Discrimination, Migration, and Economic Outcomes: Evidence from World War I*

Andreas Ferrara
University of Pittsburgh

Price Fishback
University of Arizona

This version: November 30th, 2020

Abstract

Are the costs of discrimination mainly borne by the targeted group or also by society? This paper examines both individual and aggregate costs of ethnic discrimination. Studying Germans in the U.S. during World War I, an event that abruptly downgraded their previously high social standing, we propose a novel measure of local anti-German sentiment based on counties' casualties in the war. We show that Germans fled high casualty rate counties where anti-German sentiment intensified. German movers had worse occupational outcomes after the war but also the discriminating communities paid a substantial cost. Counties with larger outflows of Germans, who tended to be well-trained manufacturing workers, saw worse post-war employment outcomes for Americans and a drop in average annual manufacturing wages of 1-7 percent that lasted until 1940. The price of a few years of intense anti-German sentiment was reflected in worse economic outcomes that lasted for more than a decade.

Keywords: ANTI-GERMAN SENTIMENT; WORLD WAR I; COSTS OF DISCRIMINATION
JEL Classification: N42, J15, R11

Acknowledgments: We thank Sascha O. Becker, Karen Clay, Nathan Eisner, Bishnupriya Gupta, James Fenske, Joung Ha, Taylor Jaworski, Daniel Jones, Keith Meyers, Allison Shertzer, Max Steinhardt, Randy Walsh, and Marianne Wanamaker as well as seminar participants at Free University Berlin, Monash University, University of Southern Denmark, University of Pittsburgh, and University of Warwick for helpful discussions and comments. We also thank Mike Matheis who kindly shared his manufacturing data with us. Corresponding author: a.ferrara@pitt.edu

1 Introduction

Since Becker's (1957) seminal book on the economics of discrimination, a substantial literature has studied the effects of discrimination on individuals of the affected group.[1] While the effects at the individual level are better known, a less explored question is whether discrimination against a certain group generates costs that are borne by this group alone or whether the local communities that chose to discriminate also have to pay a price. Such costs can arise theoretically in labor markets with frictions that are amplified by this type of discrimination (Lang, Manove and Dickens, 2005), or in practice via reduced innovation (Cook, 2014) and deadweight losses paid by workers to avoid workplace interactions with the discriminated group (Heedegaard and Tyran, 2018).[2] A major difficulty in studying this question is to differentiate between taste-based and statistical discrimination (Phelps, 1972), and to find plausibly exogenous variation in discriminatory behavior.

To this end, a recent literature has focused on Germans in the United States during World War I starting with Moser (2012). Being the largest, economically successful, and socially accepted immigrant group (Higham, 1998), the war temporarily but substantially turned this previous acceptance into stark discrimination. Prior work on this unexpected shift in taste-based discrimination has mainly focused on the Germans themselves and their assimilation efforts in particular (e.g. Fouka, 2019, 2020). In this paper, we exploit this natural experiment to causally estimate the costs to both Germans and the local communities that chose to discriminate against them. We show that relocation, as opposed to assimilation, was a frequent but individually costly reaction of Germans to high levels of discrimination. We also show that the discriminating communities had to pay a price. Germans tended to be well-trained manufacturing workers before the war.[3] With these specialized workers leaving high-discrimination counties, the costs of a few years of intense discrimination were reflected in decreased average wages in manufacturing in those local communities which lasted until the 1940s.

[1] This includes discrimination in labor markets (Bertrand, Chugh and Mullainathan, 2004; Charles and Guryan, 2008; Heedegaard and Tyran, 2018), education (Card and Krueger, 1992; Lang and Manove, 2011), health care (Alsan and Wanamaker, 2018; Alsan, Garrick and Graziani, 2019), and others to provide just some examples. For reviews see Lang and Lehmann (2012) and Neumark (2018).

[2] At a national level, other studies have calibrated macroeconomic models to estimate the overall output costs of discrimination. They find that discrimination against women or African Americans reduced output in the long-run (Cavalcanti and Tavares, 2016), e.g. via the misallocation of talent (Hsieh, Hurst, Jones and Klenow, 2019).

[3] According to the 1910 Census, only 15.7% of the labor force worked in manufacturing, a sector which accounted for 42% of national output (Kendrick, 1961). 23.6% of male German-born labor force participants worked in manufacturing, and 37.1% of them worked in operatives and craftsmen occupations. This compares to 26.5% of non-Germans in operatives and craftsmen jobs, and 15.4% in manufacturing.

We first propose a novel measure of localized anti-German sentiment based on war casualties. We digitized and geo-located information on more than 71,000 soldiers who died in the war. With newspapers reporting the daily casualty lists and highlighting losses in the local communities, casualties increased animosity towards Germans in precisely those communities. We formalize this idea in a regression framework by showing a significant positive correlation between our casualty rate and more noisy measures of anti-German sentiment that we collected from digitized newspapers during the war years. Counties with higher World War I casualties had a higher share of newspaper articles calling Germans *enemies* or *huns*,[4] or that reported tarrings and featherings of Germans in 1917-18.[5]

Using both county-level and linked individual Census data in a difference-in-differences setting, we then show that German-born individuals moved away from counties in the top part of the casualty rate distribution while counties in the bottom part of the distribution saw an increase in their German population shares.[6] This relationship between war casualties and migration patterns existed for Germans from before to after the war, but not in other decades such as the period from 1900 to 1910. The relationship also did not exist for other immigrant groups such as Swedes or Italians. Another contributing factor was salience as the outflow of Germans was strongest in the Midwest, where Germans historically had been the largest immigrant group, and the inflow was most pronounced in the South.[7] We are the first to document this substantial internal migration of Germans in response to discrimination as previous work mainly focused on assimilation, or lack thereof, on part of the Germans (e.g. Fouka, 2019, 2020).

