel B tells a similar story for the FNI. Including the *free* signal strength, I estimate a reduction of $-.077$ associated to the inverse hyperbolic sine (IHS) of the FNI, or about an 8% decrease in the raw index. In the most saturated specification in column 6, a one standard deviation increase in radio network signal strength lowers FNI by 9%.[19] I obtain these estimates by constructing the FNI using the last twenty birth cohorts prior to the cohort of interest (Abramitzky et al. 2020). Appendix Table A.9 shows that the conclusion of the analysis is robust to alternative numbers of birth cohorts or the functional form used to measure the FNI.

The magnitude of the coefficients in column 6 indicate an increase of naturalization of about 5% of a standard deviation of the pre-treatment outcome, and slightly more than a 12% decrease of the foreignness of the names adopted by immigrant households. Table A.8 in the Appendix presents the same results for female immigrants, once again pointing to qualitatively similar results although with smaller magnitudes. Radio networks increases naturalization for female immigrants by 5% of a standard deviation, whereas it reduces IHS FNI by 8% of a standard deviation.

The name-based assimilation induced by the exposure to radio networks is not small. Abramitzky et al. (2020) find that, during the Age of the Mass Migration, children of immigrant households who resided in the US for twenty years display an FNI that is on average 9 p.p. lower then newly arrived immigrants. Using this magnitude as a reference and considering the coefficient in column 6 of panel B of Table 3, one standard deviation increase in radio network coverage is comparable to anticipating the year of arrival for an immigrant family who will then have a male newborn by approximately 10 years, a significant fastening of name-based assimilation.[20]

[19]Bellemare and Wichman (2020) suggest that scaling can significantly affect the interpretation of estimates when using inverse hyperbolic sine transformations. The authors suggest transforming the data with IHS only if the average non transformed value is above 10. In the paper, the average FNI value is 55, well above the suggested cutoff.

[20]Evaluating the impact of radio networks at the average FNI for male immigrants (about 55) we get a reduction in FNI of 4.95 p.p. on average, which is more than half of the reduction accumulated in 20 years documented by Abramitzky et al. (2020).

5.1.3 English Proficiency

The fourth measure of social assimilation I investigate is the ability to speak English. The question regarding English proficiency is included in the US census until the 1930 wave, hence the scope of the analysis is reduced with respect to the other outcomes. Nonetheless, I provide evidence that for recently arrived immigrants, who arguably had the highest returns from exposure, radio networks significantly boosted English proficiency.

[Figure 3 about here.]

Figure 3 plots the effect of radio networks on English proficiency splitting the sample according to the range of years since first arrival.[21] In the analysis, I use the most saturated regression including the full set of controls similarly to column 6 from the rest of the tables in the section. The Figure shows clearly how immigrants who arrived in the US within the past five years benefitted significantly from being exposed to the networks in terms of English proficiency. Immigrants experiencing a one standard deviation increase in radio network coverage were more likely to declare they were able to speak English by 7 or 8 p.p. depending on gender. The same variation in coverage did not affect the ability of immigrants who arrived earlier in the country.

[Table 4 about here.]

Table 4 formalizes the main results on English proficiency for recent immigrants. In the Table, I once again show that the naive variation in radio networks has a downward bias that is immediately reversed once I control for the *free* signal strength. Adding the controls from columns 2 to 6 leads to a coefficient equal to .079 for male immigrants, which indicates an increase of English proficiency of about 16% of a pre-treatment standard deviation, consistent with the magnitudes presented for the other outcomes. Also in this case, female immigrants show similar but smaller assimilation response. Table A.10 shows that the most saturated specification points to an increase in the likelihood of speaking English correctly of 7.3 p.p.

5.2 Evidence from Event Study Analysis

To strengthen the causal interpretation of the main results just presented, I show graphical evidence that the network coverage variation I build does not show pre-trends, a key piece of evidence supporting the identification strategy. I do so by running an event study analysis

[21] To split immigrants across different years of arrivals, I employ a variable from the US census that splits immigrant arrivals in bins of five years.

of the four outcomes of the paper. Equation 2 shows the empirical model I estimate which mirrors the most stringent specification from the main analysis.

$$
\begin{aligned}
Y_{ict} &= \sum_{t=1870}^{1940} \beta_t Network_c D_t + \beta_f Free_c \alpha_t + \gamma X_c \alpha_t + \delta I_{ict} + \alpha_c + \alpha_o + \alpha_t + \alpha_{tr} + \varepsilon_{ict} \\
D_t &= 1[t = 1870, 1880, 1900, 1910, 1930, 1940]
\end{aligned} \tag{2}
$$

Since treatment is not staggered and all cross-sections post 1929 are continuously treated while all cross-section pre 1929 are control, event year is identical to calendar year. Running the event study in this case is equivalent to interacting decade fixed effects with the continuous variation in radio network signal strength. The natural reference year is 1920, the last decade before treatment.

In the equation, Y_{ict} is one of the four outcome variables. The coefficients of interest are the β_t which are associated to the impact of radio networks at decade t. D_t is an indicator equal to one in decade t, and interacted with the continuous variation in radio networks captures the dynamic impact of the networks with respect to the reference decade, in this case 1920. The rest of the specification is identical to the most saturated one in the main analysis. It includes the full set of geographic, socio-demographic and political controls, all interacted with decade fixed effects ($X_c\alpha_t$); crucially, it also includes the free signal strength interacted with decade fixed effects ($Free\_S_c\alpha_t$) and the age of the immigrant (I_{ict}). To further test the robustness of the results, I augment the specification by including decade by region fixed effects (α_{tr}).

Equation 2 suggests that I can inspect pre-trends all the way back to 1870. This is only possible for naming patterns. For naturalization and English proficiency, I can obtain data since 1900, whereas for intermarriage since 1880. Naturalization can be traced back to 1900 only for male immigrants due to sample restrictions in the census. Similarly, the US census does not report information on English proficiency in 1940 which limits my post-treatment analysis for that outcome.

[Figure 4 about here.]

Figure 4 displays the results of the event study analysis, where I plot the estimates of β_t. Panels a, b and c show a similar dynamic: radio networks produce no significant trend before treatment onset. In 1930, only one year after the rise of the networks, estimates start showing some sing of assimilation, although mostly not statistically significant. By 1940, the full effect of radio networks kicks in with estimates that are sizable in magnitude and statistically significant. The event study analysis highlights that the impact of radio networks does not kick in instantly but accumulates over time and persist for at least a decade.

Panel d of Figure 4 presents the event study analysis of English proficiency for immigrants who arrived in the country in the last five years. As shown in Figure 3, the impact is concentrated in recent arrivals for which exposure to radio networks should have the highest returns. Similarly to other panels, the graph documents the absence of pre-trends in the decades leading to treatment. Given the lack of information on English proficiency in the 1940 census wave, the post treatment coefficient is virtually identical to the one presented in the last column of Table 4. Consistently with what I showed in the previous section, all panels report similar dynamics of assimilation by gender.

Within Household Variation in Naming Patterns

An important shortcoming of the event study analysis presented above is that I can only leverage decennial variation. I circumvent this limitation by exploiting children of immigrants who resided in the parental household until the age of 18. Using the 1940 census wave and exploiting their year of birth, I reconstruct pseudo birth cohorts with yearly variation in naming patterns between 1922 and 1940 (Abramitzky et al. 2020). This allows me to compare the names of second generation immigrants from the same households born before and after the rise of radio networks. I implement a similar event study analysis to the one shown in equation 2 with the key difference that I swap county level fixed effects and county controls for household fixed effects. Including household fixed effects also implies the absorption of ancestry fixed effects.

[Figure 5 about here.]

Figure 5 plots the estimates of the event study analysis, where I pick 1928 as the reference year. Some estimates are noisy but the general picture depicts a clear decrease in the foreignness of the names adopted by immigrant households. After the onset of radio networks in 1929, the estimates turn negative, and slowly increase their magnitude. Starting 1933 and onward, the coefficients are consistently negative, and statistically significant with the exception of 1939. While the exact magnitudes between this exercise and the one using repeated cross-section are not the same, they are consistent with each other for male immigrants. Crucially, Figure 5 brings supporting evidence for the lack of pre-trends in FNI before the rise of radio networks. The graph highlights that all estimated coefficients before 1929 hover around zero with no particular pattern.

While the results for male immigrants indicate a decrease in the FNI, results for female immigrants are less stark. Appendix Figure A.4 shows that the estimates are consistent with a lack a pre-trend but show less clearly whether in the post-period there is a decrease in the FNI. Indeed the difference-in-difference coefficient associated to the event study analysis is negative (-0.017, s.e. 0.015) but it is not statistically different from zero.

Overall the evidence in Figures 4 and Figures 5 strengthen the assumption that the radio network variation I reconstruct is as good as randomly assigned across US counties, and, thus, the estimates can be interpreted causally.

5.3 Robustness

This section provides evidence assessing the robustness of the main analysis. Firstly, I show that other important historical shocks happening during the interwar period are not confounding factors. Then, I openly address concerns related to urbanization and agglomeration as potentially driving both assimilation and radio networks coverage. Third, I check that different sample selections do not alter the results of the analysis. Fourth, I test the robustness of the main results to the inclusion of potentially bad controls and other types of fixed effects. Finally, I provide evidence that alternative methods of calculating standard errors leaves the significance of the estimates stable.

There are several historical events that characterize the American interwar period. In principle, there is no reason why the radio networks variation I reconstruct should be confounded by any of these events. Nonetheless, some of these historical shocks have been associated to the assimilation of immigrants or had large economic repercussions. To rule out that any of these events might be a confounding factor in my analysis, I sequentially include them in the analysis and verify the stability of the radio networks estimate. Tables A.11 and A.12 present the results of the robustness checks. To ease comparisons, column 1 displays the estimate in the most saturated regressions in the main analysis (column 6 of Tables 2, 3 and 4).

A recent contribution shows that the influx of African Americans during the Great Migration fastened the assimilation of European immigrants (Fouka et al. 2021). Columns 2 and 3 of both Tables include the change in the share of African Americans in the population between 1910 and 1930 (total and male only). The radio network estimates remain unchanged eliminating concerns over the influence of the Great Migration on the link between radio networks and assimilation.

Ager et al. (2023) argue that immigration quotas enacted in the 1920s decreased mortality rates mostly due to the living conditions of immigrants. Differential exposure to quotas might have altered the composition of immigrants and, possibly interact with the assimilation spurred by radio networks. I include county-level variation in the exposure to the quotas interacted with the post indicator in column 4. The coefficient associated to radio networks remains virtually identical, eliminating concerns related to the confounding effects of the quotas.

Columns 5 to 13 of Tables A.11 and A.12 rule out, using different measures, that the Great Depression might have confounded the effect of radio networks on immigrant assimilation. I sequentially include several measures of the severity of the Great Depression (columns 5 to 11) and county-level variation in New Deal spending (Fishback et al. 2006; Fishback and Shawn 2018).[22] For most estimates differences appear only after the second digit.

Finally, column 14 presents results controlling for the coverage of radio stations not affiliated to any of the national networks. Once again, the coefficient estimated associated with radio network coverage is essentially unchanged with respect to the estimate presented in Column 1. The result establishes that radio networks variation is not confounded by generalized radio access.[23]

A different source of concern relates to the role of agglomeration forces determining the location of radio stations and simultaneously affecting assimilation. My identification strategy tackles this concern by conditioning on the hypothetical signal which also takes into account the distance to the station responsible for the coverage in a given county (Durante et al. 2019). However, the function generating the hypothetical signal strength is highly non linear, and its inclusion in the regression model might not be sufficient to account for the distance to the transmitting antennas. Thus, I test the robustness of the main results to the inclusion of a set of distances to the nearest stations or relevant nearby hubs.

In Table A.13, I investigate how robust are the main estimates to the inclusion of the following distances interacted with the post dummy: distance to the closest antenna transmitting network radio in column 2; distance to the closest non-network radio station in column 3; column 4 swaps distance to non-network transmitter for the five closest antennas regardless of network affiliation; finally, I separately include distance to the state capital in column 5 and to Washington, DC in column 6. Across all four panels and all six columns the estimated coefficients remain virtually unchanged with respect to column 1, my preferred and most saturated estimate from the main analysis. These results, together with the balancedness tests on 1920 pre-determined observables, help ruling out concerns pertaining to the confounding effect of proximity to urban areas.

[22]More precisely, severity of the Great Depression is typically measured as the change in retail sales. I look at the change in retail sale per capita between 1939 and 1933, 1933 and 1929, 1939 and 1929; and as the change in the log sales per capita (growth rate) between 1933 and 1929, 1939 and 1933, 1939 and 1929, 1935 and 1929. New Deal spending is measured as the per capita spending in AAA grants between 1933 and 1937 and per capita AAA grants from 1933 and through 1935.

[23]Non-network radio coverage has no impact on immigrant assimilation. Results available upon request.

In Table A.14, I show that alternative sample selections do not change significantly my results. In particular, I replicate my main results excluding first the geographic US South and then states that only recently were incorporated in the US, ie New Mexico and Arizona. Column 4 shows that the results are robust also to the exclusion of the counties with an active antenna transmitting network radio station from the estimating sample. This last piece of evidence further corroborates the hypothesis that the impact of radio network on assimilation is not just an urban phenomenon and there is little heterogeneity depending on the presence of antennas.

Tables A.15 present results that advocate for the robustness of the main results to the inclusion of additional individual level controls and additional fixed effects. Column 1 shows the estimate from the most saturated specification in the main analysis. Columns 2 through 8 sequentially add: number of family members, employment status, occupational score, home ownership status, urban status, region by decade FEs and ancestry by decade FEs.[24] In panel C where the sample is composed by immigrant children, employment and occupational score refer to the household head. The individual level controls represent potentially bad controls as they could be affected by radio networks in the post period. Despite that, adding them leaves the coefficient estimates across the four panel virtually unchanged from column 1. In columns 7 and 8, I include region by decade and ancestry by decade fixed effects to absorb variation across regions or ancestries over time. Once again adding such fixed effects does not alter significantly the point estimate.

The last robustness check I perform pertains to inference. All the results of the main analysis in section 5.1 have standard errors clustered at the county-level. Table A.16 and A.17 show that the significance levels of these estimates hold similarly when I cluster at the state-level or when I use spatially corrected standard errors (Conley 1999).

6 Role Models, Performance and Assimilation

During the interwar period, radio networks played a significant role in fostering immigrant assimilation, with role models featured on these networks serving as a crucial channel for this process. In this section, I provide supporting evidence for this claim, focusing on the live broadcasting of baseball, a prominent segment of the Golden Age of Radio. I argue that baseball players who shared ancestral origins with immigrants and those who achieved success were important drivers of name adoption among immigrants. Additionally, I find that natives also altered their naming decisions in response to the performance of players

[24] In US historical censuses income is not available until the 1940 wave. Occupational score refers to the median total income in that particular occupation.

featured on the networks. This suggests that media role models may have contributed to the reduction of negative stereotypes held by natives towards immigrants.

A sizable literature in economics and in the social sciences suggests that role models are a key driver of societal change (La Ferrara 2016; Bandura and Walters 1977; Bandura et al. 1999). Historical records suggest that during the interwar period, numerous baseball players rose to national prominence, exerting a significant impact on societal norms and customs. Notably, this period also witnessed the emergence of many second-generation immigrants from recent immigrant groups as professional baseball players (Baldassaro 2011; Levine 1993). Building upon this historical evidence and insights from the media literature on role models, I test the hypothesis that the ascent of these new popular players from historically discriminated immigrant backgrounds, coinciding with the expansion of radio networks, played a crucial role in creating national media role models for immigrant assimilation.

The first and more direct impact of media role models is on the aspirations to assimilate of immigrants. To empirically assess whether baseball players featured on the networks served as media role models for immigrants, I draw upon insights from social psychology literature to identify key dimensions associated with role models. According to social psychologists, role models are typically perceived to be similar to their target audience (identification) or to reach their objectives successfully (aspiration) (Lockwood and Kunda 1997; Bandura and Walters 1977; Bandura et al. 1999). Leveraging variation in these role modeling characteristics, I employ the adoption of first names as an indicator that parents viewed these players as role models of assimilation for their offspring (La Ferrara et al. 2012). To measure these dimensions, I utilize rich historical data on baseball players (Friendly et al. 2019). Specifically, I use the perceived ancestral origin of the player, proxied by their last name, to gauge whether immigrants felt represented by a player. Additionally, I identify successful players using various indicators of performance during a season.

I implement a linear probability model to estimate the likelihood a children born in a given year was named after a baseball player active in the same year depending on radio network exposure interacted by the characteristics of a player's name. Similarly to the FNI analysis, I rely on birth cohorts of naming patterns between 1922 and 1940 extracted by the 1940 census wave. This is my preferred specification and data to investigate naming patterns since it allows me to exploit yearly variation and to include household fixed effects. In practice, I compare yearly naming decisions within the same household before and after the rise of radio networks. I then expand the model by interacting radio networks variation with indicators of performance or ancestral origin associated to a player's name and determine what characteristics of the players amplified or curbed the impact of the networks on naming decisions.[25]

$$
\begin{aligned}
Baseball_{iht} = \beta_0 \; Network\_S_c Post_t + \beta_1 \; Network\_S_c Post_t D_{it} + \\
\beta_f \; Free\_S_c \alpha_t + \gamma \; ihsPOP\alpha_t + \alpha_{hh} + \alpha_t + \varepsilon_{iht}
\end{aligned}
\tag{3}
$$

Equation 3 formalize the model I estimate. The parameters of interest are β_0 and β_1 which capture the own impact of radio networks and the differential effect of the name characteristics on name adoption, respectively. D_{it} measures the imputed ancestry or the performance level associated to the name. Controls mirror the analysis on FNI in the previous section, accounting for IHS 1920 population and the *free* radio network signal strength interacted with year fixed effects. I also explicitly control for the share of games played in a season to directly account for the differential exposure that different names had in the games and, hence, on the media.

Impact of Players' Ancestry. I lead the analysis by showing how the match between the perceived ancestry of baseball players and the ancestry of immigrant households affects immigrants' naming decisions. Table 5 displays the results of the analysis focusing on immigrant households composed by both parents foreign-born having a male newborn. Column 1 shows that the own impact of radio networks is small and not statistically significant. However, column 2 shows a large positive effect on adopting a player's name when the country of origin of immigrants matches the perceived ancestry of the baseball player. Specifically, I find that increasing radio networks coverage by 1 standard deviation boosts the likelihood immigrants named their child after a baseball player by 3.2 p.p. if immigrants and players shared a common ancestry.

[Table 5 about here.]

Column 3 and 4 of Table 5 investigate whether this result is driven by some ancestries in particular. According to the hypothesis underlying this mechanism, the results should be concentrated among baseball players from groups that faced xenophobia and discrimination during the early 1900s. This is what I find in column 3: the point estimate indicates that, facilitated by the networks, immigrants are more inclined to select names from players perceived to originate from Eastern Europe, Southern Europe, Germany, or Russia. Baseball

[25] Since there were multiple players for a given name in a given season, I collapse the ancestral and performance information at the name-level. When dealing with performance, I take the weighted average of the performance associated to a name weighing by the number of games played in a season. To determine the ancestral origin associated to a name, I pick the ancestry associated to the player who played the most games. Section 3.2 provides further details on the data construction.

names associated with these ancestries exhibit a 1.7 percentage point higher likelihood of being chosen by immigrants. Correspondingly, in line with the estimates in column 2 and 3, column 4 demonstrates a heterogeneous impact of the networks on immigrants from Eastern Europe, Southern Europe, Germany, and Russia. Immigrants from these countries are 2 percentage points more likely to select baseball players' names they hear through the radio networks.

To contextualize the findings from the mechanisms within the broader results of the paper, I present evidence indicating that the main results on assimilation exhibit a similar heterogeneous pattern across countries of origin. Figure 6 illustrates the outcomes of the analysis, where European immigrants are divided into eight macro regions. Although confidence intervals widen considerably, the magnitudes of the coefficients convey a consistent narrative. Across all four panels, immigrants from Eastern Europe, Russia, and Germany demonstrate the most pronounced response to exposure to radio networks, consistently ranking in the top three estimates for all outcomes. These findings reinforce the interpretation that radio networks had a substantial impact on the assimilation of immigrants from recent immigrant waves, aligning with the influence of media role models who affected the aspirations of similar immigrant groups.

[Figure 6 about here.]

Impact of Players' Performance. The second dimension through which baseball players might have acted as positive role models for assimilation is their performance. Together with identification, success is another pillar of role models that actively spur positive aspirations of emulation (Bandura and Walters 1977; Bandura et al. 1999). I test whether baseball players performance affected the naming decisions of immigrant households in Table 6. Column 1 once again indicates that the average impact of the networks is not significantly different from zero. In column 2, I interact the effect of radio networks with the index of successful baseball names based on players receiving any award during the season. The estimate associated with the interaction term suggests that immigrant households were 2.9 percentage points more likely to select names of baseball players they heard on the networks if those names were linked to particularly successful players. Appendix Table A.19 presents similar findings when measuring success with specific accolades such as being featured in the Baseball Magazine All-Star list or in the TSN All-Star list.

[Table 6 about here.]

I obtain similar results if I measure success with the WAR statistic constructed to gauge performance. Column 3 presents the results of the interaction between radio network variation and the index measuring whether a player's name was associated to a WAR value at least one standard deviation above the mean in a given season. The point estimate shows that successful baseball names are 1.1 p.p. more likely to be used by immigrants.

A limitation to the interpretation that good performance spurred assimilation is that successful players were also more likely to appear more often on the networks, possibly leading to a repetition effect. Although controlling for the share of games played partially accounts for this, it is only a proxy for their presence on the media. To rule out the possibility that the sheer repetition of names on the networks is a confounder of performance, I include in the analysis the impact of negative performance in column 4. Specifically, I interact radio network variation with the index measuring names associated with performance one standard deviation below the season average. According to the point estimate, immigrant households were 1.3 percentage points less likely to name their children after players who underperformed in a season. Column 5 includes both positive and negative performance together, leaving the point estimates unchanged. These results support the interpretation that immigrants did not simply select names they heard repeatedly on the media but paid close attention to how players' names were depicted, preferring successful names and avoiding those described negatively.

Impact of Performance on Natives. Finally, in Appendix Tables A.18 and A.20 I replicate the performance analysis for native children. I find results very similar to the ones I document for the immigrant population. Sample size increases by more than 15 times making estimates generally more significant than in the immigrant sample. The rest of the results tells a similar story. Radio networks increased the likelihood that natives picked names from successful baseball players, particularly those players who received an award during the season. I do not find that positive WAR deviations increased name adoption of players, but also natives, although to a lesser extent, were less likely to name their children after underperforming baseball names. In particular, I show that a one standard deviation boost in radio networks coverage led to a 3 percentage point increase in the use of successful players' names. The same variation reduced the likelihood of adopting underperforming names by 0.5 percentage points.

7 Conclusion

The anti-immigration rhetoric often asserts that immigrant assimilation is inherently challenging and that societal efforts to promote assimilation are pointless, as it relies solely on immigrants' decisions to integrate. In this paper, I show that the projection of positive role models on media is a powerful tool to foster the assimilation of immigrants. I provide evidence on this topic by exploiting the rise of radio networks, the first national media platform of US history. The networks expanded during one of the most diverse periods in US history when, at its peak, 15% of the US population was foreign-born and anti-immigrant rhetoric were rampant in society.

Leveraging quasi exogenous variation in the initial expansion of radio networks, I document that the networks were responsible for a faster integration of immigrants as measured in four key dimensions: intermarriage with natives, naturalization, naming patterns and English proficiency. I then focus on a salient segment of radio networks during the interwar period, baseball games, to argue that media role models were a key channel of assimilation. I document that baseball players acted as crucial role models through the networks, shaping the sense of belongings and the aspirations of immigrants through their ancestry and on-field performance. However, role models not only worked on immigrants but also helped the reduction of negative stereotypes from natives, generating momentum towards the successful assimilation of millions of immigrants at a critical time in history.

Deriving direct policy prescriptions from these results is challenging given the differences between the media landscape during the Golden Age of Radio and today. Nevertheless, the conclusions of the paper carry important implications for the ongoing discourse on inclusivity in the media and its impact on minority groups like immigrants. This paper furnishes systematic evidence that can be leveraged by media regulators to advocate for a fair and enhanced representation of minorities in the media, employing personalities that embody the key aspects of role models as proposed by the social psychology literature. In particular, the findings of the paper suggest that a crucial factor for successful assimilation is the representation of immigrants on the media in settings where integration with natives is conducive to success, mirroring the prominence of baseball players of immigrant origins who excelled as top athletes in the *national pastime.*

References

Abramitzky, Ran and Leah Boustan. 2017. "Immigration in American Economic History." *Journal of Economic Literature* 55:1311–45.

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson. 2020. "Do Immigrants Assimilate More Slowly Today Than in the Past?" *American Economic Review: Insights* 2:125–41.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122:467–506.

Adda, Jerome, Paolo Pinotti, and Giulia Tura. 2020. "There's More to Marriage than Love: The Effect of Legal Status and Cultural Distance on Intermarriages and Separations." CEPR Discussion Papers 14432, C.E.P.R. Discussion Papers.

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya. 2015. "Radio and the Rise of the Nazis in Prewar Germany." *The Quarterly Journal of Economics* 130:1885–1939.

Ager, Philipp, James J Feigenbaum, Casper Worm Hansen, and Hui Ren Tan. 2023. "How the Other Half Died: Immigration and Mortality in US Cities." *The Review of Economic Studies* 91:1–44.

Algan, Yann, Clément Malgouyres, Thierry Mayer, and Mathias Thoenig. 2022. "The Economic Incentives of Cultural Transmission: Spatial Evidence from Naming Patterns Across France." *The Economic Journal* 132:437–470.

Ang, Desmond. 2023. "*The Birth of a Nation*: Media and Racial Hate." *American Economic Review* 113:1424–60.

Armand, Alex, Paul Atwell, Joseph Gomes, and Yannik Schenk. 2023. "It's a Bird, It's a Plane, It's Superman!: Using Mass Media to Fight Intolerance." Cepr discussion paper no. 18207, CEPR Press, Paris & London.

Armand, Alex, Paul Atwell, and Joseph F. Gomes. 2020. "The Reach of Radio: Ending Civil Conflict through Rebel Demobilization." *American Economic Review* 110:1395–1429.

Baldassaro, Lawrence. 2011. *Beyond DiMaggio: Italian Americans in Baseball*. University of Nebraska Press.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. 2019. "Nation-building Through Compulsory Schooling during the Age of Mass Migration." *The Economic Journal* 129:62–109.

Bandura, Albert, William H Freeman, and Richard Lightsey. 1999. *Self-efficacy: The Exercise of Control*. Springer.

Bandura, Albert and Richard H Walters. 1977. *Social Learning Theory*, volume 1. Englewood cliffs Prentice Hall.

Baseball Reference. 2024. *WAR Explained*. *https://www.baseball-reference.com*.

Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilasse. 2020. "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States." *Econometrica* 88:2329–2368.

Bellemare, Marc F. and Casey J. Wichman. 2020. "Elasticities and the Inverse Hyperbolic Sine Transformation." *Oxford Bulletin of Economics and Statistics* 82:50–61.

Berkes, Enrico, Ezra Karger, and Peter Nencka. 2023. "The census place project: A method for geolocating unstructured place names." *Explorations in Economic History* 87:101477. Methodological Advances in the Extraction and Analysis of Historical Data.

Bisin, Alberto, Giorgio Topa, and Thierry Verdier. 2004. "Religious Intermarriage and Socialization in the United States." *Journal of Political Economy* 112:615–664.

Bisin, Alberto and Giulia Tura. 2019. "Marriage, Fertility, and Cultural Integration in Italy." Working Paper 26303, National Bureau of Economic Research.

Bisin, Alberto and Thierry Verdier. 2000. ""Beyond the Melting Pot": Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits." *The Quarterly Journal of Economics* 115:955–988.

Bleakley, Hoyt and Aimee Chin. 2004. "Language Skills and Earnings: Evidence from Childhood Immigrants." *The Review of Economics and Statistics* 86:481–496.

Blouin, Arthur and Sharun W. Mukand. 2019. "Erasing Ethnicity? Propaganda, Nation Building, and Identity in Rwanda." *Journal of Political Economy* 127:1008–1062.

Bolger, Eilleen. 2013. *Background History of the United States Naturalization Process*. https://socialwelfare.library.vcu.edu/federal/naturalization-process-in-u-s-early-history.

Citrin, Jack, Donald P. Green, Christopher Muste, and Cara Wong. 1997. "Public Opinion Toward Immigration Reform: The Role of Economic Motivations." *The Journal of Politics* 59:858–881.

Clubb, Jerome M., William H. Flanigan, and Nancy H. Zingale. 2006. *Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972 [Database]*. Inter-university Consortium for Political and Social Research.

Cohen, Lizabeth. 1990. *Making a new deal: Industrial workers in Chicago, 1919-1939*. Cambridge University Press.

Conley, T.G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92:1–45.

Consultative Committee on International Radio. 1992. "World Atlas of Ground Conductivities." Recommendation 832, Consultative Committee on International Radio.

DellaVigna, Stefano, Ruben Enikolopov, Vera Mironova, Maria Petrova, and Ekaterina Zhuravskaya. 2014. “Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia.” *American Economic Journal: Applied Economics* 6:103–32.

DellaVigna, Stefano and Eliana La Ferrara. 2015. “Chapter 19 - Economic and Social Impacts of the Media.” In *Handbook of Media Economics*, edited by Simon P. Anderson, Joel Waldfogel, and David Strömberg, volume 1 of *Handbook of Media Economics*, pp. 723–768. North-Holland.

Doerry, Armin W. 2013. “Radar Range Measurements in the Atmosphere.” Technical Report for the US Department of Energy, Sandia National Laboratories.

Durante, Ruben, Paolo Pinotti, and Andrea Tesei. 2019. “The Political Legacy of Entertainment TV.” *American Economic Review* 109:2497–2530.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. “Media and Political Persuasion: Evidence from Russia.” *American Economic Review* 101:3253–85.

Esposito, Elena, Tiziano Rotesi, Alessandro Saia, and Mathias Thoenig. 2023. “Reconciliation Narratives: *The Birth of a Nation* after the US Civil War.” *American Economic Review* 113:1461–1504.

Farré, Lídia and Francesco Fasani. 2013. “Media exposure and internal migration — Evidence from Indonesia.” *Journal of Development Economics* 102:48–61. Migration and Development.

Federal Communication Commission. 1947. *Standards of Good Engineering Practice Concerning Standard Broadcast Stations (550-1600 kc.)*. United States Government Printing Office.

Ferrara, Andreas, Patrick A. Testa, and Liyang Zhou. 2021. “New Area- and Population-based Geographic Crosswalks for U.S. Counties and Congressional Districts, 1790-2020.” CAGE Online Working Paper Series 588, Competitive Advantage in the Global Economy (CAGE).

Fishback, Price and Kantor Shawn. 2018. *New Deal Studies [Database]*. Inter-university Consortium for Political and Social Research.

Fishback, Price V., William C. Horrace, and Shawn Kantor. 2006. “The impact of New Deal expenditures on mobility during the Great Depression.” *Explorations in Economic History* 43:179 – 222.

Fouka, Vasiliki. 2019a. “Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I.” *The Review of Economic Studies* 87:204–239.

Fouka, Vasiliki. 2019b. “How Do Immigrants Respond to Discrimination? The Case of Germans in the US During World War I.” *American Political Science Review* 113:405–422.

Fouka, Vasiliki, Soumyajit Mazumder, and Marco Tabellini. 2021. "From Immigrants to Americans: Race and Assimilation during the Great Migration." *The Review of Economic Studies* 89:811–842.

Friendly, Michael, Chris Dalzell, Martin Monkman, Dennis Murphy, Vanessa Foot, and Justeena Zaki-Azat. 2019. *Sean 'Lahman' Baseball Database*. *http://lahman.r-forge.r-project.org/*.

Fryer, Roland G., Jr. and Steven D. Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *The Quarterly Journal of Economics* 119:767–805.

Gagliarducci, Stefano, Massimiliano Gaetano Onorato, Francesco Sobbrio, and Guido Tabellini. 2020. "War of the Waves: Radio and Resistance during World War II." *American Economic Journal: Applied Economics* 12:1–38.

Gagliarducci, Stefano and Marco Tabellini. 2022. "Faith and Assimilation: Italian Immigrants in the US." Working Paper 30003, National Bureau of Economic Research.

Gentzkow, Matthew. 2006. "Television and Voter Turnout." *The Quarterly Journal of Economics* 121:931–972.

Gentzkow, Matthew, Jesse M. Shapiro, and Michael Sinkinson. 2011. "The Effect of Newspaper Entry and Exit on Electoral Politics." *American Economic Review* 101:2980–3018.

Gordon, Milton Myron. 1964. *Assimilation in American Life: The Role of Race, Religion, and National Origins*. OUP E-Books. Oxford University Press.

Gordon, Robert J. 2017. *The Rise and Fall of American Growth: The US Standard of Living Since the Civil War*, volume 70. Princeton University Press.

Gureckis, Todd M. and Robert L. Goldstone. 2011. "How You Named Your Child: Understanding the Relationship Between Individual Decision Making and Collective Outcomes." *Topics in Cognitive Science* 1:651–674.

Haines, M.R., Inter university Consortium for Political, and Social Research. 2010. *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002*. ICPSR (Series). Inter-university Consortium for Political and Social Research.

Hainmueller, Jens and Michael J. Hiscox. 2010. "Attitudes toward Highly Skilled and Low-skilled Immigration: Evidence from a Survey Experiment." *American Political Science Review* 104:61–84.

Halper, Donna L. 2017. *Broadcasting Red Sox Baseball: How the Arrival of Radio Impacted the Team and the Fans*. Society for American Baseball Research.

Higham, John. 2002. *Strangers in the land: Patterns of American nativism, 1860-1925*. Rutgers University Press.

Hilmes, Michele. 1997. *Radio voices: American broadcasting, 1922-1952*. U of Minnesota Press.

International Telecommunication Union. 2014. *Handbook On Ground Wave Propagation.* Radiocommunication Bureau.

Jensen, Robert and Emily Oster. 2009. "The Power of TV: Cable Television and Women's Status in India." *The Quarterly Journal of Economics* 124:1057–1094.

Kaya, Ayhan. 2023. *Influence of Populism on the European Migration Agenda.* IEMed Mediterranean Yearbook. *https://www.iemed.org/publication/influence-of-populism-on-the-european-migration-agenda/.*

Kirby, Robert S, J. C. Harman, F. M. Capps, and R. N. Jones. 1954. *Effective Radio Ground-Conductivity Measurements in the United States*, volume 546. United States Department of Commerce.

Kirchick, James. 2019. "European populism and immigration." *Great Decisions* pp. 51–58.

La Ferrara, Eliana. 2016. "Mass Media and Social Change: Can We Use Television To Fight Poverty?" *Journal of the European Economic Association* 14:791–827.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics* 4:1–31.

Lenthall, B. 2008. *Radio's America: The Great Depression and the Rise of Modern Mass Culture*. University of Chicago Press.

Levine, Peter. 1993. *Ellis Island to Ebbets Field: Sport and the American Jewish Experience*. Oxford University Press.

Lieberson, Stanley and Eleanor O. Bell. 1992. "Children's First Names: An Empirical Study of Social Taste." *American Journal of Sociology* 98:511–554.

Lippmann, Stephen. 2008. "Rationalization, Standardization, or Market Diversity?: Station Networks and Market Structure in US Broadcasting, 1927–1950." *Social Science History* 32:405–436.

Lleras-Muney, Adriana and Allison Shertzer. 2015. "Did the Americanization Movement Succeed? An Evaluation of the Effect of English-Only and Compulsory Schooling Laws on Immigrants." *American Economic Journal: Economic Policy* 7:258–90.

Lockwood, Penelope and Ziva Kunda. 1997. "Superstars and me: Predicting the impact of role models on the self." *Journal of personality and social psychology* 73:91.

Major League Baseball. 2024. *Wins Above Replacement (WAR).* https://www.mlb.com/glossary/advanced-stats/wins-above-replacement.

Olken, Benjamin A. 2009. "Do Television and Radio Destroy Social Capital? Evidence from Indonesian Villages." *American Economic Journal: Applied Economics* 1:1–33.

Riley, Emma. 2022. "Role Models in Movies: The Impact of Queen of Katwe on Students' Educational Attainment." *The Review of Economics and Statistics* 106:1–48.

Ruggles, Steven, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Matt A. Nelson, Evan Roberts, Megan Schouweiler, and Matthew Sobek. 2021. *IPUMS Ancestry Full Count Data: Version 3.0 [dataset]*. Minneapolis, MN: IPUMS.

Scott, Carole. 2008. *History of the Radio Industry in the United States to 1940*. EH.Net Encyclopedia, edited by Robert Whaples. *http://eh.net/encyclopedia/the-history-of-the-radio-industry-in-the-united-states-to-1940/.*

Sterling, C.H. and J.M. Kittross. 2001. *Stay Tuned: A History of American Broadcasting*. Routledge Communication Series. Taylor & Francis.

Stromberg, David. 2004. "Radio's Impact on Public Spending." *The Quarterly Journal of Economics* 119:189–221.

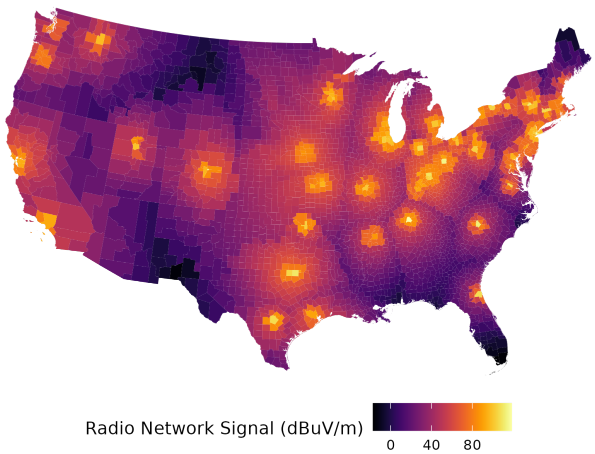
Tabellini, Marco. 2019. "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration." *The Review of Economic Studies* 87:454–486.