The short-run costs to Germans who moved due to this casualty shock were substantial. Linked Census data from the 1910 and 1920 full count Censuses for almost 150,000 German-born men show that movers were willing to relocate by 585 miles (941km) on average to leave discriminating counties.[8] In 1920, they were significantly more likely to live in the South, resist naturalization, or to have experienced an occupational downgrade from before to after

[4]The term *huns* was used as ethnically-loaded and derogatory term for Germans (similar to *kraut* during World War II) based on the *Hun speech* by Emperor Wilhelm II on July 27, 1900.

[5]Both violent and non-violent forms of discrimination against Germans were common during the war years but spiked particularly once the U.S. officially joined the conflict in April 1917 (Lübke, 1974).

[6]It should be noted that we use the terms Germans and German-born interchangeably for brevity and to avoid repetition.

[7]From 1910 to 1920, the share of Germans increased by 0.47 percentage points in the South but declined by -0.21 percentage points outside the South. This pattern is shown in panel (a) of Figure 3.

[8]This is approximately the distance from Pittsburgh in Pennsylvania, to Atlanta, Georgia; or from Milwaukee in Wisconsin, to Nashville, Tennessee. The finding is also robust to quality of the links produced by the record linkage algorithm we used (Abramitzky, Boustan and Eriksson, 2014).

the war.[9] The occupational downgrading was mainly driven by moving from non-agricultural work before the war to jobs in agriculture after the war. The occupational change toward farming may have occurred because the German-born movers mostly went South, where there was less manufacturing activity. Another advantage of becoming a farmer and to live in a more rural setting was that the German-born were able to reduce the day-to-day contact with workers, employers, and customers, and thus the sting of discrimination.

Anti-German discrimination and the resulting outflow of Germans had negative effects on economic growth in discriminating counties in the post-war decades that persisted until the 1940s. This was particularly true for the manufacturing sector, where Germans tended to be concentrated before the war. Using county-level Census data from 1900 to 1940 in a difference-in-differences setting, we show that counties that experienced an outflow of Germans between 1910 and 1920 saw a reduction in the average log annual earnings in manufacturing of between 1 to 7 percent.[10] The results are robust to the inclusion of pre-war county characteristics such as population size, the share of Germans, share of manufacturing employment, the male-to-female ratio, percent urban, as well as county-specific linear or quadratic time trends. The same relationship between German outflows and manufacturing wages did not exist for the outflow of Germans from 1900-10, nor for the outflow of other immigrant groups from 1910-20 such as Swedes or Italians. We also mirror the previous linked-Census data exercise by examining the effect of German outflows on American-born workers who remained in counties with higher exits of Germans between 1910 and 1920. We show that such outflows are associated with lower mobility, increased chances of occupational downgrading, decreased occupational income scores, and no increase in the probability of manufacturing employment for American-born workers despite leaving Germans freeing up such jobs.

As a robustness check we also employ an instrumental variables strategy in which we instrument for the outflow of Germans from 1910 to 1920 with the World War I casualty rate in the respective counties.[11] This gives the outflow effect the interpretation of being the chan-

[9] We refer to occupational downgrading as moving from a high-skilled occupation in 1910 to a semi-skilled or low-skilled occupation in 1920, or from a semi-skilled to a low-skilled occupation from before to after the war.

[10] We also test for the effect on counties that received Germans and find a small positive but insignificant effect. This is mostly explained by the fact that relocating Germans took up agricultural occupations in which they did not necessarily have a comparative advantage if they worked in manufacturing before. Also, the Southern manufacturing sector at the time was much less developed, hampering any positive impacts of new German workers with skills in this sector.

[11] We rule out selection into higher casualty rates based on literacy, wealth, skill, racial and ethnic composition. We also show that casualty rates are as good as randomly assigned in space once conditioned on county fixed effects. The average number of casualties was 29 soldiers in the top quintile and thus was not sufficiently high to disturb the labor market itself but discriminatory pressure amplified via newspaper reporting and hence drove the

nel through which discrimination affected economic outcomes in the local communities. We rule out a direct labor market effect of those casualties by estimating the regressions not only with pre-war county characteristics but also with time-varying controls for total population and male population size. This shuts down the direct labor market effects that these casualties might have had through their impact on labor supply.[12] The results confirm the baseline difference-in-differences results. The negative wage effect was strongest in 1920 and then slowly faded out until wages reached their pre-war level again in 1940. The costs for the few years of intense anti-German sentiment were therefore reflected in a reduction in average wages that lasted for more than a decade.

This paper is the first to quantify the direct effects of the war on local anti-German sentiment and its effect on the relocation decisions of Germans in the U.S., as well as on the economic outcomes of the local communities that chose to discriminate.[13] Previous work has used this sudden but temporary taste-based discrimination shock to study assimilation (Fouka, 2019, 2020) or employment outcomes (Moser, 2012), innovation and patenting (Moser and Voena, 2012; Baten, Bianchi and Moser, 2017), and spending on urban schools (Schmick and Shertzer, 2020). We add to this literature by introducing a novel measure of anti-German sentiment based on war casualties. Using this measure, we provide evidence that the decision to migrate substantially affected not only Germans, but also the counties that discriminated away their German population. Our setting provides a rare opportunity in which the costs of discrimination can be studied for both the affected group as well as for the discriminating local communities.