United Nations, Department of Economic and Population Division Social Affairs. 2019. "International Migration 2019: Highlights (ST/ESA/SER.A/439)."

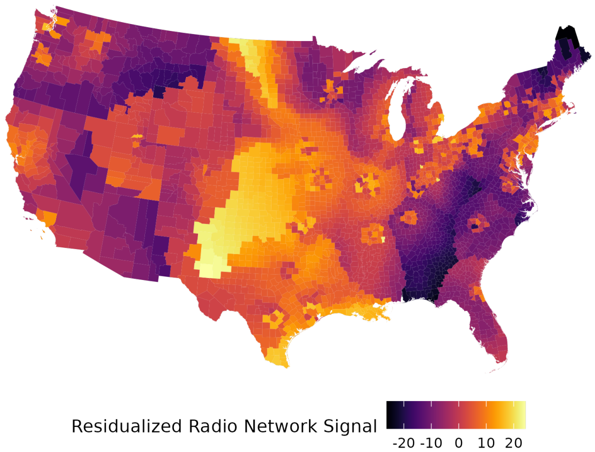
Wang, Tianyi. 2021. "Media, Pulpit, and Populist Persuasion: Evidence from Father Coughlin." *American Economic Review* 111:3064–92.

Ward, Zachary. 2019. "The Low Return to English Fluency During the Age of Mass Migration." *European Review of Economic History* 24:219–242.

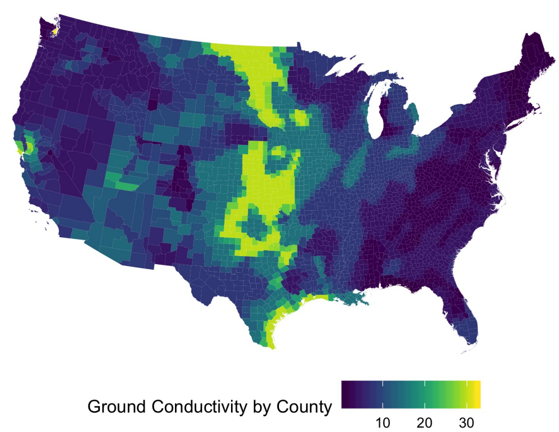
Yanagizawa-Drott, David. 2014. "Propaganda and Conflict: Evidence from the Rwandan Genocide." *The Quarterly Journal of Economics* 129:1947–1994.



(a) Actual Signal Strength Variation



(b) Residualized Signal Strength Variation



(c) Ground Conductivity Variation

Figure 1: County-Level Radio Networks Signal Strength Variation

Notes: Panel (a) shows the actual radio networks signal strength variation; Panel (b) shows the residualized variation after partialing out geographic controls and the free network signal strength; Panel (c) presents county-level ground conductivity, the main driver of AM radio signal.

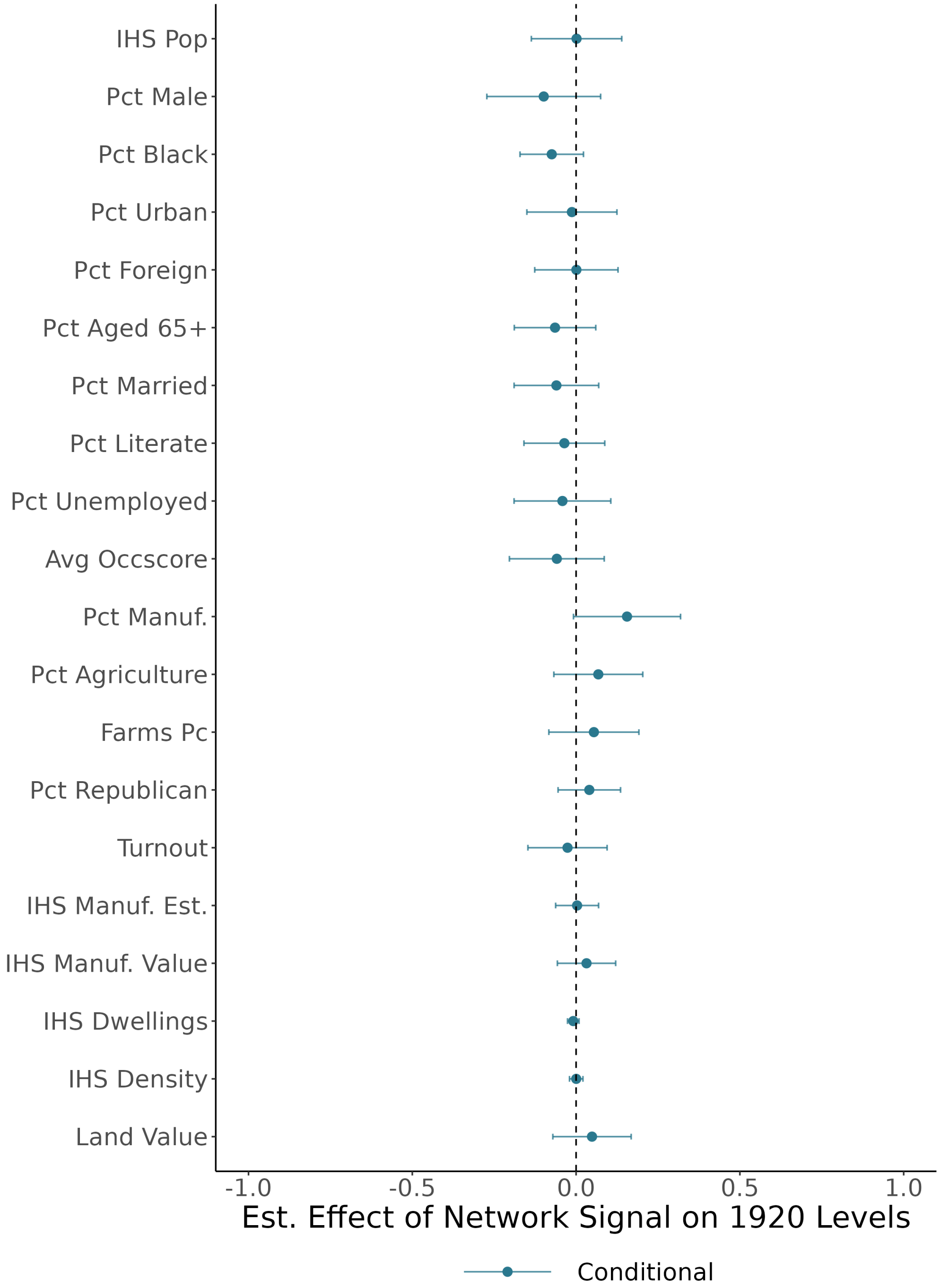


Figure 2: Balancedness Test on 1920 County-Level Observables

Notes: The figure plots estimated coefficients from the balancedness test for radio networks coverage. I sequentially regress the 1920 county-level observables indicated on the y-axis on the radio networks signal strength conditional on state fixed effects, the *free* signal strength, 1920 IHS population, geographic controls and IHS distances to the closest network antenna and non-network antenna. Both radio networks signal strength and 1920 county-level variables are expressed in standard deviations. I plot 95% confidence intervals obtained with standard errors clustered at the county-level.

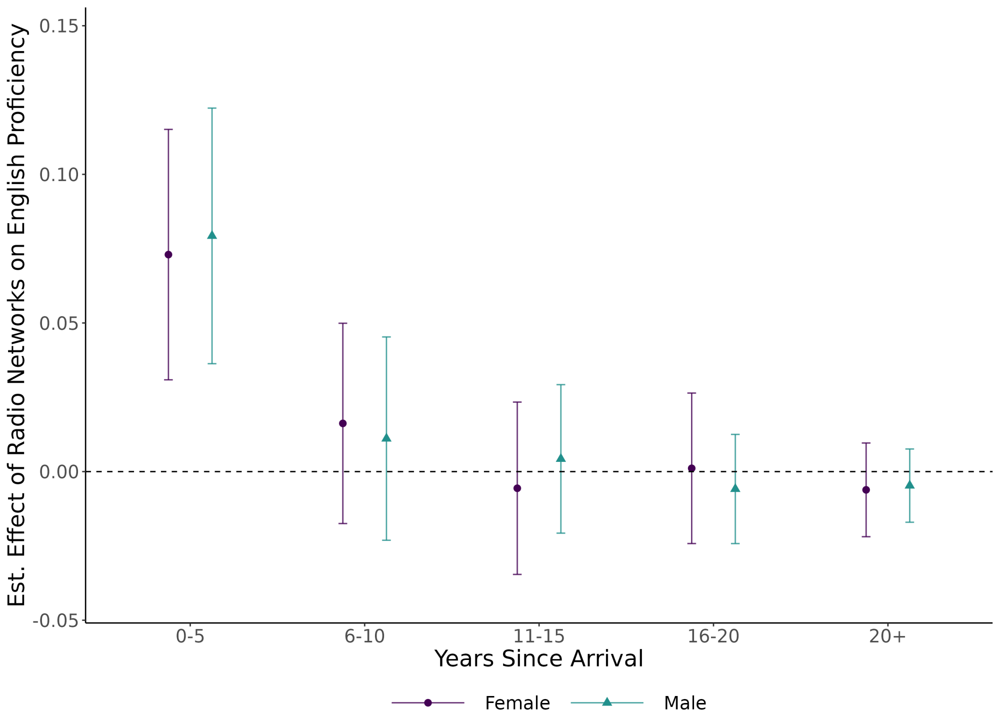
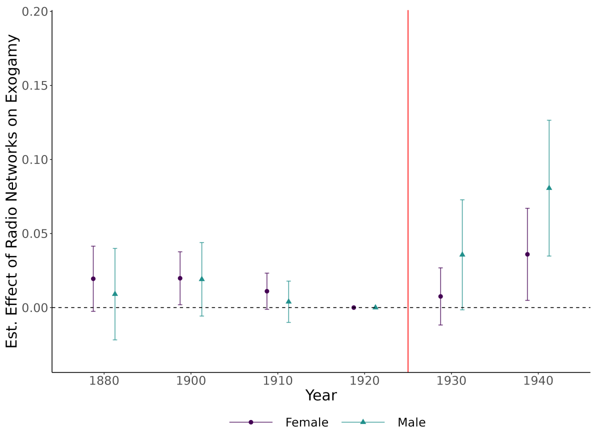


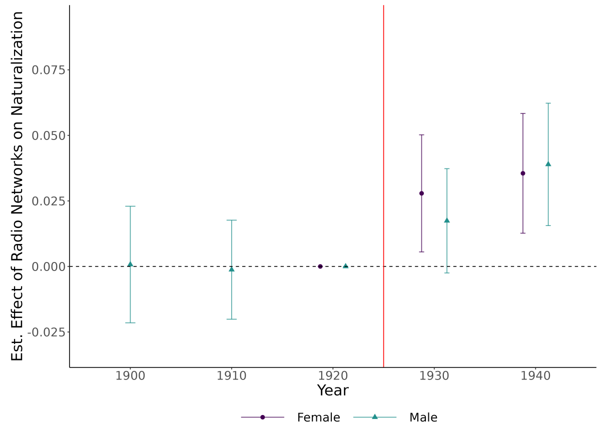
Figure 3: Estimated Effect of Radio Networks on Self-Reported English Proficiency Interacted with Year of Arrival

Notes: The figure plots estimated coefficients of the impact of radio networks on English proficiency by year of arrival. Sample includes repeated cross-sections of all immigrants above the age of 10 between 1900 and 1940 (1920 and 1940 for female). I exclude from the analysis immigrants from countries where English was the first language. The ranges for year of arrival are determined by the US census. Radio networks signal strength is expressed in standard deviations. The specification used to obtain the estimates is the most saturated one from column 6 of Table 4. I plot 95% confidence intervals obtained with standard errors clustered at the county-level.

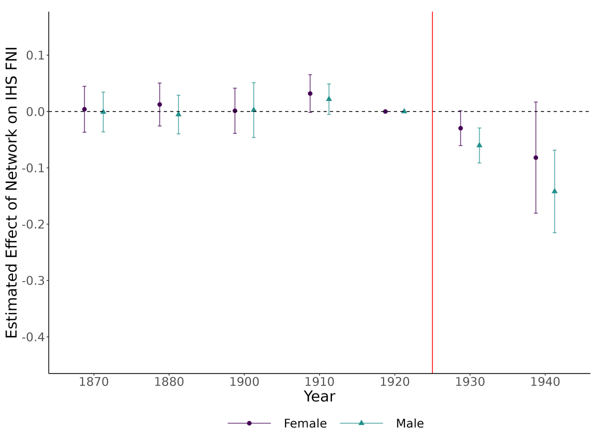
Figure 4: Event Study on the Effects of Radio Networks on Immigrant Assimilation



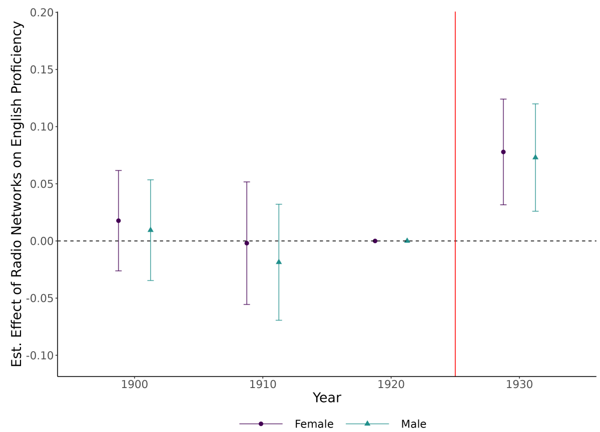
(a) Intermarriage



(b) Naturalization



(c) IHS FNI



(d) English Proficiency

Notes: The figure plots estimated coefficients from the event study analysis of the impact of radio networks on four outcome variables. The list of controls in the specifications used to obtain the estimates is the same as in the most saturated specification in column 6 of Tables 2, 3 and 4. To further assess robustness, the exercise also includes decade by region fixed effects. Radio network signal strength is expressed in standard deviations. I plot 95% confidence intervals obtained with standard errors clustered at the county-level.

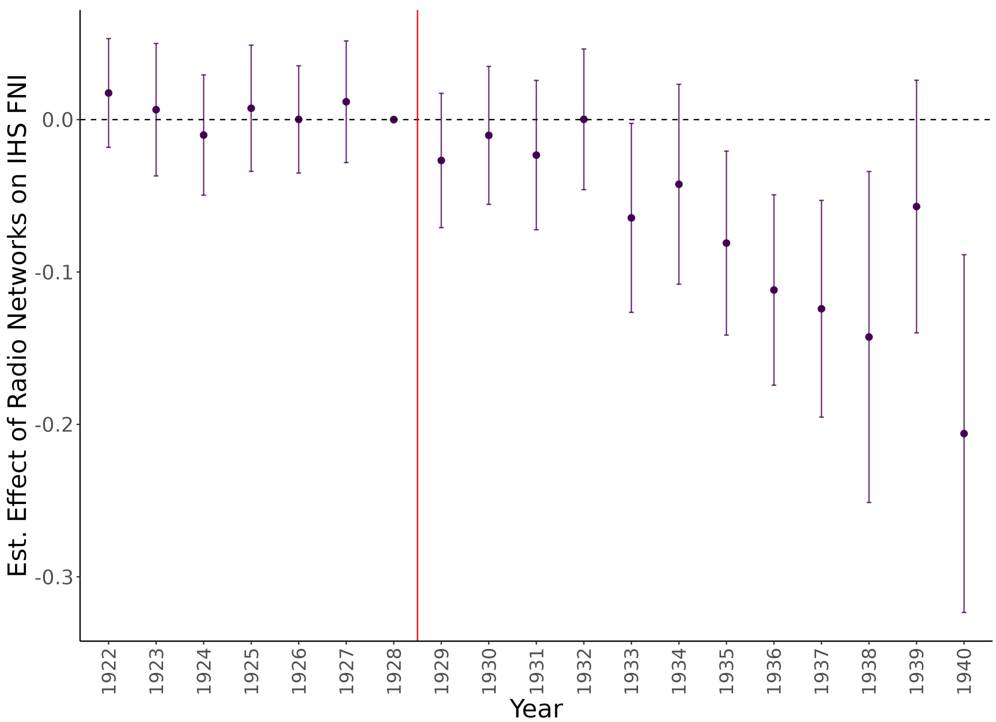
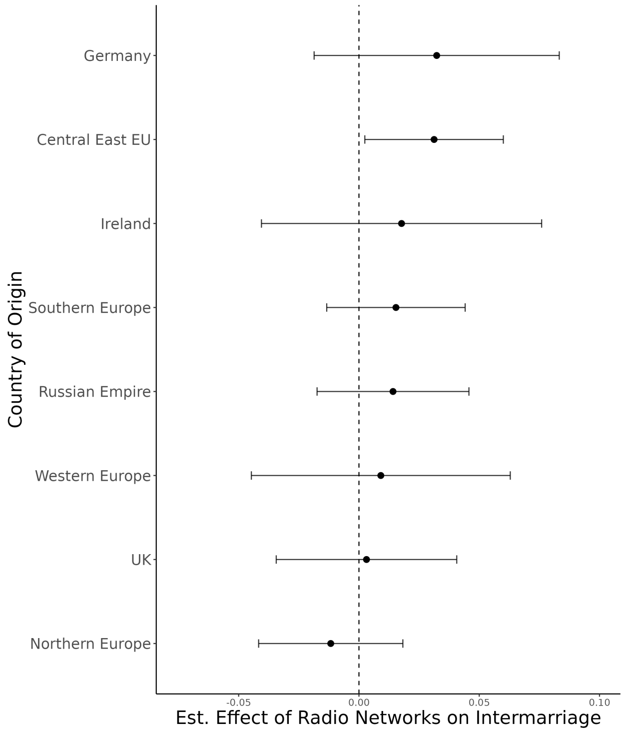


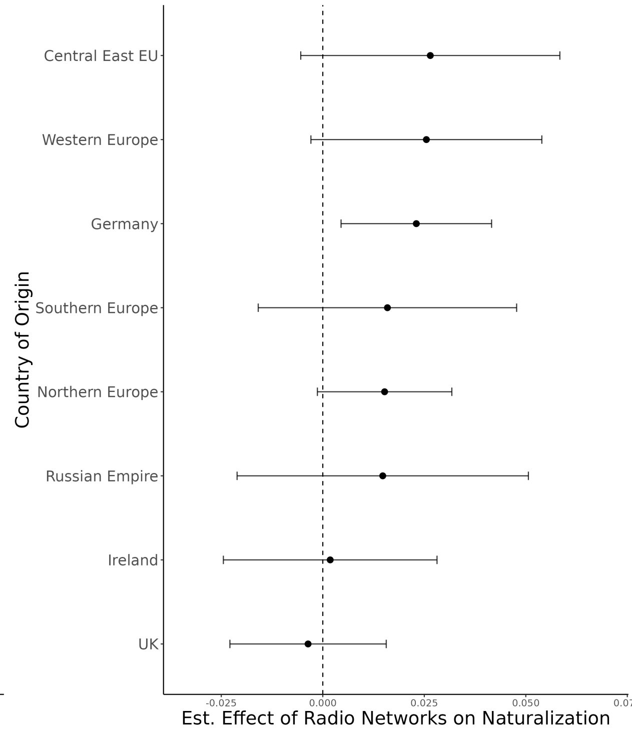
Figure 5: Radio Networks Impact on FNI, Event Study Analysis Using Within Household Variation

Notes: The figure plots estimated coefficients from the event study analysis of the impact of radio networks on IHS FNI. The sample is composed by birth cohorts of male immigrants born between 1922 and 1940 alive in 1940 census year. The empirical model estimated compares naming patterns of male immigrants within the same households, before and after the rise of radio networks controlling for IHS population in 1920 and *free* signal strength interacted with birth cohort fixed effects. The model accounts for household and birth cohort fixed effects: $Y_{ict} = \sum_{t=1922}^{1927} \beta_t^{pre} Network_c D_t + \sum_{t=1929}^{1940} \beta_t^{post} Network_c D_t + \beta_f Free_c \alpha_t + \gamma Pop \alpha_t + \alpha_{hh} + \alpha_t + \varepsilon_{iht}$. The diff-in-diff counterpart of the event study estimate is equal to -0.051 (s.e. 0.015), significant at the 1% level. Radio network signal strength is expressed in standard deviations. I plot 95% confidence intervals obtained with standard errors clustered at the county-level.