The negative impact on the discriminating communities is similar to the recent work on forced migration which has mainly focused on state mandated expulsions, one of the most extreme types of discrimination (see Pascali, 2016; Testa, 2018; Becker, Grosfeld, Grosjean, Voigtländer and Zhuravskaya, 2020). The survey article by Becker and Ferrara (2019) highlights that forced migration is well studied with respect to its effects on the migrants and the receiving economies. We provide a new estimate for the effects on both the migrants and the so-called sending communities by showing that discrimination harmed the local economies through the channel of outmigration of a relatively skilled group in response to ethnic animus.

disproportional outflow of Germans from these counties.

[12] It should be noted that the 110,000 fallen soldiers made up only a small part of the overall population of 100 million in 1910. In fact, our results hold with and without such time-varying population controls.

[13] Joe Price, Kerwin Charles, Tanner Eastmond, and Daniel Rees are currently working on a similar project that exploits WWI casualties and the discrimination against Germans to study German long-term outcomes related to migration and intergenerational mobility.

The main take-away is that even though the discriminatory shock lasted for a few years, the economic cost of discrimination to those communities lasted for more than a decade.

More broadly, we also contribute to our knowledge of the intangible benefits of certain groups to society that only become tangible once said group disappears. This includes benefits arising from different factors such as birth place diversity and its positive impact on economic prosperity (Alesina, Harnoss and Rapoport, 2016), the return from specific skills within a group that benefit the broader society via the provision innovation (Cook, 2014), education (Akbulut-Yuksel and Yuksel, 2015) or talent (Hsieh et al., 2019). The associated long-run costs of losing those benefits because of the disappearance of such groups can be substantial (see D'Acunto, Prokopczuk and Weber, 2019; Lee, Peri and Yasenov, 2019), and it can alter how those groups are treated subsequently (Feigenbaum, Mazumder and Smith, 2019). We complement this literature by showing that discrimination against a previously well-integrated group can create significant costs for both the discriminated group and the discriminating societies even when such discrimination is temporary.

2 Historical Background and Related Literature

Germans were among the largest immigrant groups in the United States until the early 20th century. In the 1910 Census, the share of German-born and second generation Germans among the total population in the U.S. was over 10%. Not only were Germans a large share of the population, they also were known for their economic successes. Higham (1998) cites a survey of businessmen from 1908, who ranked immigrant nationalities by traits. They ranked Germans above English immigrants and even attributed to them more positive traits than Americans in some respects. Abramitzky et al. (2014) compute earnings penalties for different nationalities based on occupations and show that Germans had the lowest penalty relative to Americans. The penalty disappeared entirely for those who had stayed in the country for more than 30 years.

German, Swedish, and English immigrants had similar levels of literacy, urbanization, and naturalization to citizenship. They had higher average education than other immigrant groups, and their occupational distribution had earnings that were similar to those with American-born individuals whose parents were also born in America. Their home ownership rates were 4.4 percentage points higher than those of American-born individuals and they owned businesses at roughly the same rate (see Appendix Table A1).[14]

[14]In this particular table we defined Americans as American-born individuals whose parents were also born in

Germans were known for being hard working and economically successful, but also for tending to their language and customs. For example, Germans were 8 to 14 percent less likely to speak English than Swedish and English immigrants. Cities with larger German populations even offered bilingual education (Fouka, 2020). Other examples include the gymnastics (*Turnvereine*) and shooting (*Schützenvereine*) societies, German language newspapers, and German churches (Lübke, 1974). German culture and their local communities left a permanent mark on the landscape of the U.S. Many Germans settled in the Midwest and many towns had German names such as Berlin, Wisconsin, or Bremen, Indiana. Despite their social preferences, Germans and German-Americans were well respected as hard-working, rapidly assimilating, and patriotic members of society (Higham, 1998). For decades they had managed a balance between their old and new home, which is reflected in the saying: "*Germania my mother, Columbia my bride*" (Lübke, 1974, p. 48)

Attitudes changed dramatically with the onset of World War I (1914-18). Germans and German-Americans experienced increased animosity from the beginning of the war, which increased when some German churches and societies tried to raise funds for the German war effort or lobbied for the U.S. to remain neutral during the conflict (Lübke, 1999). However, the peak of Anti-Germanism in the U.S. was reached after the country eventually entered the war in 1917.

A key channel through which anti-German sentiment was spread were newspapers since the radio did not diffuse until the 1920s. Figure 1 and panel (a) of Figure 2 show examples of anti-German slurs in the news. Appendix Figure A1 plots the share of newspaper articles including the words *enemy* or *huns* among articles mentioning *Germans*. In reading a large number of newspapers, we discovered that the word "Hun" stood out as derogatory term. Usage of the word *enemy* in relation to Germans saw a surge after 1914 and then doubled again with the entry of the U.S. into the war. This is also when the share of articles about Germans using the word *huns* spikes.[15]

Figure 1 displays different depictions of Germans in the American press or books. Panel (a) shows the "German Hun" as rapist who is stopped by an American soldier in an advertisement for war bonds, and panel (b) depicts a German soldier as a goose-stepping child murderer. These types of depictions were targeted not at only at German-born citizens but also against

the U.S. to give a more stringent definition of the established local population. The wording was chosen to avoid confusion with American natives.

[15]Data for this graph were taken from Chronicling America, a source which is described in more detail in the later part of this paper.

naturalized German-Americans. Panel (c) displays a spy who, under the cover of citizenship, seeks to sabotage the U.S. war economy. Germans in the U.S. were frequently accused of spying for the Empire, and were under constant surveillance by para-official organizations such as the American Protective League (APL).[16] Germans were forced to buy war bonds, to kiss the U.S. flag, and to denounce the German Emperor (Lübke, 1974).

The level and extent of Anti-Germanism reached into all parts of life. Moser (2012) shows that the share of operas by German composers fell from 43% to less than 7%, the use of Otto or Wilhelm as first names for newborn children declined dramatically, and applicants to the NYSE with a German sounding surname were twice as likely to be rejected during the war years. Sauerkraut consumption fell by 75% from 1914 to 1918, hamburgers were renamed *liberty steaks* (Fouka, 2020), and substantial amounts of school resources were reallocated in response to anti-German hysteria (Schmick and Shertzer, 2020). Aside from the economic and social discrimination, Germans also had to fear physical harm. Robert Prager was lynched on April 5, 1918, in Collinsville, Illinois, and beatings or tarring and feathering were other more common forms of assault (Lübke, 1974). Several men involved in the Prager lynching subsequently faced trial in a court of law. None of them were convicted.

The state not only turned a blind eye but actively benefited from expropriating German property. The Office of the Alien Property Custodian was established in October 1917 and tasked with expropriating German assets. Miller (1922) details the corporations, firms, trademarks, copyrights, and patents seized and sold under the office in relation to the Trading With the Enemy Act of October 6, 1917. The total value of these exceeded more than $ 500 million in 1919, which corresponds to $ 7.5bn in 2018 dollars using the CPI from the BLS inflation calculator.

Moser and Voena (2012) show that expropriation and reselling of German chemical patents to domestic inventors raised domestic patenting by 20% relative to the average patenting activity between 1919 and 1939. Most of the gains came from improvements based on learning-by-doing, which tended to follow after a lag of eight to nine years. German firms responded by patenting more in these same fields after the war (Baten et al., 2017).

A large number of schools prohibited the teaching of the German language (Lleras-Muney and Shertzer, 2015). In areas where this occurred Fouka (2020) shows that the children of German immigrants reacted to this effort to force assimilation by volunteering less for military

[16]The APL was founded in 1917 and at its peak had 250,000 members in over 600 cities who were looking out under cover for enemy activities or possible spies (Higham, 1998).

service during World War II, marrying more within their own group, and giving their children typical German names more often.

A significant number of German immigrants responded to the prospect of discrimination by assimilating in various ways. According to Fouka (2019), many Germans chose to anglicize their names and to petition for naturalization, especially in states with higher incidences of violence against Germans. German parents who had been in the U.S. longer were more likely to give their children English names. Biavaschi, Giulietti and Siddique (2017) provide evidence of general positive payoffs for name Americanization by migrants in the early 20th century. Some Germans sought to prove their loyalty to the U.S. by volunteering for military service. The volunteers later were more likely to marry Americans and to become naturalized citizens, albeit at a lesser rate than other immigrant groups such as Italians or Eastern Europeans (Mazumder, 2018).

Another response to anti-German attitudes was to move to a more receptive location, which is a focus of our paper. Between 1910 and 1920 German-Americans migrated to new locations in the U.S. at a significantly higher rate than any other group. Panel (a) of Figure 3 maps the change in the share of German population (net of total population changes) from 1910 to 1920 using county-level Census data. The largest outflows occurred in the Midwest where Germans were a large and salient immigrant group. With the war ending the Age of Mass Migration (Abramitzky et al., 2014), very few Germans left the U.S. and the evidence we present in this paper shows that they tended to relocate to areas with lower anti-German sentiment.

A major question we address in this paper is the extent to which German migration flows affected economic outcomes not only for the moving Germans, but also for the communities that these Germans fled due to the increased discriminatory pressure. In a recent survey, Becker and Ferrara (2019) provide an overview of the few existing examples in which forced migration led to lasting effects on the economies of the sending locations. Huguenots who moved to Prussia after being expelled from France in 1685 brought knowledge and technology with them that subsequently raised Prussian productivity in the long-run (Hornung, 2014). The expulsion of 3 million Germans from the Czech borderlands after World War II negatively affected economic growth in these areas by eroding property rights and reducing the benefits of agglomeration economies (Testa, 2018). Pascali (2016) shows that Italian municipalities that expelled their Jewish population during the 15th and 16th century have lower incomes today and a less developed banking system.

In this paper we examine the forced movement by Germans in the U.S. that was not driven by state mandated expulsion but by differential increases in discriminatory pressure during the war from the rest of the local populations. Unlike other studies in this area, we can observe the economic outcomes for both the discriminated group as well as the sending communities within the legal, economic, and institutional framework of the same country. This provides us with a unique setting to study the costs of discrimination for those who were discriminated against and those who chose to discriminate due to a short-run shock to anti-group sentiment.

3 WWI Casualties, Discrimination and Patterns of German Migration

3.1 War Casualties and Anti-German Sentiment in the News

To effectively study the causal impact of discrimination on the German response and economic outcomes at the local level, we need a measure of the types of discriminatory attitudes that German immigrants perceived and a driving source of variation in that measure that is exogenous. We propose measures of anti-German sentiment based on text from 1.9 million newspaper articles we obtained from Chronicling America for the war years. In particular, we extracted all articles that mentioned Germans in any context and then calculated the share of articles that mentioned Germans as *enemies*, used the derogatory term *huns*, or reported tarring and feathering of Germans. The term *enemies* could merely reflect war reporting. The use of the term *huns* in print, however, suggests that both the newspaper staff and the readers found it acceptable to use a nasty, anti-German slur in a public manner. Descriptions of actual bodily harm at the local level showed that some segment of the population was attacking German immigrants (Lübke, 1974).