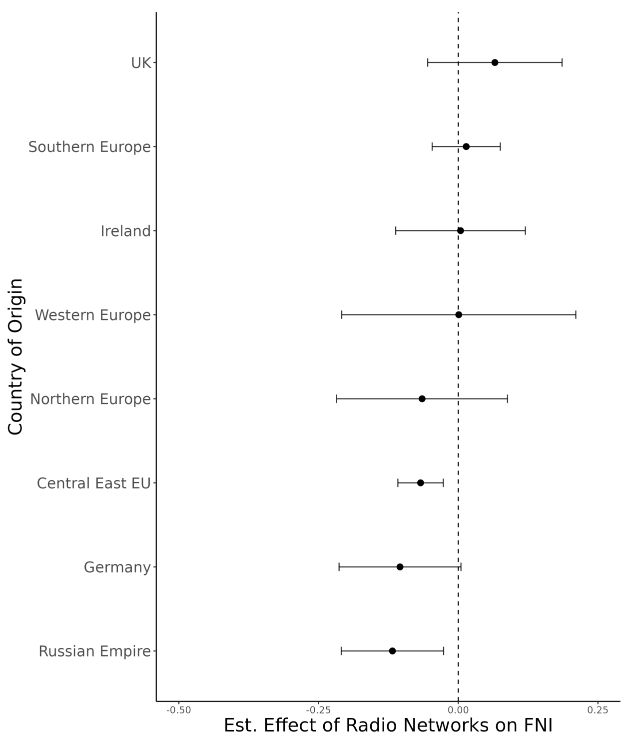
Figure 6: Heterogeneity Across Immigrant Country of Origin



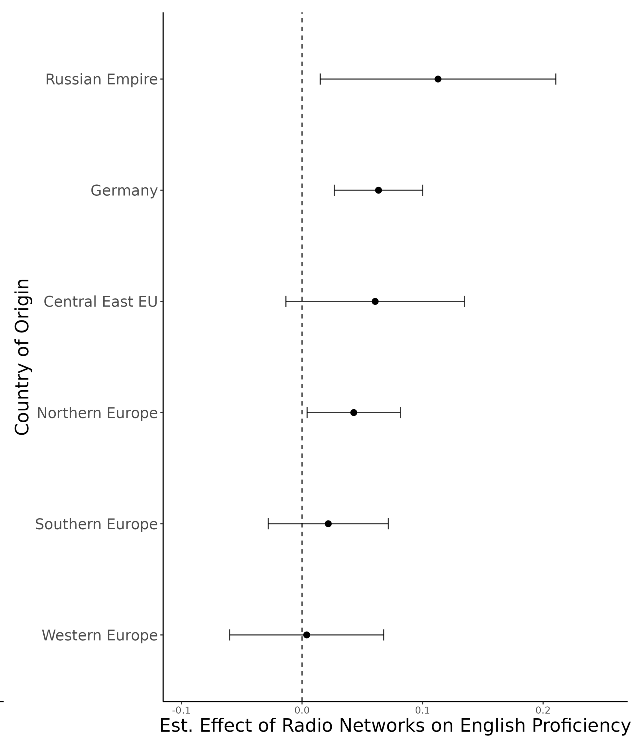
(a) Intermarriage



(b) Naturalization



(c) IHS FNI



(d) English Proficiency

Notes: The figure plots estimated coefficients of the impact of radio networks on four outcome variables by immigrant country of origin. The specifications used to obtain the estimates are the most saturated ones from column 6 of Tables 2, 3 and 4. Sample is restricted to male immigrants from European countries. Radio network signal strength is expressed in standard deviations. I plot 95% confidence intervals obtained with standard errors clustered at the county-level.

Table 1: Network Radio Orthogonal to Outcome in Levels in the Pre-Period, 1920

	IHS FNI (1)	Intermarriage (2)	Naturalization (3)	English Proficiency (4)
Network	−0.014	−0.003	0.012	0.002
	(0.016)	(0.038)	(0.012)	(0.009)
Free Signal	✓	✓	✓	✓
Geographic Controls	✓	✓	✓	✓
IHS Pop	✓	✓	✓	✓
State FE	✓	✓	✓	✓
Observations	2,137	2,137	2,137	2,137
Adjusted R^2	0.307	0.131	0.577	0.602
Mean Outcome	0.404	4.528	0.586	0.874
Sd Outcome	0.190	0.496	0.181	0.138

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Estimated effect of radio networks on 1920 immigrants' assimilation outcomes. Sample includes the set of counties with no missing on all outcome variables. Dependent variables are measured at the 1920 county-level and include: the average IHS FNI, intermarriage rate, naturalization rate and English proficiency rate. All columns control for the *free* signal strength, the IHS 1920 population, the geographic controls and their squared term, as well as state fixed effects. Geographic controls include average temperature, average rainfall, average elevation, average ruggedness, distance to closest coast, distance to closest river, distance to closest lake, distance to nearest historical portage site, distance to nearest historical Indian battle, distance to nearest mine site in 1890, average rainfall risk and average potential agricultural productivity. I also control for the distance to the closest network antenna, the distance to the closest non-affiliated antenna and the area of the county.

Table 2: Network Radio Increased Intermarriage with Natives for Male Immigrants

	Likelihood Intermarriage					
	(1)	(2)	(3)	(4)	(5)	(6)
Network × Post	−0.015	0.063***	0.069***	0.071***	0.059***	0.060***
	(0.009)	(0.022)	(0.023)	(0.023)	(0.021)	(0.020)
Baseline	✓	✓	✓	✓	✓	✓
Free Signal		✓	✓	✓	✓	✓
Geographic			✓	✓	✓	✓
Age				✓	✓	✓
County					✓	✓
Political						✓
County FE	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓
Observations	4,575,324	4,575,324	4,575,324	4,575,324	4,575,324	4,575,324
Adjusted R^2	0.16	0.16	0.16	0.17	0.17	0.17
Mean Outcome	0.285	0.285	0.285	0.285	0.285	0.285
Sd Outcome	0.452	0.452	0.452	0.452	0.452	0.452

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the estimated effect of radio networks signal strength on the likelihood the immigrant's spouse is US born. The sample includes repeated cross-sections of male immigrants below the age of 35 between 1880 and 1940. Radio network signal strength is expressed in standard deviations. All regressions include county, decade and country of origin fixed effects and a set of baseline controls all measured in 1920 and interacted with decade fixed effects: log population, share Black, share urban, share foreign born, share literate and share in agriculture. Free signal refers to the hypothetical radio network signal strength in the absence of ground conductivity impediment interacted with decade fixed effects. Geographic controls measure average temperature, average rainfall, average elevation, average ruggedness, distance to closest coast, distance to closest river, distance to closest lake, distance to nearest historical portage site, distance to nearest historical Indian battle, distance to nearest mine site in 1890, average rainfall risk and average potential agricultural productivity, all interacted with decade fixed effects. County-level controls adds a larger set of county characteristics measured in 1920 interacted with decade fixed effects: share male, share older than 65, share married, share unemployed, average occupational score, share in manufacturing and number of farms per capita. Mean and standard deviations of the outcome variables refer to the pre-treatment period. Standard errors in parentheses are clustered at the county-level.

Table 3: Network Radio Increased Assimilation Effort by Immigrants

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Naturalization					
Network × Post	0.008	0.019	0.035***	0.026***	0.025***	0.025***
	(0.005)	(0.012)	(0.011)	(0.010)	(0.009)	(0.009)
Observations	27,956,713	27,956,713	27,956,713	27,956,713	27,956,713	27,956,713
Adjusted R^2	0.15	0.15	0.15	0.18	0.19	0.19
Mean Outcome	0.588	0.588	0.588	0.588	0.588	0.588
Sd Outcome	0.492	0.492	0.492	0.492	0.492	0.492
	Panel B: IHS FNI					
Network × Post	−0.009	−0.077***	−0.094***	−0.092***	−0.089***	−0.091***
	(0.007)	(0.018)	(0.021)	(0.020)	(0.020)	(0.018)
Observations	754,491	754,491	754,491	754,491	754,491	754,491
Adjusted R^2	0.07	0.07	0.07	0.07	0.07	0.07
Mean Outcome	4.582	4.582	4.582	4.582	4.582	4.582
Sd Outcome	0.701	0.701	0.701	0.701	0.701	0.701
Baseline	✓	✓	✓	✓	✓	✓
Free Signal		✓	✓	✓	✓	✓
Geographic			✓	✓	✓	✓
Age				✓	✓	✓
County					✓	✓
Political						✓
County FE	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the estimated impact of radio networks on naturalization and naming decisions. In panel A the sample includes all male children born at census year between 1870 to 1940 from foreign born parents. The dependent variable is the inverse hyperbolic sine of the name foreignness index assigned to the newborns. In panel B the sample is composed by all male immigrants older than twenty years old present in the US at census year from 1900 to 1940. The dependent variable is an indicator equal to one if the immigrant is naturalized US citizen. Radio network signal strength is expressed in standard deviations. All regressions include county, decade and country of origin fixed effects and a set of baseline controls all measured in 1920 and interacted with decade fixed effects: log population, share Black, share urban, share foreign born, share literate and share in agriculture. Free signal refers to the hypothetical radio network signal strength in the absence of ground conductivity impediment interacted with decade fixed effects. Geographic controls measure average temperature, average rainfall, average elevation, average ruggedness, distance to closest coast, distance to closest river, distance to closest lake, distance to nearest historical portage site, distance to nearest historical Indian battle, distance to nearest mine site in 1890, average rainfall risk and average potential agricultural productivity, all interacted with decade fixed effects. County-level controls adds a larger set of county characteristics measured in 1920 interacted with decade fixed effects: share male, share older than 65, share married, share unemployed, average occupational score, share in manufacturing and number of farms per capita. Mean and standard deviations of the outcome variables refer to the pre-treatment period. Standard errors in parentheses are clustered at the county-level.

Table 4: Network Radio Boosted English Proficiency for Recent Immigrants

	(1)	(2)	(3)	(4)	(5)	(6)
Network × Post	−0.017	0.069***	0.093***	0.093***	0.084***	0.079***
	(0.012)	(0.025)	(0.025)	(0.025)	(0.022)	(0.022)
Baseline	✓	✓	✓	✓	✓	✓
Free Signal		✓	✓	✓	✓	✓
Geographic			✓	✓	✓	✓
Age				✓	✓	✓
County					✓	✓
Political						✓
County FE	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓
Observations	1,985,965	1,985,965	1,985,965	1,985,965	1,985,965	1,985,965
Adjusted R^2	0.19	0.19	0.19	0.20	0.20	0.20
Mean Outcome	0.535	0.535	0.535	0.535	0.535	0.535
Sd Outcome	0.499	0.499	0.499	0.499	0.499	0.499

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the estimated effect of radio networks signal strength on the likelihood the immigrant's declare to speak proper English. The sample includes repeated cross-sections of male immigrants above the age of 10 between 1900 and 1930. I restrict the sample to recent immigrants who arrived in the US within five years from census year and exclude from the analysis immigrants from countries where English was the first language. Radio network signal strength is expressed in standard deviations. All regressions include county, decade and country of origin fixed effects and a set of baseline controls all measured in 1920 and interacted with decade fixed effects: log population, share Black, share urban, share foreign born, share literate and share in agriculture. Free signal refers to the hypothetical radio network signal strength in the absence of ground conductivity impediment interacted with decade fixed effects. Geographic controls measure average temperature, average rainfall, average elevation, average ruggedness, distance to closest coast, distance to closest river, distance to closest lake, distance to nearest historical portage site, distance to nearest historical Indian battle, distance to nearest mine site in 1890, average rainfall risk and average potential agricultural productivity, all interacted with decade fixed effects. County-level controls adds a larger set of county characteristics measured in 1920 interacted with decade fixed effects: share male, share older than 65, share married, share unemployed, average occupational score, share in manufacturing and number of farms per capita. Mean and standard deviations of the outcome variables refer to the pre-treatment period. Standard errors in parentheses are clustered at the county-level.

Table 5: Match with Player's Ancestry on the Networks Drives Immigrant Naming Choices

	Baseball Name			
	(1)	(2)	(3)	(4)
Post × Network	0.008	0.007	0.005	−0.013
	(0.007)	(0.007)	(0.005)	(0.008)
– × Same Country		0.032***		
		(0.011)		
– × Player Recent Imm Group			0.017***	
			(0.005)	
– × HH Recent Imm Group				0.030***
				(0.005)
Birth Cohort FE	✓	✓	✓	✓
Household FE	✓	✓	✓	✓
Observations	1,849,074	1,849,074	1,849,074	1,849,074
Adjusted R^2	0.71	0.71	0.73	0.71
Mean Outcome	0.281	0.281	0.281	0.281
Sd Outcome	0.450	0.450	0.450	0.450

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Estimated effect of standardized network signal strength on the likelihood of naming a child after a baseball player active during the same year the child was born interacted with the ancestry of either the player or the immigrant household. The sample is composed by white children from a household with both parents foreign. All regressions account for household and birth cohort fixed effects. Controls include 1920 IHS population, the *free* network coverage interacted with birth cohort fixed effects and the share of games played in a season by the player. The model estimates the interaction between radio network signal variation, the post-1929 dummy and a series of indicators relating to ancestry: a dummy equal to one if the player's ancestry is equal to the country of origin of the household, a dummy equal to one if the player's last name suggests that the player's ancestry is from either Eastern Europe, Southern Europe, Germany, or Russia, and a dummy equal to one if the immigrant household immigrated from either Eastern Europe, Southern Europe, Germany, or Russia. Although not shown in the table, the model estimates all lower levels interaction. I cluster standard errors at the county-level.

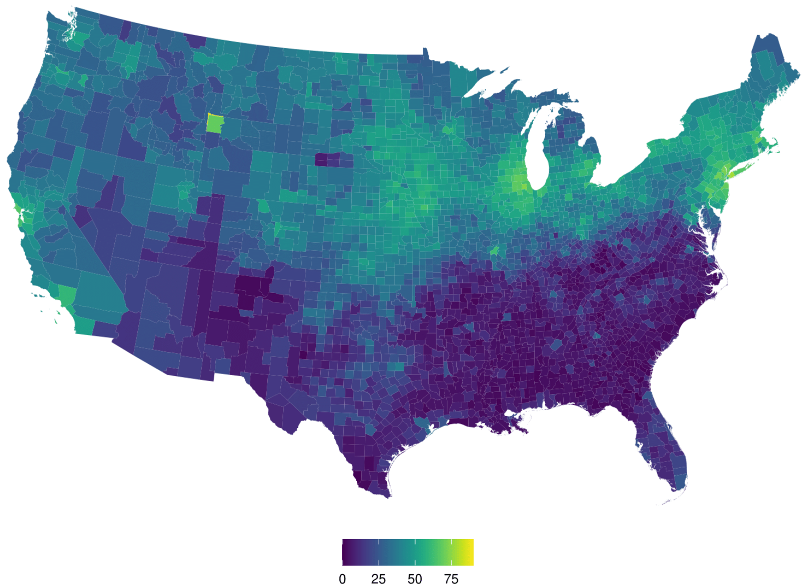
Table 6: Networks Increased Naming After Success for Immigrants

	Baseball Name				
	(1)	(2)	(3)	(4)	(5)
Network × Post	0.008	0.008	0.008	0.011	0.011
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
– × Any Award		0.029***			
		(0.011)			
– × WAR^{+}_{1sd}			0.011***		0.011***
			(0.004)		(0.004)
– × WAR^{-}_{1sd}				−0.012**	−0.013**
				(0.006)	(0.006)
Birth Cohort FE	✓	✓	✓	✓	✓
Household FE	✓	✓	✓	✓	✓
Observations	1,849,074	1,849,074	1,849,074	1,849,074	1,849,074
Adjusted R^2	0.71	0.71	0.71	0.74	0.74
Mean Outcome	0.281	0.281	0.281	0.281	0.281
Sd Outcome	0.450	0.450	0.450	0.450	0.450

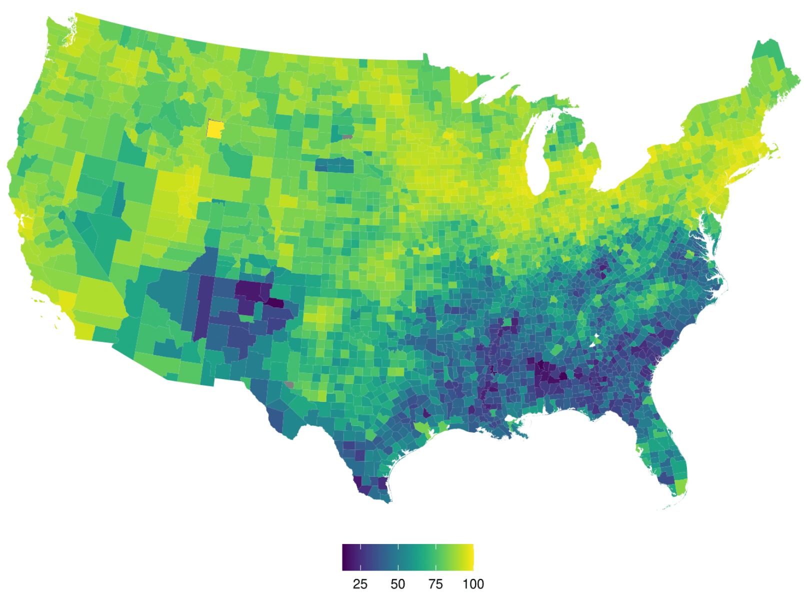
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Estimated effect of standardized network signal strength on the likelihood of naming a child after a baseball player active during the same year the child was born interacted with the performance of the player. The sample is composed by white children from a household with both parents foreign. All regressions account for household and birth cohort fixed effects. Controls include 1920 IHS population, *free* network coverage interacted with birth cohort fixed effects and the share of games played associated to the baseball name. The model estimates the interaction between radio network signal variation, the post-1929 dummy and a series of performance indicators: any award during the season, scoring a WAR value that is one standard deviation above the season mean WAR or scoring a WAR value that is one standard deviation below the season mean WAR. Although not shown in the table, the model estimates all lower levels interaction. I cluster standard errors at the county-level.

A Online Appendix (Not for publication)



(a) Share of Households With a Radio Set in 1930



(b) Share of Households With a Radio Set in 1940

Figure A.1: Share of US Households Owning a Radio Set in 1930 and 1940. Source: Haines et al. (2010)

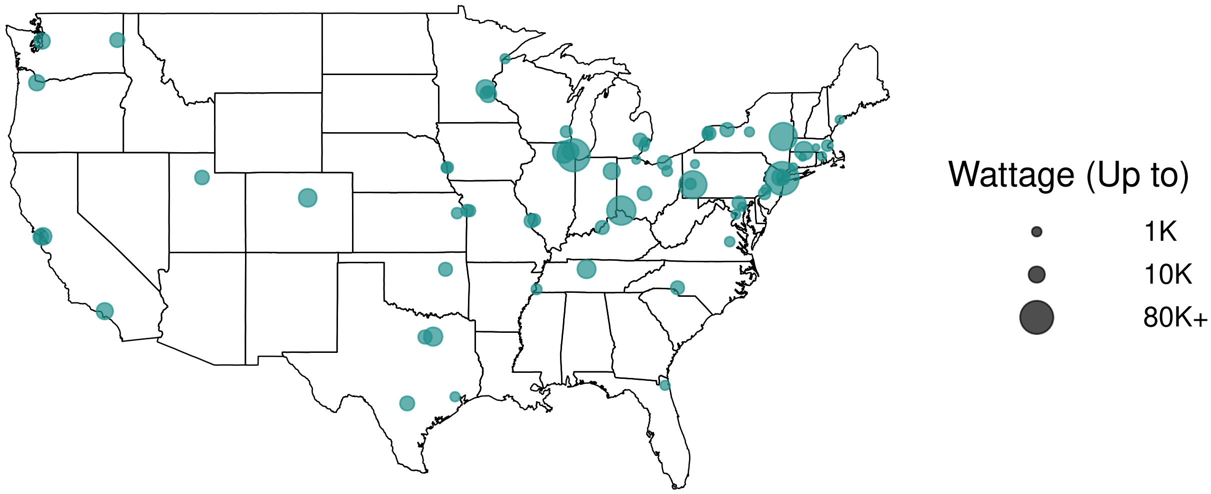
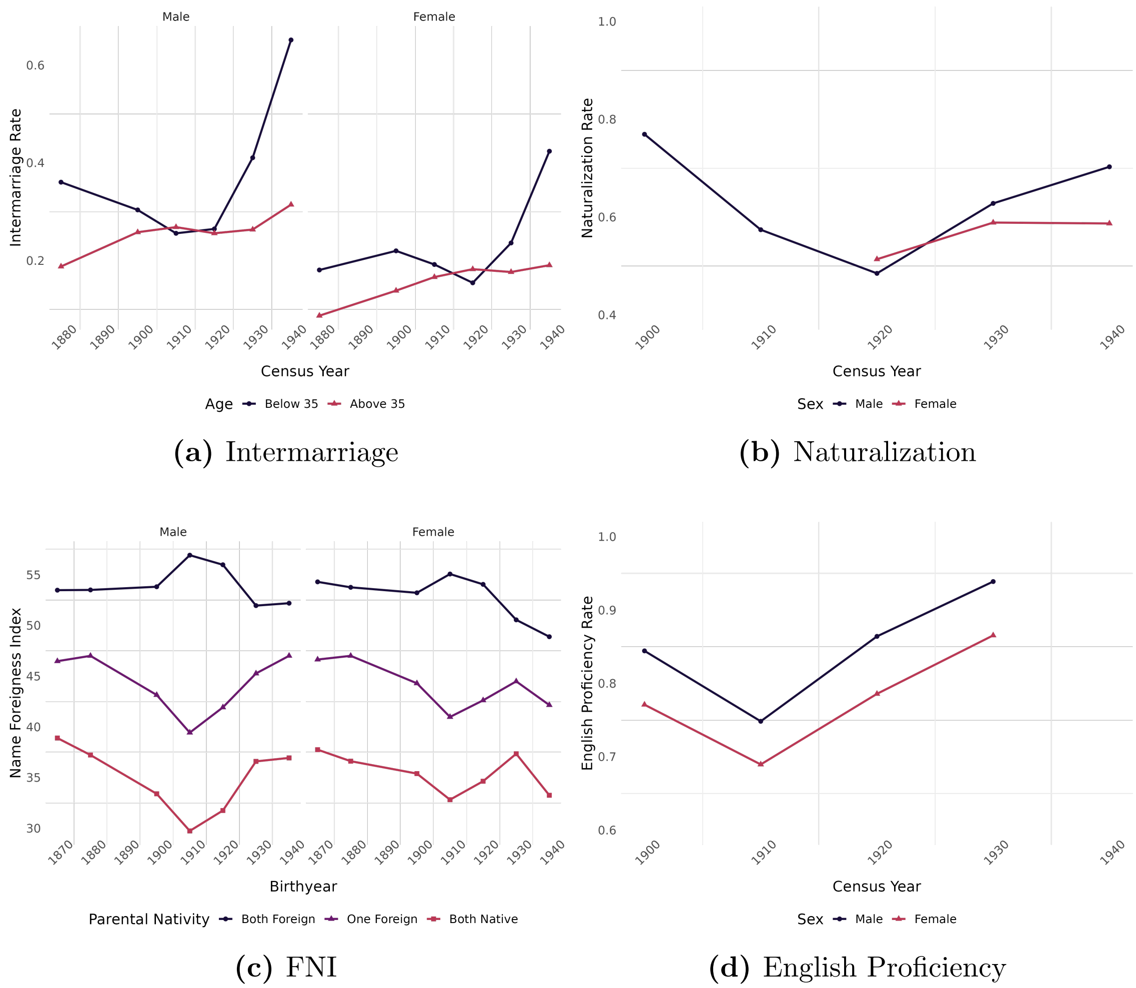


Figure A.2: Locations with at Least One Active Radio Station with a Network Affiliation in 1929. Size of the dots indicates power of the antennas. Source of the data is the 1929 January edition of the *White's Radio Log*.

Figure A.3: Evolution of Assimilation Outcomes Over Time



(a) Intermarriage

(b) Naturalization

(c) FNI

(d) English Proficiency

Notes: Descriptive statistics elaborated from the full count US census (Ruggles et al. 2021)

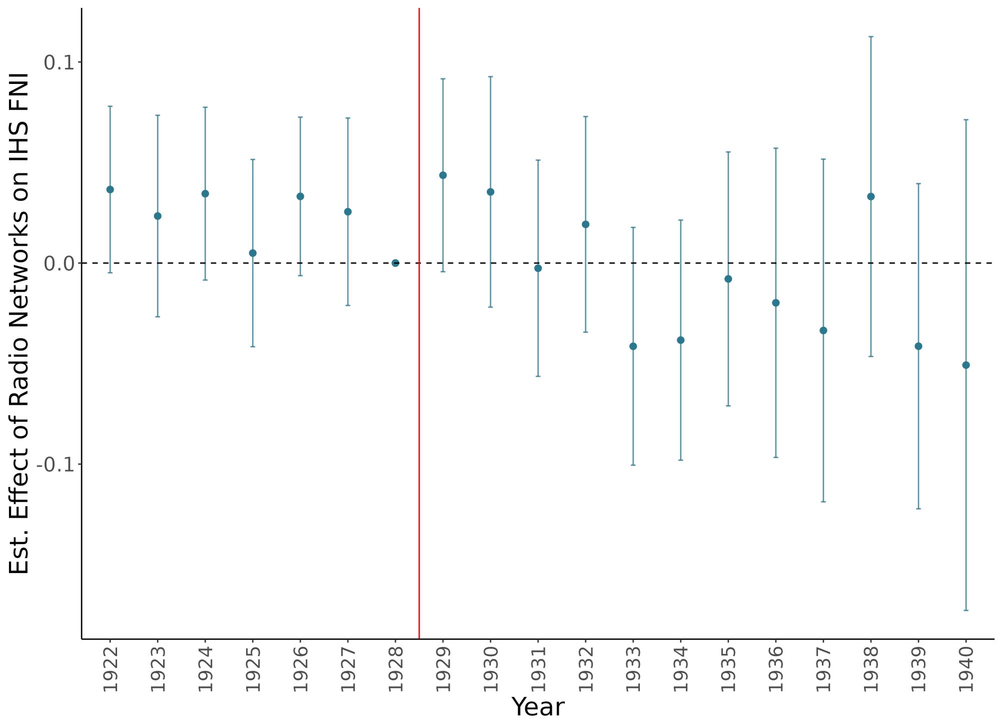


Figure A.4: Radio Networks Impact on FNI, Event Study Analysis Using Within Household Variation, Female Immigrants

Notes: The figure plots estimated coefficients from the event study analysis of the impact of radio networks on IHS FNI for female immigrants. The sample is composed by birth cohorts of immigrants born between 1922 and 1940 alive in 1940 census year. The empirical model estimated compares naming patterns of female immigrants within the same households, before and after the rise of radio networks controlling for IHS population in 1920 and *free* signal strength interacted with year fixed effects: $Y_{ict} = \sum_{t=1922}^{1927} \beta_t^{pre} Network_c D_t + \sum_{t=1929}^{1940} \beta_t^{post} Network_c D_t + \beta_f Free_c \alpha_t + \gamma ihspop \alpha_t + \alpha_{hh} + \alpha_t + \varepsilon_{iht}$. The diff-in-diff counterpart of the event study estimate is equal to -0.017 (s.e. 0.015). Radio networks signal strength is expressed in standard deviations. I plot 95% confidence intervals obtained with standard errors clustered at the county-level.

Table A.1: County-Level Descriptive Statistics, 1920

	N	Mean	Median	SD	Min	Max
Intermarriage Rate	2137	0.4	0.4	0.19	0	1
Naturalization Rate	2137	0.59	0.62	0.18	0.0047	0.97
FNI	2137	50	50	14	0	95
English Proficiency Rate	2137	0.87	0.91	0.14	0.094	1
Population	2137	41729	20499	125364	273	3022912
Percent Male	2137	52	51	2.4	46	67
Percent Black	2137	7.3	0.43	14	0	79
Percent Urban	2137	24	19	25	0	100
Percent Foreign-Born	2137	9.1	7.3	8.3	0.034	54
Percent 65+	2137	5.1	4.7	2	1.2	14
Percent Married	2137	40	40	3.4	28	50
Percent Literate	2137	72	75	7.7	32	85
Percent Unemployed	2137	29	29	4.1	13	56
Occupational Score	2137	5.7	5.5	1.2	0.56	11
Percent Manufacturing	2137	10	5.4	11	0.19	76
Percent Agriculture	2137	50	54	22	0.54	94
Farms pc	2137	0.097	0.1	0.046	0.0000022	0.28
Republican Share	2137	0.56	0.6	0.2	0	0.94
Turnout	2137	11219	5038	33445	12	893137
Manuf. Ests	2137	121	39	649	0.28	22981
Manuf. Value (000s)	2137	27047	2518	151946	7.3	3908354
Dwellings	2137	8120	4508	17646	54	399684
Pop Density	2137	168	32	2446	0.2	103823
Land Value (000s)	2137	21980	13682	24947	15	340682

Note: Descriptive statistics of the baseline county sample in 1920 census year, the last census year before treatment occurs. Statistics elaborated from the US full count census (Ruggles et al. 2021)

Table A.2: Baseball Names Dataset, Descriptive Statistics

	N	Mean	Median	SD	Min	Max
Rate Games Played	5068	0.43	0.34	0.31	0.0063	1
Recent Imm Group Ancestry	5068	0.35	0	0.48	0	1
WAR	5068	1.2	0.68	1.8	-3.2	12
WAR^{+}_{1sd}	5068	0.15	0	0.36	0	1
WAR^{+}_{1sd}	5068	0.07	0	0.26	0	1
Any Award	5068	0.053	0	0.2	0	1
BM Allstar	5068	0.05	0	0.2	0	1
TSN Allstar	5068	0.018	0	0.11	0	1
TSN MVP	5068	0.0025	0	0.042	0	1
N Awards	5068	0.076	0	0.32	0	4.8

Note: Descriptive statistics of the baseball names data. Variation is at the name by year level. Recent Imm Group Ancestry refers to player's ancestry belonging to Southern Europe, Eastern Europe, Russia or Germany. $WAR^{+/-}_{1sd}$ is an indicator equal to one if the WAR statistic in a season is above or below the mean WAR of the season. Other awards refer to being featured in the Baseball Magazine All Star team, being featured in the TSN All Star team or being featured as the TSN MVP. Award indicators are collapsed-weighted at the name-level by the share of games played in a season.Â Source of the data is the Lahman Baseball data (Friendly et al. 2019).

Table A.3: Newborn Pseudo Birth Cohorts, Descriptive Statistics 1922-1940

	N	Mean	Median	SD	Min	Max
FNI	3120985	52	54	20	0	100
Baseball Name	3120985	0.13	0	0.34	0	1
Recent Imm Group Ancestry	3120985	0.68	1	0.47	0	1
Male	3120985	0.51	1	0.5	0	1
Age Household Head	3120985	36	35	7.4	15	99
Employed Household Head	3120985	0.98	1	0.16	0	1
Occscore Household Head	3120985	25	23	9.2	3	80
Urban	3120985	0.78	1	0.42	0	1
Farm	3120985	0.098	0	0.3	0	1

Note: Descriptive statistics of the 1922-1940 birth cohort sample extracted from the 1940 sample. Statistics elaborated from the US full count census (Ruggles et al. 2021)

Table A.4: Balancedness Test on Pre-Treatments Observables, 1920

	Unconditional	Conditional	Test Difference (1-2)
IHS Pop	0.42***	0.001	34.738
	(0.024)	(0.07)	[0.000]
Pct Male	−0.175***	−0.099	0.877
	(0.022)	(0.089)	[0.349]
Pct Black	−0.268***	−0.074	15.189
	(0.023)	(0.049)	[0.000]
Pct Urban	0.32***	−0.013	22.469
	(0.024)	(0.07)	[0.000]
Pct Foreign	0.112***	0.001	3.131
	(0.025)	(0.065)	[0.077]
Pct Aged 65+	0.372***	−0.064	48.261
	(0.021)	(0.063)	[0.000]
Pct Married	0.359***	−0.060	41.331
	(0.02)	(0.066)	[0.000]
Pct Literate	0.436***	−0.036	59.355
	(0.021)	(0.063)	[0.000]
Pct Unemployed	0.385***	−0.042	34.372
	(0.021)	(0.075)	[0.000]
Occscore	0.266***	−0.059	20.087
	(0.023)	(0.074)	[0.000]
Pct Manuf.	0.202***	0.155*	0.355
	(0.023)	(0.083)	[0.551]
Pct Agriculture	−0.273***	0.068	25.442
	(0.023)	(0.069)	[0.000]
Farms Pc	−0.214***	0.054	15.245
	(0.025)	(0.070)	[0.000]
Pct Republican	0.245***	0.040	17.553
	(0.021)	(0.049)	[0.013]
Turnout	0.332***	−0.026	24.078
	(0.046)	(0.062)	[0.000]
IHS Manuf. Est.	0.434***	0.003	119.449
	(0.023)	(0.033)	[0.000]
IHS Manuf. Value	0.387***	0.032	55.132
	(0.021)	(0.045)	[0.000]
IHS Dwellings	0.435***	−0.009	328.529
	(0.023)	(0.009)	[0.000]
IHS Density	0.524***	0.000	424.003
	(0.023)	(0.010)	[0.000]
Land Value	0.253***	0.049	12.554
	(0.02)	(0.061)	[0.000]
Observations	2,137	2,137	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Table presents balancedness test for radio networks signal strength. Each row displays a separate OLS estimate where the dependent variable is a 1920 county-level variable and the independent variable is radio networks signal expressed in standard deviations. The left-most column indicates the outcome variable; column 1 shows the estimates of radio networks signal in an unconditional specification; column 2 shows the same estimates presented in Figure 2 which controls for the radio networks signal free of topographic obstacles, geographic characteristics and state fixed effects; columns 1 and 2 show standard errors clustered at the county-level below the point estimate in parenthesis. Column 3 presents the Wald test for the difference in the coefficients and in brackets the p-value associated to the test.

Table A.5: Most and Least Distinctively Foreign Names, 1930

	Most Foreign		Least Foreign	
	Males	Females	Males	Females
1	Heinz	Erika	Booker	Callie
2	Horst	Ilse	Odis	Robbie
3	Helmut	Herta	Denver	Willie Mae
4	Gunther	Irmgard	Olen	Mozelle
5	Kurt	Gisela	Virgie	Vergie
6	Egon	Elfriede	Jewell	Annie Mae
7	Sven	Elly	Early	Mittie
8	Hans	Gerda	Otha	Twila
9	Siegfried	Simone	Kyle	Ruby Lee
10	Sergio	Britta	Talmage	Mozell

Note: Table presents ten most and least distinctively foreign names from the 1930 birth cohort for names with a at least 100 observations. Source of the data is the full count US census (Ruggles et al. 2021)

Table A.6: Network Radio Increased Intermarriage with Natives for Female Immigrants

	Likelihood Intermarriage					
	(1)	(2)	(3)	(4)	(5)	(6)
Network × Post	-0.015^{**}	0.032^{**}	0.042^{***}	0.044^{***}	0.028^{**}	0.029^{**}
	(0.007)	(0.015)	(0.015)	(0.016)	(0.014)	(0.013)
Baseline	✓	✓	✓	✓	✓	✓
Free Signal		✓	✓	✓	✓	✓
Geographic			✓	✓	✓	✓
Age				✓	✓	✓
County					✓	✓
Political						✓
County FE	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓
Observations	5,918,031	5,918,031	5,918,031	5,918,031	5,918,031	5,918,031
Adjusted R^2	0.19	0.19	0.19	0.20	0.20	0.20
Mean Outcome	0.184	0.184	0.184	0.184	0.184	0.184
Sd Outcome	0.388	0.388	0.388	0.388	0.388	0.388

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the estimated effect of radio networks signal strength on the likelihood the immigrant's spouse is US born. The sample includes repeated cross-sections of female immigrants below the age of 35 between 1880 and 1940. Radio networks signal strength is expressed in standard deviations. All regressions include county, decade and country of origin fixed effects and a set of baseline controls all measured in 1920 and interacted with decade fixed effects: log population, share Black, share urban, share foreign born, share literate and share in agriculture. Free signal refers to the hypothetical radio networks signal strength in the absence of ground conductivity impediment interacted with decade fixed effects. Geographic controls measure average temperature, average rainfall, average elevation, average ruggedness, distance to closest coast, distance to closest river, distance to closest lake, distance to nearest historical portage site, distance to nearest historical Indian battle, distance to nearest mine site in 1890, average rainfall risk and average potential agricultural productivity, all interacted with decade fixed effects. County-level controls adds a larger set of county characteristics measured in 1920 interacted with decade fixed effects: share male, share older than 65, share married, share unemployed, average occupational score, share in manufacturing and number of farms per capita. Mean and standard deviations of the outcome variables refer to the pre-treatment period. Standard errors in parentheses are clustered at the county-level.