A plausibly exogenous force that would have influenced local anti-German attitudes during World War I was the extent to which members of the local population became casualties in battles against German forces during the War. Recent work has used war deaths as exogenous shocks in different contexts such as to study marriage markets in France after WWI (Abramitzky, Delavande and Vasconcelos, 2011; Boehnke and Gay, 2020) or labor markets and outcomes of blacks in the U.S. after WWII (Ferrara, 2018). We hypothesize that such deaths suffered by the local communities abroad increased anti-German sentiment in those communities. Newspapers were the primary source of information because radios did not diffuse until the 1920s. They were also instrumental in spreading anti-German slurs as well as information on war casualties abroad for which they would highlight fallen soldiers from the local commu-

nities. Figure 2 shows an example of such a daily report on casualties abroad that highlights a death from the local community in panel (a), and an example of anti-German slurs in panel (b). The casualty and newspaper data are described further in the data appendix.

The relationship between anti-German sentiment and WWI casualties can be estimated with the following regression

$$D_c = \alpha_s + \phi \text{WWI Casualty Rate}_c + X_c'\gamma + \nu_c \tag{1}$$

where D_c is one of three measures of anti-German animus: i) the percentage of newspaper articles in county c mentioning Germans as *enemies* out of all articles mentioning Germans in any context in 1917 and 1918, i.e. the war years with U.S. involvement, ii) the percentage mentioning Germans as *huns*, and iii) the percentage reporting tarring and feathering of Germans.[17] The vector α_s contains state fixed effects, and X_c is a vector of pre-war controls measured at the county-level in 1910 and includes the percentage of German-born individuals, population size, male-to-female ratio, the percentage of employment in manufacturing, and the percent urban. Also included are the draft rate and the total number of articles published in the different counties which are indexed by c. The denominator of the casualty and draft rate is the number of voting-aged males as proxy for the number of service eligible men.[18] The distributions of the casualty and draft rate are plotted in Appendix Figure A2.

Around 17% of all counties had at least one newspaper outlet in Chronicling America that published at least one article during 1917 and 1918. For counties that did not have a newspaper outlet or none that was published during the war years and included in Chronicling America, we assigned the measures D_c from the nearest county with a publishing newspaper and weight the corresponding observations by the inverse distance to the assigned newspaper county.

The results from this correlation exercise are reported in Table 1. A one percentage point increase in the WWI casualty rate is associated with a 4.75 percentage points increase in the share of articles mentioning Germans as *enemies*, a 2.37 percentage points increase in the share of articles mentioning them as *huns*, and a 0.67 percentage point increase in reported tarrings and featherings of Germans. Relative to the outcome averages, these effects correspond to a

[17] Note that these are not necessarily tarrings and featherings that also occurred in county c but are typically either in the county or in a nearby county.

[18] The county-level aggregate data from the 1910 Census do not provide more finely grained age groups. Our results are robust to using the total population aged 10-20 as denominator, however, since this also includes women it is a less accurate measure of potential soldiers in a county.

26%, 49%, and 38% increase, respectively. War casualties therefore appear to be a significant shifter in anti-German sentiment at the local level.

In what follows, we will rely on the WWI casualty rate as the main independent variable associated with anti-German animus. While newspaper based measures of anti-German sentiment are more in line with capturing anti-German sentiment per se, the number of counties with recorded articles in 1917 and 1918 in Chronicling America is only 17 percent. We also do not know the geographic range and volume of circulation for each newspaper. Thus, we know that the county had access to newspaper content at the time, but we do not know how broadly its information was disseminated among the county's population and to the population in surrounding counties. Given the strong relationship between casualties and anti-German sentiment shown in Table 1 and the availability of casualty rates for every county, it makes sense to use the casualty rate as measure of the shift in anti-German sentiment.

3.2 War Casualties and German Migration Patterns

3.2.1 Data

With anti-German sentiment peaking during 1917/18, Germans had two options. First, they could attempt to assimilate and demonstrate their loyalty to America as argued by Fouka (2019). The alternative was relocation. The war severely restricted cross-Atlantic migration and also made a return to the German Empire unattractive, therefore migrating Germans were most likely to relocate to other counties in the U.S. Panel (a) of Figure 3 shows the change in the German-born population from 1910 to 1920 across U.S. counties.[19] The German share of the population fell the most in the Midwest and rose in the South and to a lesser extent in New England.

Discrimination against Germans was particularly pronounced in the Midwest: "The midwestern state councils, in particular, got a notorious reputation for patriotic vigilantism [...] Most of this coercion was directed at the German-American communities" (Breen, 1984, p. 79). Hegi (2005) describes the case of Fred Tenekheig in Audubon, Iowa, who was dragged across the public square with a rope around his neck by parents of American soldiers, and then forced to leave the town with his family a few days later. He also discusses the much lower degree of violence against Germans in Texas. Potential reasons for this are the rural character

[19]General time and population growth effects have been removed from the mapped values to avoid underlying confounding trends. This ensures that it is the numerator, the share of German-born individuals, and not the change in total population in the denominator that is driving these patterns.

of the South at the time as well as relatively little exposure to German immigrants in general, making it harder to identify Germans by their accents or names.

A striking geographic correlation between casualty rates and changes in the German population share can be seen in the two maps in Figure 3. The strong relationship is not being driven by strong correlations with other observables. As a correlational exercise, we estimate a LASSO selection model to determine which observables from the 1910 Census were most likely to have a statistically significant relationships with the casualty rate and the draft rates during World War I.[20] The correlates with the strongest relationships as ranked by the respective coefficients' t-statistic from highest to lowest are shown in Appendix Figure A3.[21]

As expected, the casualty rate has the strongest statistically significant relationship with the draft rate because counties that sent a higher share of their male population to fight had a higher share of their men at risk. The only other statistically significant relationships were with the population in cities with more than 25,000 inhabitants and with less densely populated, more rural areas. The latter also has the most statistically significant relationship with the draft rate. As before, we therefore continue to condition on pre-war urbanization rates in 1910. It is also important to note that per capita variables such as the number of illiterates, manufacturing labor, farm or house owners, blacks, or foreigners by nationality were not selected by the LASSO regression as significant predictors of the casualty rate. Only the number of manufacturing firms per capita is marginally significant but positively related to the draft and negatively to casualty rates. This implies that World War I deaths were unlikely to be selected along wealth, skill, or ethnic dimensions, which otherwise would potentially raise issues of selection bias. Note that for consistency we continue to control for the same variables as in section 3.1,[22] and do not rely on the LASSO to select the controls.[23]

Another potential confounding factor could be a spatial correlation between changes in

[20] All variables were standardized by population size. The LASSO regression then solves the problem arg $\min_\beta \sum_{i=1}^{N}(y_i - X'\beta)^2 + \lambda_{j=1}^{p} \mid \beta_j \mid$ to select the most significant regressors while shrinking unimportant ones towards zero via the regularized penalization terms that is added to the standard least squares minimization problem.

[21] The 1910 Census has almost 200 variables from which we selected. Figure A3 therefore only plots a certain number of the most significant coefficients. For the casualty rate only the top three most significant predictors are significant at at least the 5% level whereas for the draft rate that number is five coefficients.

[22] These are the 1910 levels of the percentage of German-born individuals, population size, male-to-female ratio, the percentage of employment in manufacturing, and the percent urban, all of which are interacted with a post-treatment dummy, as well as the WWI draft rate.

[23] The LASSO might omit variables that are potentially important but highly correlated with other controls and hence their coefficients would be shrunk to zero. Also, selecting controls based only on correlations with the treatment may lead to biased estimates since it ignores the potential correlations with the outcome (Belloni, Chernozhukov and Hansen, 2014).

the German-born population share from 1910 to 1920 and the casualty rate. Tests for spatial clustering in Appendix Table A2 show that spatial autocorrelation almost disappears entirely once coarse geographic controls, such as state fixed effects, and pre-war county controls are included in the estimations.[24]

The strong raw negative correlation between the WWI casualty rates and the change in the German-born population share between 1910 and 1920 can also be seen in the binned scatter plot in panel (a) of Figure 4. The figure also shows that counties in the bottom of that casualty rate distribution experienced an increase in the share of German-born population whereas those in the top of the distribution saw a decrease. This differential treatment response motivates why we consider a difference-in-differences regression using quintiles of the casualty rate distribution in the next section.

To provide a placebo test, panel (b) of the same figure shows that there was no negative relationship between the World War I casualty rate and the change in the German population share between 1900 and 1910. There were a large number of Swedish-born and Italian-born immigrants in the U.S. and their home countries did not fight against the U.S. in World War I. Panels (c) and (d) show that there was no negative relationship between the WWI casualty rate and the changes in their shares of the population across counties between 1910 and 1920. Thus, the negative relationship between casualty rates and the German share occurs only during the World War I decade and that relationship is not present for other immigrant groups who were not fighting against the U.S.

3.2.2 Difference-in-Differences Model

To formalize the analysis, we combine the casualty rate measure with county-level data from the U.S. decennial Census from 1900 to 1940 and estimate the following regression,[25]

$$\begin{aligned} \%\text{ German population}_{ct} = &\sum_{q=1,q\neq 3}^{5} \tau_q Q_q(\text{WWI Casualty rate})_c \times \text{Post-WWI}_t \\ &+ X_{ct}'\gamma + \alpha_c + \lambda_t + \epsilon_{ct} \end{aligned} \tag{2}$$

where the outcome is the share of the German-born population of a county's total population in county c in Census year t. To capture the differential treatment response across the casualty

[24]The routine for estimating the Getis-Ord statistic for local spatial autocorrelation was provided by Kondo (2016).

[25]Summary statistics for the county level data are reported in Table 2.

rate distribution highlighted in Figure 4, we split the casualty rate into five bins and let Q_q denote the τ^{th} quintile of the World War I casualty rate distribution. Each quintile indicator is interacted with a post-war indicator that equals one after 1910 and is zero otherwise. Quintile three is omitted and acts as the baseline for comparison. If local casualties led to anti-German sentiment, and if Germans sought to evade such animosity, then those counties with lower casualty rates should receive Germans and those with higher losses should lose them, i.e. τ_q should be decreasing in q.

We control for time-invariant county characteristics with county fixed effects α_c and aggregate shocks common to all counties are absorbed by time fixed effects λ_t. The vector X'_{ct} includes the World War I draft rate, as well as pre-war characteristics such as population size, the share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate. All of these are measured in 1910 and are interacted with time fixed effects, hence the t subscript. All unexplained variation remains in the error term ϵ_{ct}. To account for heteroscedasticity and autocorrelation we cluster standard errors at the county level.