Table A.7: Placebo Analysis on Intermarriage for Immigrants not in Wedding Age

	Likelihood Intermarriage							
	Male				Female			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Network × Post	−0.011***	0.010	0.009	0.009	−0.003	0.007	0.004	0.003
	(0.004)	(0.008)	(0.007)	(0.007)	(0.002)	(0.005)	(0.005)	(0.005)
Baseline	✓	✓	✓	✓	✓	✓	✓	✓
Free Signal		✓	✓	✓		✓	✓	✓
Geographic			✓	✓			✓	✓
Individual			✓	✓			✓	✓
Aggregate			✓	✓			✓	✓
Political				✓				✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	15,779,444	15,779,444	15,779,444	15,779,444	11,951,922	11,951,922	11,951,922	11,951,922
Adjusted R^2	0.13	0.13	0.14	0.14	0.15	0.15	0.16	0.16
Mean Outcome	0.249	0.249	0.249	0.249	0.249	0.249	0.249	0.249
Sd Outcome	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the estimated effect of radio networks signal strength on the likelihood the immigrant's spouse is US born. The sample includes repeated cross-sections of immigrants above the age of 35 between 1880 and 1940. Columns 1 to 4 show results on male immigrants whereas columns 5 to 8 on female immigrants. Radio networks signal strength is expressed in standard deviations. All regressions include county, decade and country of origin fixed effects and a set of baseline controls all measured in 1920 and interacted with decade fixed effects: log population, share Black, share urban, share foreign born, share literate and share in agriculture. Free signal refers to the hypothetical radio networks signal strength in the absence of ground conductivity impediment interacted with decade fixed effects. Geographic controls measure average temperature, average rainfall, average elevation, average ruggedness, distance to closest coast, distance to closest river, distance to closest lake, distance to nearest historical portage site, distance to nearest historical Indian battle, distance to nearest mine site in 1890, average rainfall risk and average potential agricultural productivity, all interacted with decade fixed effects. County-level controls adds a larger set of county characteristics measured in 1920 interacted with decade fixed effects: share male, share older than 65, share married, share unemployed, average occupational score, share in manufacturing and number of farms per capita. Mean and standard deviations of the outcome variables refer to the pre-treatment period. Standard errors in parentheses are clustered at the county-level.

Table A.8: Network Radio Increased Assimilation Effort by Immigrants, Female Sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Naturalization					
Network × Post	0.005	0.029**	0.025**	0.020*	0.025**	0.026**
	(0.004)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)
Observations	16,630,591	16,630,591	16,630,591	16,630,591	16,630,591	16,630,591
Adjusted R^2	0.11	0.11	0.11	0.13	0.13	0.13
Mean Outcome	0.514	0.514	0.514	0.514	0.514	0.514
Sd Outcome	0.500	0.500	0.500	0.500	0.500	0.500
	Panel B: IHS FNI					
Network × Post	−0.008	−0.058**	−0.057***	−0.055***	−0.061***	−0.063***
	(0.007)	(0.023)	(0.021)	(0.021)	(0.020)	(0.020)
Observations	735,780	735,780	735,780	735,780	735,780	735,780
Adjusted R^2	0.05	0.05	0.05	0.05	0.05	0.05
Mean Outcome	4.549	4.549	4.549	4.549	4.549	4.549
Sd Outcome	0.716	0.716	0.716	0.716	0.716	0.716
Baseline	✓	✓	✓	✓	✓	✓
Free Signal		✓	✓	✓	✓	✓
Geographic			✓	✓	✓	✓
Age				✓	✓	✓
County					✓	✓
Political						✓
County FE	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the estimated impact of radio networks on naturalization and naming decisions. In panel A the sample includes all female children born at census year between 1870 to 1940 from foreign born parents. The dependent variable is the inverse hyperbolic sine of the name foreignness index assigned to the newborns. In panel B the sample is composed by all female immigrants older than twenty years old present in the US at census year from 1920 to 1940. The dependent variable is an indicator equal to one if the immigrant is naturalized US citizen. Radio networks signal strength is expressed in standard deviations. All regressions include county, decade and country of origin fixed effects and a set of baseline controls all measured in 1920 and interacted with decade fixed effects: log population, share Black, share urban, share foreign born, share literate and share in agriculture. Free signal refers to the hypothetical radio networks signal strength in the absence of ground conductivity impediment interacted with decade fixed effects. Geographic controls measure average temperature, average rainfall, average elevation, average ruggedness, distance to closest coast, distance to closest river, distance to closest lake, distance to nearest historical portage site, distance to nearest historical Indian battle, distance to nearest mine site in 1890, average rainfall risk and average potential agricultural productivity, all interacted with decade fixed effects. County-level controls adds a larger set of county characteristics measured in 1920 interacted with decade fixed effects: share male, share older than 65, share married, share unemployed, average occupational score, share in manufacturing and number of farms per capita. Mean and standard deviations of the outcome variables refer to the pre-treatment period. Standard errors in parentheses are clustered at the county-level.

Table A.9: Alternative Methods of Measuring FNI Leaves Results Unchanged

	IHS FNI			Log FNI			FNI		
N. Cohorts	20	10	5	20	10	5	20	10	5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Network × Post	−0.091***	−0.098***	−0.094***	−0.124***	−0.146***	−0.162***	−3.422***	−3.521***	−2.930***
	(0.018)	(0.020)	(0.028)	(0.031)	(0.040)	(0.059)	(0.601)	(0.581)	(0.621)
Baseline	✓	✓	✓	✓	✓	✓	✓	✓	✓
Free Signal	✓	✓	✓	✓	✓	✓	✓	✓	✓
Geographic	✓	✓	✓	✓	✓	✓	✓	✓	✓
Age	✓	✓	✓	✓	✓	✓	✓	✓	✓
County	✓	✓	✓	✓	✓	✓	✓	✓	✓
Political	✓	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	741,175	741,151	741,078	741,175	741,151	741,078	741,175	741,151	741,078
Adjusted R^2	0.07	0.04	0.03	0.03	0.02	0.02	0.14	0.12	0.10
Mean Outcome	4.571	4.396	4.515	3.797	3.660	3.389	55.216	53.824	51.473

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents robustness checks pertaining the measurement of the FNI. FNI measured in IHS, logs and in levels and constructed using either 5 10 or 20 birth cohorts of names. Sample includes all male children born on every census decade from 1870 to 1940 from foreign born parents. Controls and fixed effects are the same as in the most saturated specification from column 6 of main Table 3. Standard errors in parentheses are clustered at the county-level.

Table A.10: Network Radio Boosted English Proficiency for Recent Female Immigrants

	(1)	(2)	(3)	(4)	(5)	(6)
Network × Post	-0.021^{***}	0.056^{**}	0.087^{***}	0.091^{***}	0.078^{***}	0.073^{***}
	(0.008)	(0.022)	(0.026)	(0.026)	(0.022)	(0.021)
Baseline	✓	✓	✓	✓	✓	✓
Free Signal		✓	✓	✓	✓	✓
Geographic			✓	✓	✓	✓
Age				✓	✓	✓
County					✓	✓
Political						✓
County FE	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓
Observations	1,804,892	1,804,892	1,804,892	1,804,892	1,804,892	1,804,892
Adjusted R^2	0.19	0.19	0.19	0.22	0.22	0.22
Mean Outcome	0.534	0.534	0.534	0.534	0.534	0.534
Sd Outcome	0.499	0.499	0.499	0.499	0.499	0.499

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the estimated effect of radio networks signal strength on the likelihood the immigrant's declare to speak proper English. The sample includes repeated cross-sections of female immigrants above the age of 10 between 1900 and 1930. I restrict the sample to recent immigrants who arrived in the US within five years from census year and exclude from the analysis immigrants from countries where English was the first language. Radio networks signal strength is expressed in standard deviations. All regressions include county, decade and country of origin fixed effects and a set of baseline controls all measured in 1920 and interacted with decade fixed effects: log population, share Black, share urban, share foreign born, share literate and share in agriculture. Free signal refers to the hypothetical radio networks signal strength in the absence of ground conductivity impediment interacted with decade fixed effects. Geographic controls measure average temperature, average rainfall, average elevation, average ruggedness, distance to closest coast, distance to closest river, distance to closest lake, distance to nearest historical portage site, distance to nearest historical Indian battle, distance to nearest mine site in 1890, average rainfall risk and average potential agricultural productivity, all interacted with decade fixed effects. County-level controls adds a larger set of county characteristics measured in 1920 interacted with decade fixed effects: share male, share older than 65, share married, share unemployed, average occupational score, share in manufacturing and number of farms per capita. Mean and standard deviations of the outcome variables refer to the pre-treatment period. Standard errors in parentheses are clustered at the county-level.

Table A.11: Intermarriage and Naturalization Results Robust to Potential Historical Confounders

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Panel A: Intermarriage													
Network × Post	0.060***	0.059***	0.059***	0.055***	0.060***	0.055***	0.055***	0.059***	0.060***	0.058***	0.058***	0.054***	0.053***	0.060***
	(0.020)	(0.020)	(0.020)	(0.018)	(0.020)	(0.020)	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)
Observations	4,575,324	4,575,312	4,575,312	4,575,274	4,572,555	4,572,528	4,572,528	4,572,528	4,572,555	4,572,528	4,572,517	4,572,555	4,572,555	4,575,324
Adjusted R^2	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
	Panel B: Naturalization													
Network × Post	0.025***	0.025***	0.025***	0.024**	0.024***	0.024**	0.025***	0.025***	0.025***	0.025***	0.025***	0.023**	0.022**	0.024**
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
Observations	27,956,723	27,956,656	27,956,656	27,956,335	27,931,783	27,931,565	27,931,565	27,931,565	27,931,783	27,931,565	27,931,565	27,931,783	27,931,783	27,956,723
Adjusted R^2	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
Post × *Historical Shock*:														
Post × $\Delta Black_{30-10}$		✓												
– × $\Delta Black_{m,30-10}$			✓											
– × Quota Exp.				✓										
– × RDF_{39-33}					✓									
– × RDF_{33-29}						✓								
– × RDF_{39-29}							✓							
– × $RLDF_{33-29}$								✓						
– × $RLDF_{39-33}$									✓					
– × $RLDF_{39-29}$										✓				
– × RDF_{35-29}											✓			
– × DRPCAA												✓		
– × DRPCAAA													✓	
– × Non-Network														✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the robustness tests to the sequential inclusion of a set of potential historical confounding factors. Panel A shows the results for intermarriage where column 1 displays the estimate from column 6 of main Table 2. Panel B shows the results for naturalization where column 1 displays the estimate from column 6 of main Table 3. Radio networks signal strength is expressed in standard deviations. Controls and fixed effects as in the most saturated specification in the main analysis. The historical shocks investigated are the following: the change in the share of the Black population between 1930 and 1910 and the change in the share of the Black male population between 1930 and 1910; county level exposure to the quota effect from (Ager et al. 2023); Measures of exposure to the Great Depression, that is the change in the retail sale per capita between 1939 and 1933, 1933 and 1929, 1939 and 1929; and the change in the log sales per capita (growth rate) between 1933 and 1929, 1939 and 1933, 1939 and 1929, 1935 and 1929; New Deal spending measured as the per capita spending in AAA grants between 1933 and 1937 and per capita AAA grants from 1933 and through 1935; signal strength variation in non-network signal strength (Fishback and Shawn 2018). Standard errors in parentheses are clustered at the county-level.

Table A.12: FNI and English Proficiency Results Robust to Potential Historical Confounders

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Panel A: IHS FNI													
Network × Post	−0.091***	−0.091***	−0.091***	−0.091***	−0.091***	−0.094***	−0.093***	−0.091***	−0.090***	−0.093***	−0.092***	−0.080***	−0.079***	−0.086***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)	(0.018)	(0.017)	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)
Observations	741,443	741,443	741,443	741,440	741,032	741,025	741,025	741,025	741,032	741,025	741,024	741,032	741,032	741,443
Adjusted R^2	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
	Panel B: English Proficiency													
Network × Post	0.079***	0.080***	0.080***	0.082***	0.079***	0.081***	0.082***	0.080***	0.080***	0.080***	0.080***	0.081***	0.080***	0.075***
	(0.022)	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.021)	(0.021)	(0.022)	(0.021)	(0.022)	(0.022)	(0.022)	(0.023)
Observations	1,985,965	1,985,960	1,985,960	1,985,949	1,985,116	1,985,104	1,985,104	1,985,104	1,985,116	1,985,104	1,985,104	1,985,116	1,985,116	1,985,965
Adjusted R^2	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Post × *Historical Shock*:														
– × $\Delta Black_{30-10}$		✓												
– × $\Delta Black_{m,30-10}$			✓											
– × Quota Exp.				✓										
– × RDF_{39-33}					✓									
– × RDF_{33-29}						✓								
– × RDF_{39-29}							✓							
– × $RLDF_{33-29}$								✓						
– × $RLDF_{39-33}$									✓					
– × $RLDF_{39-29}$										✓				
– × RDF_{35-29}											✓			
– × DRPCAA												✓		
– × DRPCAAA													✓	
– × Non-Network														✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents robustness tests to the sequential inclusion of a set of potential historical confounding factors. Panel A shows the results for FNI where column 1 displays the estimate from column 6 of main Table 3. Panel B shows the results for English proficiency where column 1 displays the estimate from column 6 of main Table 4. Radio networks signal strength is expressed in standard deviations. Controls and fixed effects as in the most saturated specification in the main analysis. The historical shocks investigated are the following: the change in the share of the Black population between 1930 and 1910 and the change in the share of the Black male population between 1930 and 1910; county level exposure to the quota effect from (Ager et al. 2023); Measures of exposure to the Great Depression, that is the change in the retail sale per capita between 1939 and 1933, 1933 and 1929, 1939 and 1929; and the change in the log sales per capita (growth rate) between 1933 and 1929, 1939 and 1933, 1939 and 1929, 1935 and 1929; New Deal spending measured as the per capita spending in AAA grants between 1933 and 1937 and per capita AAA grants from 1933 and through 1935; signal strength variation in non-network signal strength (Fishback and Shawn 2018). Standard errors in parentheses are clustered at the county-level.

Table A.13: Main Results Robust to Distances to Stations and Urban Centers

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Intermarriage					
Network × Post	0.060***	0.059***	0.059***	0.054***	0.060***	0.054***
	(0.020)	(0.021)	(0.021)	(0.019)	(0.020)	(0.020)
Observations	4,575,324	4,575,324	4,575,324	4,575,324	4,575,324	4,575,324
Adjusted R^2	0.17	0.17	0.17	0.17	0.17	0.17
	Panel B: Naturalization					
Network × Post	0.025***	0.026***	0.026***	0.028***	0.025***	0.030***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)
Observations	27,956,723	27,956,723	27,956,723	27,956,723	27,956,723	27,956,723
Adjusted R^2	0.19	0.19	0.19	0.19	0.19	0.19
	Panel C: IHS FNI					
Network × Post	−0.091***	−0.096***	−0.096***	−0.097***	−0.090***	−0.086***
	(0.018)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)
Observations	741,443	741,443	741,443	741,443	741,443	741,443
Adjusted R^2	0.07	0.07	0.07	0.07	0.07	0.07
	Panel D: English Proficiency					
Network × Post	0.079***	0.081***	0.080***	0.075***	0.075***	0.077***
	(0.022)	(0.022)	(0.022)	(0.021)	(0.021)	(0.021)
Observations	1,985,965	1,985,965	1,985,965	1,985,965	1,985,965	1,985,965
Adjusted R^2	0.20	0.20	0.20	0.20	0.20	0.20
Post × *Distance to*:						
– × Network Antenna		✓	✓	✓		
– × Non-Network Antenna			✓			
– × Five Antennas				✓		
– × State Capital					✓	
– × DC						✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the robustness tests to the sequential inclusion of a set of distances to radio stations and urban centers. Panel A shows results for intermarriage, Panel B for naturalization, Panel C for FNI and Panel D for English proficiency. Column 1 displays the estimate from column 6 of main Tables 2, 3 and 4. Radio networks signal strength is expressed in standard deviations. Controls and fixed effects as in the most saturated specification in the main analysis. The distances included in the analysis are: distance to the closest network antenna, distance to the closest non-network antenna, distance to the five closest antennas, distance to the state capital and distance to DC. Standard errors in parentheses are clustered at the county-level.

Table A.14: Main Results Robust to Alternative Sample Selections

	(1)	(2)	(3)	(4)
		Panel A: Intermarriage		
Network × Post	0.060***	0.055**	0.059***	0.049***
	(0.020)	(0.021)	(0.021)	(0.017)
Observations	4,575,324	4,406,316	4,564,360	2,573,549
R^2	0.17	0.17	0.17	0.18
Mean Outcome	0.285	0.285	0.285	0.285
Sd Outcome	0.452	0.452	0.452	0.452
		Panel B: Naturalization		
Network × Post	0.025***	0.019**	0.025***	0.026***
	(0.009)	(0.010)	(0.010)	(0.009)
Observations	27,956,723	26,804,628	27,880,949	16,590,362
Adjusted R^2	0.19	0.18	0.18	0.19
Mean Outcome	0.588	0.588	0.588	0.588
Sd Outcome	0.492	0.492	0.492	0.492
		Panel C: IHS FNI		
Network × Post	−0.091***	−0.083***	−0.091***	−0.070***
	(0.018)	(0.018)	(0.018)	(0.018)
Observations	741,443	715,098	739,816	442,559
Adjusted R^2	0.07	0.07	0.07	0.07
Mean Outcome	4.571	4.571	4.571	4.571
Sd Outcome	0.700	0.700	0.700	0.700
		Panel D: English Proficiency		
Network × Post	0.079***	0.075***	0.079***	0.093***
	(0.022)	(0.023)	(0.022)	(0.020)
Observations	1,985,965	1,867,192	1,974,472	1,107,663
R^2	0.20	0.19	0.20	0.22
Mean Outcome	0.535	0.535	0.535	0.535
Sd Outcome	0.499	0.499	0.499	0.499
Sample	Full	No South	No Territories	No Antennas

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the robustness tests to alternative sample selections. Panel A shows results for intermarriage, Panel B for naturalization, Panel C for FNI and Panel D for English proficiency. Column 1 displays the estimates on the full sample from column 6 of main Tables 2, 3 and 4. Radio networks signal strength is expressed in standard deviations. Controls and fixed effects as in the most saturated specification in the main analysis. Column 2 excludes the geographical US South, column 3 excludes what were still US territories in the early 1910s, ie New Mexico and Arizona and column 4 excludes counties where the network antennas were located. Standard errors in parentheses are clustered at the county-level.