The identifying assumptions are that the share of the German-born population evolved in a parallel way across low- and high-casualty rate counties over time prior to the war and that there are no unobserved time-varying factors that are correlated with both the WWI casualties and the post-war measures of the outcome. To relax this assumption, we also include linear county-specific time trends $\alpha_c t$, and quadratic time trends $\alpha_c t^2$, in other specifications to probe for robustness of our findings with respect to underlying differential time trends in high and low casualty rate counties that might be driving the results.

3.2.3 Results

The results from estimating eq. (2) are reported in Table 3. The τ_q coefficients for the casualty-rate bins are negative and statistically significant for counties with the highest casualty rates. As the casualty rates fall, the coefficients eventually turn positive and become statistically significant again in areas with the lowest casualty rates. Counties in the top quintile experienced an outflow that ranges from -0.310 percentage points in specification 2 with the controls in equation 2 to -0.195 percentage points when we control for county-specific time trends. Relative to the average share of Germans of 1.815% in 1910, this is an average reduction of 11 to 17% in the German-born county population.

The percentage point reduction in the German-born share in counties in the top casualty-

rate quintile is nearly matched by the increase in the German-born share in counties in the lowest quintile with a range of increases of 0.154 to 0.377 percentage points. Thus, Germans appear to have relocated within the U.S. rather than leaving the country entirely. This finding is consistent with robustness checks by Fouka (2019), who finds that her results are also not driven by Germans exiting the U.S.

We also estimated an alternative event-study specification to test for the presents of pre-trends before the war decade and to measure the timing of the post-war changes in the German-born shares across counties. Figure 5 plots the coefficients from regressing eq. (2) where Post-WWI$_t$ was replaced with year dummies, leaving out the indicator for the year 1910 which acts as baseline.

Error bars show the 95% confidence intervals. None of the coefficients prior to 1910 are statistically different from zero, consistent with the absence of any uncontrolled pre-trends. After the war, the share of Germans in counties with the highest casualty rates fell by -0.25 percentage points in 1920, -0.3 percentage points by 1930 and -0.44 percentage points in 1940. All of these coefficients are statistically significant. The shares also fell for counties in the second highest quintile, but the coefficients were smaller and not statistically significant.

Germans were most likely to shift into counties in the lowest casualty-rate quintile. The German-born share rose by 0.25 percentage points in 1920, 0.3 percentage points by 1930, and 0.4 percentage points by 1940. The coefficients for the second lowest casualty-rate quintiles also were positive, but were not statistically significant. The increasing effects of the casualty rates over time in the highest and lowest casualty-rate counties suggest that relocation was a permanent decision, setting sending and receiving counties on differential population growth trajectories for their German-born population.

4 Outcomes of Migrating Germans

How did the German-born fare when they moved away from counties with the highest discriminatory pressure? To answer this question, we link individual data from the full count Censuses of 1910 and 1920. We match German-born men aged between 15 and 60 in 1910 to individuals in 1920 following the procedure developed by Abramitzky et al. (2014) based on first and last name, place of birth, and a two year interval around a person's year of birth.[26] We obtained a matched sample of almost 150,000 individuals who were either uniquely matched, or where the

[26] We are linking men only due to the well known issue that women tend to change their surname upon marriage.

closest match could be determined by the birth year difference and the parents' place of birth. The matching algorithm is described in more detail in the data appendix. Summary statistics for the individual linked data are presented in Appendix Table A3.

It should be noted that the resulting matched sample is for the sub-population of German-born individuals who chose not to anglicize their names.[27] Name Americanization was one way through which Germans and German-Americans sought to evade discrimination, in which case our matching algorithm would not find those individuals in the 1920 Census. However, it is not clear how successful this strategy was. Especially for German-born individuals the German accent was more likely to reveal their nationality and we show that language was a key determinant of relocation below. The group of Germans who did not change their names is the group of most interest here because they were the ones who most likely responded to discrimination by migrating rather than by assimilating.

We have shown that the largest discrimination likely occurred in counties that were in the top quintile of the casualty rate distribution and that those were also the areas that experienced the largest reduction in the German-born share of the population. To determine how German-born men who migrated to avoid the largest discriminatory pressure fared relative to other German-born men, we regress the outcomes of German-born males on an indicator for whether they lived in a county in 1910 that would ended up in the top quintile of the casualty rate distribution interacted with an indicator for the year 1920,

$$y_{ict} = \beta \left[Q_5 \left(\text{Casualty Rate} \right)_{c,1910} \times \text{Post-WWI}_t \right] + \alpha_{c,1910} + \lambda_t + X'_{ict}\Gamma + \nu_{ict} \tag{3}$$

where y_{ict} is the outcome for individual i living in county c in Census year t. The outcomes include indicators for moving to a new county between 1910 and 1920, the distance moved between 1910 and 1920 in miles,[28] living in a Southern state in 1920, not being naturalized by 1920, having experienced an occupational downgrade between the two Census years, and for working as farmer in agriculture in 1920 after having a non-farm occupation in 1910. The occupational downgrade equals one if a person had a high-skilled job in 1910 and moved to a

[27] For this section only, we will use German-born and German interchangeably for brevity and to avoid repetition. The correct term is German-born in this context.

[28] Distance is measured as minimum straight line distance, i.e. as the crow flies, from the centroid of the county of residence in 1910 to the centroid of the county of residence of an individual in 1920. Naturally, the distance for non-movers is zero.

semi-skilled or low-skilled job in 1920, or if they had a semi-skilled job in 1910 and ended up in a low-skilled job in 1920.

The top quintile casualty rate indicator is captured by the $Q_5(\cdot)$ function and is interacted with the 1920 indicator, Post-WWI$_t$. The parameter of interest is β. To stick as closely as possible to the empirical design of the previous section, we control for county fixed effects in 1910, i.e. the county for which the casualty treatment is assigned,[29] and a 1920 indicator λ_t.