Table A.15: Results are Robust to the Inclusion of Further Controls and Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Intermarriage							
Network × Post	0.060***	0.068***	0.068***	0.065***	0.065***	0.065***	0.063***	0.061***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)	(0.019)
Observations	4,575,274	4,068,309	4,068,309	4,068,309	4,068,309	4,068,309	4,068,309	4,068,309
Adjusted R^2	0.17	0.19	0.19	0.20	0.20	0.20	0.20	0.20
Mean Outcome	0.285	0.272	0.272	0.272	0.272	0.272	0.272	0.272
Sd Outcome	456	0.445	0.445	0.445	0.445	0.445	0.445	0.445
	Panel B: Naturalization							
Network × Post	0.025***	0.024**	0.023**	0.023**	0.021**	0.021**	0.021**	0.020*
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Observations	27,956,723	27,956,723	27,956,723	27,956,723	27,956,723	27,956,723	27,956,723	27,956,723
Adjusted R^2	0.19	0.19	0.20	0.21	0.22	0.22	0.22	0.22
Mean Outcome	0.588	0.588	0.588	0.588	0.588	0.588	0.588	0.588
Sd Outcome	0.492	0.492	0.492	0.492	0.492	0.492	0.492	0.492
	Panel C: IHS FNI							
Network × Post	−0.091***	−0.080***	−0.080***	−0.078***	−0.078***	−0.078***	−0.074***	−0.075***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)	(0.017)
Observations	741,175	534,759	534,759	534,759	534,759	534,759	534,759	534,759
Adjusted R^2	0.07	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Mean Outcome	4.571	4.565	4.565	4.565	4.565	4.565	4.565	4.565
Sd Outcome	0.700	0.748	0.748	0.748	0.748	0.748	0.748	0.748
	Panel D: English Proficiency							
Network × Post	0.079***	0.082***	0.083***	0.078***	0.076***	0.074***	0.070***	0.070***
	(0.022)	(0.022)	(0.022)	(0.021)	(0.021)	(0.021)	(0.020)	(0.020)
Observations	1,985,965	1,985,965	1,985,965	1,985,965	1,985,965	1,985,965	1,985,965	1,985,965
Adjusted R^2	0.20	0.21	0.22	0.24	0.24	0.24	0.24	0.24
Mean Outcome	0.535	0.535	0.535	0.535	0.535	0.535	0.535	0.535
Sd Outcome	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499
Controls	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓	✓	✓
Fam Size		✓	✓	✓	✓	✓	✓	✓
Employed			✓	✓	✓	✓	✓	✓
Occ Score				✓	✓	✓	✓	✓
Ownership					✓	✓	✓	✓
Urban						✓	✓	✓
Region by Decade FE							✓	✓
Ancestry by Decade FE								✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the robustness tests to the addition of further controls and fixed effects. Panel A shows results for intermarriage, Panel B for naturalization, Panel C for FNI and Panel D for English proficiency. Column 1 displays the estimates from column 6 of main Tables 2, 3 and 4. Radio networks signal strength is expressed in standard deviations. The set of additional controls are measured at the household level and include: number of family members, home ownership status, urban status, employment status, occupational score. In the FNI sample, the last two controls refer to the household head. The additional fixed effect included are region by decade fixed effects and ancestry by decade fixed effects. Standard errors in parentheses are clustered at the county-level.

Table A.16: Alternative Inference, Intermarriage and Naturalization

	Male			Female		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Intermarriage					
Network × Post	0.071***	0.059***	0.060***	0.044***	0.028**	0.029**
	(0.023)	(0.021)	(0.020)	(0.016)	(0.014)	(0.013)
Clustered State	[0.021]***	[0.022]**	[0.022]***	[0.015]***	[0.014]**	[0.013]**
Conley k = 100mi	[0.015]***	[0.022]***	[0.021]***	[0.014]***	[0.014]**	[0.014]**
Conley k Est.	[0.021]***	[0.018]***	[0.021]***	[0.012]***	[0.006]***	[0.007]***
Observations	4,575,324	4,575,324	4,575,324	5,918,031	5,918,031	5,918,031
Adjusted R^2	0.17	0.17	0.17	0.20	0.20	0.20
Mean Outcome	0.285	0.285	0.285	0.184	0.184	0.184
Sd Outcome	0.452	0.452	0.452	0.388	0.388	0.388
	Panel B: Naturalization					
Network × Post	0.026***	0.025***	0.025***	0.020*	0.025**	0.026**
	(0.010)	(0.009)	(0.009)	(0.011)	(0.011)	(0.011)
Clustered State	[0.010]**	[0.010]**	[0.010]**	[0.012]*	[0.011]**	[0.011]**
Conley k = 100mi	[0.012]**	[0.008]***	[0.008]***	[0.013]	[0.012]**	[0.014]*
Conley k Est.	[0.007]***	[0.013]*	[0.019]	[0.006]***	[0.012]**	[0.014]*
Observations	27,956,723	27,956,723	27,956,723	16,630,591	16,630,591	16,630,591
Adjusted R^2	0.18	0.19	0.19	0.13	0.13	0.13
Mean Outcome	0.588	0.588	0.588	0.514	0.514	0.514
Sd Outcome	0.492	0.492	0.492	0.500	0.500	0.500
Baseline	✓	✓	✓	✓	✓	✓
Free Signal	✓	✓	✓	✓	✓	✓
Geographic	✓	✓	✓	✓	✓	✓
Individual	✓	✓	✓	✓	✓	✓
Aggregate		✓	✓		✓	✓
Political			✓			✓
County FE	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents alternative inference for the main results. Panel A shows the results for intermarriage whereas Panel B for naturalization. Top row displays point estimates and standard errors referring to the last three specifications from the main results on the male and female sample, respectively. Radio networks signal strength is expressed in standard deviations. Alternative standard errors reported in square parentheses below the main row with relative significance stars: standard errors clustered at the state-level, standard errors robust to spatial correlation with cutoff at 100 miles and standard errors robust to spatial correlation with optimally estimated cutoff.

Table A.17: Alternative Inference, FNI and English Proficiency

	Male			Female		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: IHS FNI					
Network × Post	−0.092***	−0.089***	−0.091***	−0.055***	−0.061***	−0.063***
	(0.020)	(0.020)	(0.018)	(0.021)	(0.020)	(0.019)
Clustered State	[0.028]***	[0.025]***	[0.025]***	[0.027]*	[0.023]**	[0.023]***
Conley k = 100mi	[0.026]***	[0.024]***	[0.023]***	[0.027]**	[0.023]***	[0.024]***
Conley k Est.	[0.031]***	[0.031]***	[0.029]***	[0.024]**	[0.022]***	[0.022]***
Observations	741,175	741,175	741,175	722,835	722,835	722,835
Adjusted R^2	0.07	0.07	0.07	0.05	0.05	0.05
Mean Outcome	4.571	4.571	4.571	4.555	4.555	4.555
Sd Outcome	0.700	0.700	0.700	0.711	0.711	0.711
	Panel B: English Proficiency					
Network × Post	0.093***	0.084***	0.079***	0.091***	0.078***	0.073***
	(0.025)	(0.022)	(0.022)	(0.026)	(0.022)	(0.021)
Clustered State	[0.021]***	[0.017]***	[0.015]***	[0.026]***	[0.021]***	[0.020]***
Conley k = 100mi	[0.023]***	[0.016]***	[0.016]***	[0.033]***	[0.025]***	[0.024]***
Conley k Est.	[0.022]***	[0.017]***	[0.018]***	[0.027]***	[0.026]***	[0.039]*
Observations	1,985,965	1,985,965	1,985,965	1,804,892	1,804,892	1,804,892
Adjusted R^2	0.20	0.20	0.20	0.22	0.22	0.22
Mean Outcome	0.535	0.535	0.535	0.534	0.534	0.534
Sd Outcome	0.499	0.499	0.499	0.499	0.499	0.499
Baseline	✓	✓	✓	✓	✓	✓
Free Signal	✓	✓	✓	✓	✓	✓
Geographic	✓	✓	✓	✓	✓	✓
Individual	✓	✓	✓	✓	✓	✓
Aggregate		✓	✓		✓	✓
Political			✓			✓
County FE	✓	✓	✓	✓	✓	✓
Ancestry FE	✓	✓	✓	✓	✓	✓
Decade FE	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents alternative inference for the main results. Panel A shows the results for FNI whereas Panel B for English proficiency. Top row displays point estimates and standard errors referring to the last three specifications from the main results on the male and female sample, respectively. Radio networks signal strength is expressed in standard deviations. Alternative standard errors reported in square parentheses below the main row with relative significance stars: standard errors clustered at the state-level, standard errors robust to spatial correlation with cutoff at 100 miles and standard errors robust to spatial correlation with optimally estimated cutoff.

Table A.18: Radio Networks Increased Likelihood of Naming After Successful Baseball Names, Native Sample

	Baseball Name				
	(1)	(2)	(3)	(4)	(5)
Network × Post	0.009***	0.009***	0.009***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
– × Any Award		0.030***			
		(0.005)			
– × WAR^{+}_{1sd}			0.001		0.000
			(0.003)		(0.002)
– × WAR^{-}_{1sd}				−0.005***	−0.005***
				(0.002)	(0.002)
Birth Cohort FE	✓	✓	✓	✓	✓
Household FE	✓	✓	✓	✓	✓
Observations	17,247,674	17,247,674	17,247,674	17,247,674	17,247,674
Adjusted R^2	0.69	0.69	0.69	0.74	0.74
Mean Outcome	0.295	0.295	0.295	0.295	0.295
Sd Outcome	0.456	0.456	0.456	0.456	0.456

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Estimated effect of standardized network signal strength on the likelihood of naming a child after a baseball player active during the same year the child was born interacted with the performance of the player. The sample is composed by white children from a household with both parents native. All regressions account for household and birth cohort fixed effects. Controls include 1920 IHS population, *free* network coverage interacted with birth cohort fixed effects and the share of games played associated to the baseball name. The model estimates the interaction between radio networks signal variation, the post-1929 dummy and a series of performance indicators: any award during the season, scoring a WAR value that is one standard deviation above the season mean WAR or scoring a WAR value that is one standard deviation below the season mean WAR. Although not shown in the table, the model estimates all lower levels interaction. I cluster standard errors at the county-level.

Table A.19: Radio Networks Increased Likelihood of Naming After Names that Received an Award, Immigrant Sample

	Baseball Name				
	(1)	(2)	(3)	(4)	(5)
Post × Network	0.008	0.008	0.008	0.009	0.008
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
– × Any Award	0.029***				
	(0.011)				
– × TSN Allstar		0.042***			
		(0.016)			
– × TSN MVP			0.016		
			(0.043)		
– × N. awards				0.009	
				(0.006)	
– × BM Allstar					0.034***
					(0.011)
Birth Cohort FE	✓	✓	✓	✓	✓
Household FE	✓	✓	✓	✓	✓
Observations	1,849,074	1,849,074	1,849,074	1,849,074	1,849,074
Adjusted R^2	0.71	0.71	0.71	0.71	0.71
Mean Outcome	0.281	0.281	0.281	0.281	0.281
Sd Outcome	0.450	0.450	0.450	0.450	0.450

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Estimated effect of standardized network signal strength on the likelihood of naming a child after a baseball player active during the same year the child was born interacted with the performance of the player. The sample is composed by white children from a household with both parents foreign. All regressions account for household and birth cohort fixed effects. Controls include 1920 IHS population, *free* network coverage interacted with birth cohort fixed effects and the share of games played associated to the baseball name. The model estimates the interaction between radio networks signal variation, the post-1929 dummy and a series of performance indicators: any award during the season, being featured in the TSN All Star team, being featured as the TSN MVP, the numbers of awards during a season and being featured in the Baseball Magazine All Star team. Although not shown in the table, the model estimates all lower levels interaction. I cluster standard errors at the county-level.

Table A.20: Radio Networks Increased Likelihood of Naming After Names that Received an Award, Native Sample

	Baseball Name				
	(1)	(2)	(3)	(4)	(5)
Post × Network	0.009***	0.009***	0.009***	0.009***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
– × Any Award	0.030***				
	(0.005)				
– × TSN Allstar		0.064***			
		(0.006)			
– × TSN MVP			0.053***		
			(0.011)		
– × N. awards				0.013***	
				(0.003)	
– × BM Allstar					0.035***
					(0.005)
Birth Cohort FE	✓	✓	✓	✓	✓
Household FE	✓	✓	✓	✓	✓
Observations	17,247,674	17,247,674	17,247,674	17,247,674	17,247,674
Adjusted R^2	0.69	0.69	0.69	0.69	0.69
Mean Outcome	0.295	0.295	0.295	0.295	0.295
Sd Outcome	0.456	0.456	0.456	0.456	0.456

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Estimated effect of standardized network signal strength on the likelihood of naming a child after a baseball player active during the same year the child was born interacted with the performance of the player. The sample is composed by white children from a household with both parents native. All regressions account for household and birth cohort fixed effects. Controls include 1920 IHS population, *free* network coverage interacted with birth cohort fixed effects and the share of games played associated to the baseball name. The model estimates the interaction between radio networks signal variation, the post-1929 dummy and a series of performance indicators: any award during the season, being featured in the TSN All Star team, being featured as the TSN MVP, the numbers of awards during a season and being featured in the Baseball Magazine All Star team. Although not shown in the table, the model estimates all lower levels interaction. I cluster standard errors at the county-level.

B Online Appendix: Radio Signal Computation (Not for publication)[26]

B.1 AM Radio Propagation

Radio propagation is the phenomenon through which radio waves travel from one point to another. Depending what radio frequency is used to broadcast, there are different ways to transmit the electromagnetic waves that constitute the radio signal. The vast majority of broadcasters use amplitude modulation (AM) when transmitting signal on frequencies associated with short and medium wave lengths. Here, I maintain this focus, discussing exclusively the propagation of radio signal on the medium frequency band through amplitude modulation.[27]

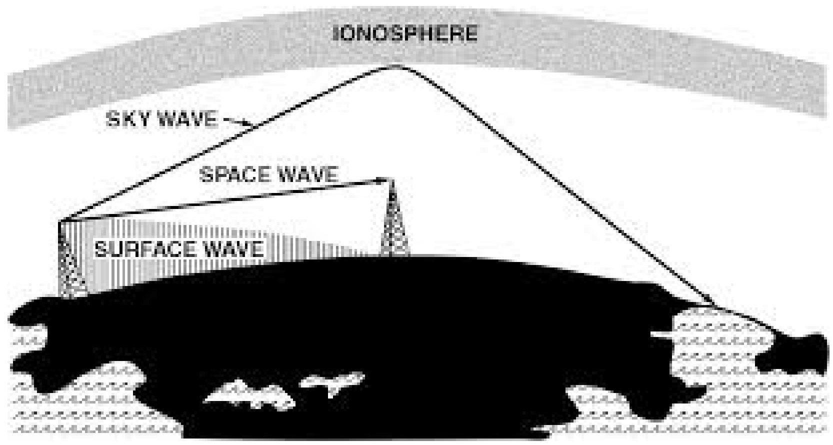
AM radio signal on the medium frequency band propagates using two type of waves: ground waves and sky waves. Sky waves rely on the ionosphere to bounce back on earth and reach receiver antennas. During the day the solar activity does not allow the ionosphere to reflect AM radio waves back on Earth. Instead, ground waves spread out from the transmitter along the surface of the Earth allowing their use at any hour of the day and the night. Hence, day time transmission of AM radio signal exclusively relies on ground waves.

To reconstruct radio coverage of US stations during the interwar period, I exclusively focus on ground waves. The reason for this is twofold. First, sky waves allow for the coverage of areas extremely far, up to 3500km from the transmitting antenna. This implies that night coverage of radio networks in the interwar period would reach the entire continental US.

[26]For fruitful conversations on the technical aspect of radio propagation I thank David Kazdan, Adjunct Assistant Professor in Electrical Engineering and Computer Science at Case Western Reserve University, William Scott (Radio-Timetraveller) and the members of the HamSCI Google group. I also would like to thank Francesco Sobbrio for helpful feedback and Caterina Mauri for sharing digitized data on US ground conductivity. Daniil Shertsnev provided excellent assistance in the implementation of the radio propagation model.

[27]The engineering definition of the portion of radio spectrum characterized by medium wave length goes from 300 kHz to 3 MHz. In practice, what is usually devoted to broadcasting goes from approximately 500 kHz to 1700 kHz. Today the Federal Communication Commission has allocated the medium frequency band to be between 530 to 1700 kHz for the United States.

Figure B.5: Sky Waves Versus Ground Waves



Second, in the second part of the paper I investigate the adoption of baseball players names. During the interwar period, games were for the vast majority played with the light of the day.[28] Thus, in what follows, I give details on the mechanics of ground wave propagation. For a discussion of sky wave propagation in economics see Gagliarducci et al. (2020).

B.1.1 Ground Waves Propagation

Ground waves encompass all waves that are transmitted by a radio antenna and propagate through the area between the surface of the earth and the ionosphere.[29] Given a transmitting antenna, the voltage (field strength) induced in the receiving antenna, at an arbitrary position can be expressed by the following expression:

$$V = QI\left\{Q_1\frac{exp(-jkr_1)}{r_1} + Q_2R\frac{exp(-jkr_2)}{r_2} + S\frac{exp(-jkr_2)}{r_2}\right\} \tag{4}$$

Where I is the current in the transmitting antenna, Q is a constant, $Q1$ and $Q2$ measure the transmitting- and receiving-antenna polar diagrams, R is a reflection coefficient, k is the radio wave number $k = 2\pi/\lambda$, r_1 is the straight line distance between transmitting and receiving antenna and r_2 is the straight line distance between transmitting antenna and the

[28]The first night game in Major League history was played in 1935. Until the 1940s only a very tiny fractions of the games were played at night.

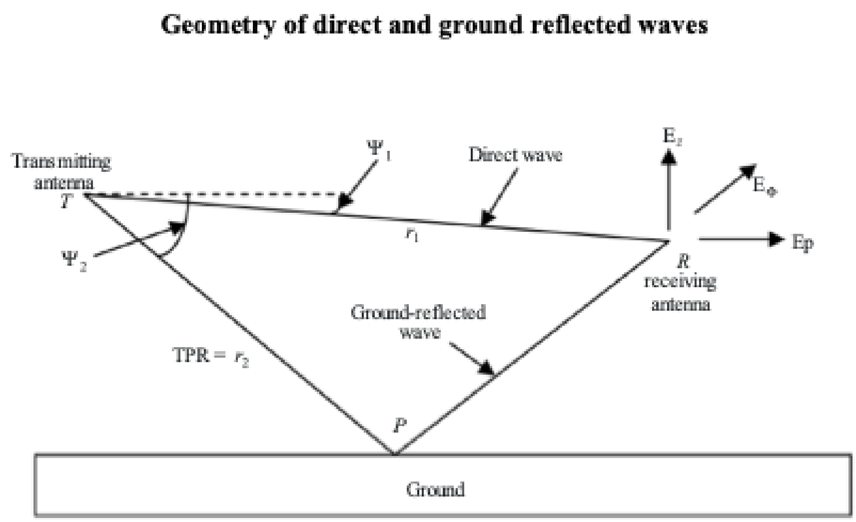
[29]This section is heavily based on the ITU handbook on radio ground waves (International Telecommunication Union 2014)

ground. The parameter S embodies an attenuation factor.

Each one of the three components in equation (4) reflects a subset of ground waves that determines field strength at the receiving end: space waves, composed by direct line waves and reflected waves, and surface waves.

Space waves are typically formed by the direct wave — transmitting antenna to receiving antenna — and the waves that leave the antenna and bounce over natural obstacles (see Figure B.6). Space waves are effective in reaching the receiving antenna only if this is within line of sight. If the receiving end is beyond the horizon, or both antennas are very close to the ground, the space waves are either zero or very negligible.

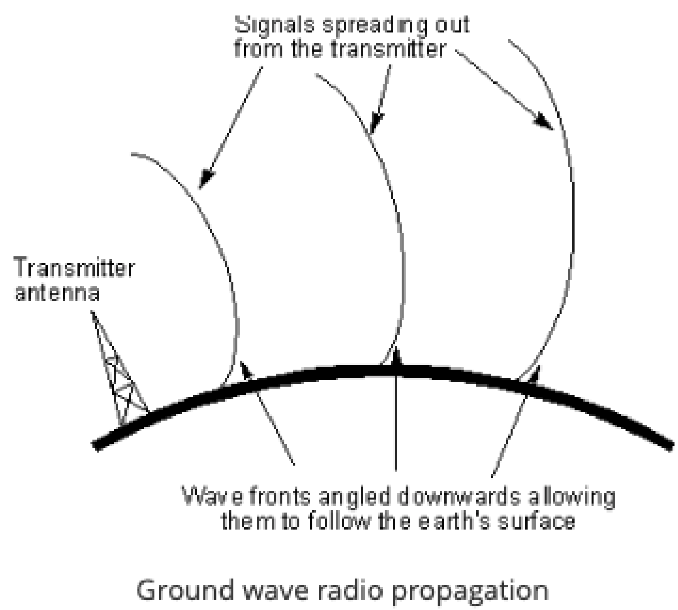
Figure B.6: Propagation of the Space Wave Component of the Ground Wave



Surface waves are sometimes referred to as ground waves themselves. Instead of just following line of sight, like direct waves do, these waves tend to follow the curvature of the Earth. The reason for this is that currents in the surface of the earth cause the front of the wave to tilt downwards towards Earth (see Figure B.7). This allows surface waves to curve around the Earth and reach receivers well beyond the horizon and beyond natural obstacles such as hills and mountains.

The rate of attenuation of the signal increases as it travels over a surface. This happens because the forward portion of the waves tilts even further. The speed with which

Figure B.7: Propagation of the Surface Wave Component of the Ground Wave



this happens is regulated by an attenuation factor called complex dielectric permittivity. Large values of the attenuation factor correspond to a small forward tilt and therefore low attenuation and vice-versa.

The attenuation factor is determined by two elements: the electrical conductivity of the surface over which the wave travels and the broadcasting frequency. Sea water has an outstandingly high conductivity and the surface wave, with a near-vertical electric field, propagates over it with relatively low attenuation. On the other hand, surface-wave attenuation is greatest over ground of low conductivity and at high radio frequencies. Within the realm of AM radio the role of the frequency band is limited. Hence, to predict field strength at the receiving end the relative importance of ground conductivity is much higher.

B.1.2 The Role of Soil Conductivity

Ground waves emanating out of the transmitting antenna pass over one or multiple surfaces before they reach a receiver. Different type of soils have different levels of electrical conductivity and affect the rate of attenuation factor of the ground wave. Hence, to confidently

Table B.21: Typical Conductivity Levels of Different Types of Soil

Soil Description	Conductivity ($\sigma_{s/m}$)	Permittivity (ϵ_r)
Polar Ice Cap	0.0001	1
Polar Ice	0.0003	3
Arctic Land	0.0005	3
Sea Ice	0.001	4
City Industrial	0.001	5
Dry, Sandy, Coastal	0.001	10
Fresh Water	0.001	80
Fertile Land	0.002	10
Rocky, Steep Hills	0.002	15
Moderate	0.003	4
Medium Hills And Forest	0.004	13
Average	0.005	13
Highly Moist Ground	0.005	30
Pastoral Hills, Rich Soil	0.007	17
Marshly Land, Densily Wooded	0.0075	12
Marshly, Forested, Flat	0.008	12
Rich Agric Land, Low Hills	0.010	15
Sea Water	5.0	81

Note: Source of the data is the Consultative Committee on International Radio (CCIR predecessor of the International Telecommunication Union) (Consultative Committee on International Radio 1992).

predict field strength at a certain distance from the transmitting antenna it is important to know the conductivity of the surface the wave is expected to travel over.

Table B.21 summarizes for some type of soil their typical levels of conductivity expressed in siemens per meter ($\sigma_{s/m}$) and their permittivity.[30] By far, sea, especially if with high salinity, is the most conductive surface with 5 $\sigma_{s/m}$. Among land surfaces, wet grounds typically have the best conductivity. This makes farming areas on average very suitable for the propagation of ground waves. Dry and desert type lands are among the least conductive.

While Table B.21 gives an idea of the conductivity for certain type of soils, measuring ground conductivity is cumbersome. Sampling soil is not enough because conductivity is de-

[30] Permittivity is closely related to conductivity. It measures the ability of a substance to hold an electrical charge. Higher permittivity decreases attenuation of the wave

termined by the type of soil even up to 25 meters below the immediate surface, requiring the use of vector network analyzers in conjunction with specific wire lines. For this reason, individuals interested in predicting field strength rely on conductivity contour maps published by International Telecommunication Union or, in the US, by the Federal Communication Commission.

B.2 Implementation of the LMLF Ground Wave Propagation Model

To predict radio ground waves we borrowed the propagation model made available from the International Telecommunication Union (ITU). The original software from the ITU is easily downloadable from the ITU webpage that collects software, data and validation exercises concerning radio wave propagation and radio noise.[31] Specifically, we retrieved the LFMF-SmoothEarth model which was designed for the ground-wave propagation for frequencies between 10 kHz and 30 MHz, a frequency span that included the frequencies used by AM radio stations during the interwar period. To fasten the computation of radio coverage we implemented the model in Python.

The model works by collecting an array of inputs and computing point to point (antenna to receiver) propagation. The output of the model is the electric field strength (signal strength) at the receiver location expressed in decibels referenced to 1 microvolt per meter (dBμV/m). The model inputs are the following:

- *h_tx_m*: height of the transmitter in meter
- *h_rx_m*: height of the receiver in meter
- *f_mhz*: frequency of broadcast in mHz
- *p_tx_watt*: power of transmission in watts
- *n_s*: surface refractivity
- *d_km*: distance in km

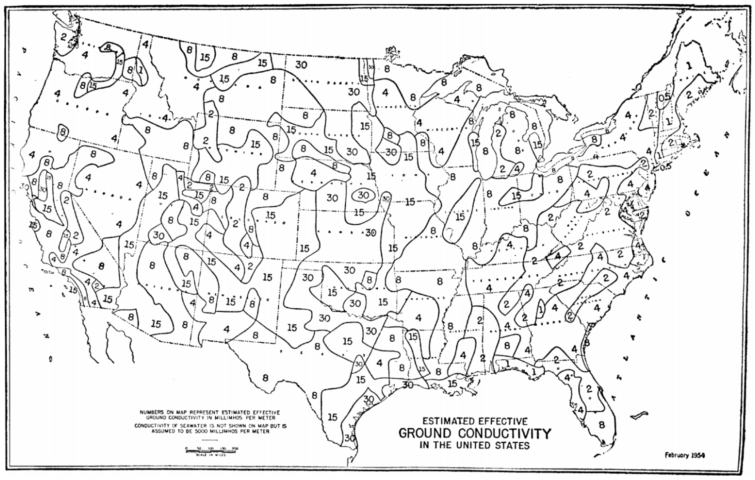
[31] Webpage available at this link

- ϵ: relative permittivity on the surface of the earth
- σ: ground conductivity (siemens by meters)
- *pol*: polarization

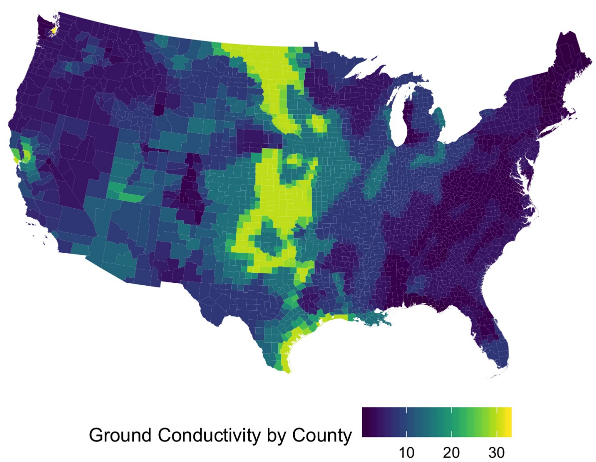
We start by collecting data on US radio stations active in January 1929. Digitizing their technical informations allows us to observe some of the parameters listed above. We run the LFMF model on each combination radio transmitter-receiver, where as receiver we use the centroid of all US counties. At this point we are able to fill in most of the inputs: transmitter heights data is partially missing, we impute it using the average in the non-missing sample. We assume receiver height to be equal to 10 meters for all centroids. We obtain frequency and power of the antennas from the station data and calculate the distance in km between all antennas and receivers. We assume surface refractivity to be constant and use the average value observed in the country that is 313 (Doerry 2013). To determine relative permittivity I use an approximation provided by the FCC that uses levels of ground conductivity to approximate values of permittivity. Given the historical and technical background, we set polarization to vertical. The only parameter left to determine is ground conductivity.

Computation of ground conductivity. We obtain county-level data on US ground conductivity. Such data is constructed by digitizing the ground conductivity map from the FCC. Figure B.8a shows the original source of the data which can be retrieved at this link. A rendering of the county level conductivity is presented in figure B.8b.

The LFMF model allows for the inclusion of a single input of soil conductivity for each station to receiver signal computation. However, in reality the radio wave emitted by a transmitter might face different types of land and conductivities. To construct this single value of conductivity I take the weighted average of all the values of conductivity faced by the signal, weighted by the length of each single conductivity segment. This procedure allows me to approximate the average conductivity faced by the radio wave, giving higher weight to the conductivities over which the wave travels longer.



(a) FCC Original Ground Conductivity Map



(b) County-Level Ground Conductivity Map

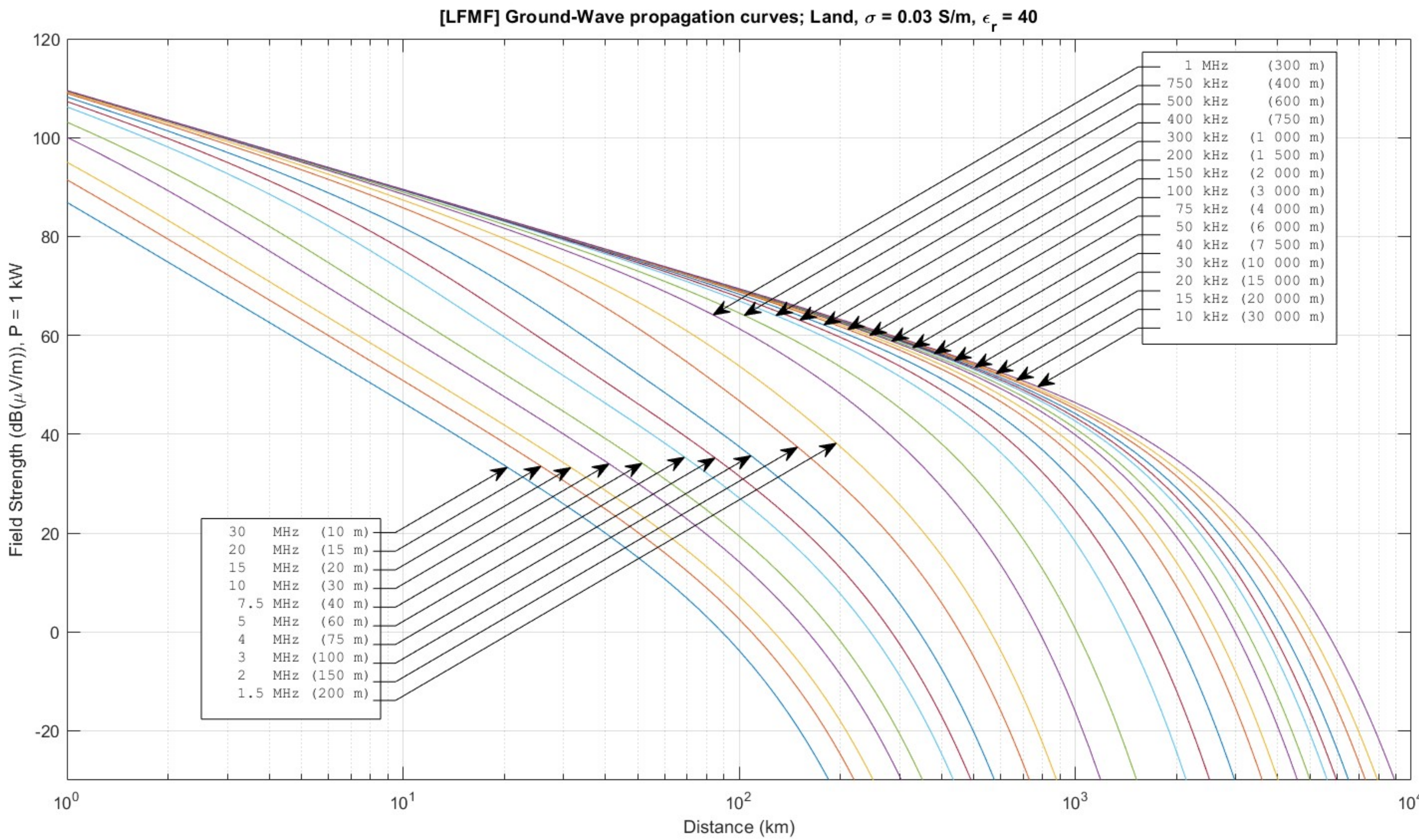
Figure B.8: US Ground Conductivity. The source of the map in panel (a) is the Federal Communication Commission and dates back to 1954. Panel (b) is the same map digitized and aggregated at the county-level.

B.2.1 Validation Exercise

To test that we correctly utilize the LFMF model we replicate one of the propagation curves provided by the ITU. Figure B.9 displays a set of propagation curves computed and provided by the ITU for a given value of ground conductivity and permittivity $\sigma = .03S/m$, $\epsilon = 40$. The graph plots the field strength values on the y-axis for different values of distance values on the x-axis. The curves show the field strength reached by a ground wave emitted by a transmitter vertically polarized with power $p\_tx = 1kW$, height $h\_tx\_m = 15$, receiver height $h\_rx\_m = 10$, for a range of frequency values f_mhz. Refractivity is fixed at mean value observer in continental US $n\_s = 315$.

We replicate one of the propagation curves above for a given value of frequency $f\_mhz = .5mHz$ in Figure B.10. We hold the rest of the values as in the ITU curves. The output of the propagation shows field strength values for the first 200km. Our replication shows a field strength of 89.17 at 10km from the transmitter. Inspecting the correct propagation curve (associated to 500 kHz, the third frequency from the top in the legend on the right side of the graph, the curve is painted in light blue), we find that the corresponding field strength at 10km falls exactly in between 80 and 100 dBμV/m. Similarly, at 100km we compute a field strength equal to 66.88dBμV/m which matches the propagation curve value at the same distance. Our short replication exercise suggests that we correctly implement the LFMF model.

Figure B.9: ITU LFMF Propagation Curves, $\sigma = .03S/m$, $\epsilon = 40$



Notes: The propagation curves are computed and provided by the ITU for several combinations of ground conductivity and permittivity. The graph plots the field strength reached by a ground wave emitted by a transmitter vertically polarized with power $p\_tx = 1kW$, height $h\_tx\_m = 15$, receiver height $h\_rx\_m = 10$, for a range of distance values and frequency values f_mhz. Refractivity is fixed at mean value observer in continental US $n\_s = 315$.

Figure B.10: Replication of ITU LFMF Propagation Curve, $\sigma = .03S/m$, $\epsilon = 40$, $f\_mhz = .5$ mHz

```
ATMOSPHERIC CONSTANTS
    REFRACTIVITY =315.00   (N-UNITS)
    SCALE HEIGHT = 7.350   KM

GROUND CONSTANTS
    RELATIVE PERMITTIVITY =     40.000
    CONDUCTIVITY =3.0000D-02  SIEMENS/METRE

VERTICAL POLARISATION

MINIMUM DISTANCE =    10.000  KILOMETRES
MAXIMUM DISTANCE =   200.000  KILOMETRES
DSTEP            =    10.000  KILOMETRES


FREQUENCY =    0.500 MHZ

TRANSMITTER HEIGHT =    15.0  METRES
RECIEVER HEIGHT    =    10.0  METRES

DISTANCE      FIELD STRENGTH     BASIC TRANSMISSION LOSS
  KM            DB(UV/M)               DB            (F)
    10.00         89.17                46.76
    20.00         82.90                53.03
    30.00         79.13                56.80
    40.00         76.38                59.55
    46.61         74.89                61.05
                                                      (R)
    46.61         74.91                61.02
    50.00         74.22                61.71
    60.00         72.39                63.55
    70.00         70.79                65.15
    80.00         69.37                66.57
    90.00         68.07                67.86
   100.00         66.88                69.05
   110.00         65.78                70.15
   120.00         64.74                71.19
   130.00         63.77                72.16
   140.00         62.84                73.09
   150.00         61.95                73.98
   160.00         61.10                74.83
   170.00         60.29                75.65
   180.00         59.50                76.44
   190.00         58.73                77.20
   200.00         57.99                77.94
```

Notes: We replicate the propagation curves in Figure B.9 for a given value of frequency $f\_mhz = .5$ mHz. We fix the rest of the values as in the ITU curves $p\_tx = 1kW$, $h\_tx\_m = 15$, $h\_rx\_m = 10$, $n\_s = 315$.