The set of controls include the county level draft rate as well as a large set of individual baseline characteristics measured in 1910 which are interacted with the 1920 indicator. The individual controls in 1910 contain binary variables for urban status, birth cohort, place of birth, literacy, occupational skill group,[30] farm status, employment status, marital status, years since entry to the U.S. in bins of 0-5, 6-10, 11-15, and 16-20 years, school attendance, labor force participation, and count measures for family size and the number of weeks spent in unemployment in 1909. Standard errors are clustered at the county level.

The β coefficient estimate in column (1) of Table 4 shows that a German-born man living in a county in 1910 that later had a WWI casualty rate in the top quintile was 63.4 percentage points more likely to have left the county by 1920. This effect is 1.3 times the baseline probability of moving of 49% in this sample.[31] Remember, however, that the sample is for Germans who did not assimilate by changing their names, so the impact may have been smaller if we looked at the German-born population that included name-changers.

Columns (2)-(6) focus on the sub-sample of movers to gauge whether relocation was economically beneficial or detrimental to those who left high-casualty rate counties. Conditional on moving, column (2) shows that Germans who left their county of residence due to the WWI casualty shock were willing to relocate by 585 miles (941km) on average. This is approximately the distance from Pittsburgh (Pennsylvania) to Atlanta (Georgia), or from Milwaukee (Wisconsin) to Nashville (Tennessee). Discrimination against Germans was widespread in the Midwest where they had traditionally been a salient minority group (Breen, 1984). Violence against Germans was significantly lower in the South (Hegi, 2005). This potentially motivates the significant relocation effort made by these individuals.

[29]The correspondence with the empirical strategy in section 3.2.2 could be achieved by collapsing the individual data at the county-year level. The benefit of the individual data are the additional observables and the ability to trace individuals to their 1920 county of residence.

[30]We use the 1950 occupation definitions of the U.S. Census Bureau and divide occupations into ten groups which are professional and technical, farmers, managers as well as officials and proprietors, clerical and kindred workers, sales workers, craftsmen, operatives, service workers, farm laborers, and laborers.

[31]See the 1910 summary statistics of the linked sample in Appendix Table A3.

Column (3) provides evidence that German- born individuals, who relocated out of a high casualty rate county from 1910 to 1920, were 3.4 percentage points more likely to live in the South in 1920. This is consistent with the average distance traveled estimated in column (2). A sign that the movers were less interested in assimilation in response to discrimination is that the movers from counties with the highest casualty rates were also 4.3 percentage points more likely to not have naturalized by 1920 in column (4). In terms of their occupational outcomes, the German-born who left counties with the highest casualty rates were 2.6 percentage points more likely to have dropped to a lower skill level in 1920 as shown in column (5). Finally, they were 20 percentage points more likely to be employed in agriculture in 1920 after having worked in a non-agricultural industry in 1910. Farming in rural and more isolated areas might have been a way to avoid discrimination though it should be noted that the South at this time had little manufacturing employment that the German-born movers could have entered.

In the next section, we examine the impact of German-born migration on manufacturing in the counties they left. Appendix Table A4 shows the results of estimating eq. (3) for individuals in three different industry groups which are measured in 1910.[32] The outcomes are the mover and distance moved variables. The group with the highest relocation probability were manufacturing workers in counties in the top WWI casualty rate quintile. They were 74.1 percentage points more likely to move as result of this casualty shock. Hence manufacturing is not only an industry in which wages are well measured before 1940,[33] but also where one would expect a significant effect of Germans migrating away.

4.1 Robustness to Linkage Errors

It appears that relocation was not beneficial for German-born individuals in economic terms. These results provide evidence for the costs that individuals are prepared to pay in order to avoid the heavy discrimination faced by Germans during the war years. As the previous section suggests, there is no evidence for return migration. One potential issue with the linked individual data is that record linkage across Census years can be prone to mismatches. In this case, individuals would be marked as *movers* when in fact they are not but simply were linked to the wrong individual in 1920. However, if such mismatches occurred, it would imply a misclassification in our binary outcomes as well as our binary treatment. This would lead to two

[32]The main industries employing Germans were agriculture, manufacturing, and retail/sales, which accounted for 57% of all employment among Germans in 1910.

[33]The Census only asked for wages from 1940 onwards but aggregate wages in manufacturing are available at the county level.

attenuation biases that come from these two misclassifications (see Meyer and Mittag, 2017).

A related issue is whether Germans purposefully misreported their place of birth in the Census to evade discrimination. This is not necessarily done specifically for the Census but general behavior that is similar to the Americanization of German names.[34] While we cannot rule this possibility out completely, the map of changes in the German population share in panel (a) of Figure 3 suggests otherwise. If a substantial majority of Germans did not leave the country but instead lied about their place of birth, we should not observe an increase in the share of German-born individuals in the South, for instance, since there was no return or incentive for lying and falsely claiming to be German-born.

To probe the sensitivity of our results with respect to potential mismatches, Appendix Table A5 reports the results from eq. (3) using high-quality matches only. This means we use individuals who appeared in the 1910 Census and had a single unique match in the 1920 Census. The results are similar to the baseline estimates reported in Table 4 both in terms of magnitude and statistical significance. When using the unique-match sample, the top-quintile casualty coefficient is larger for the relocation and farm employment outcomes but smaller for the probability of living in the South, not having naturalized by 1920, and the occupational downgrade. None of these differences between the results in Table (3) and Appendix Table A5 are statistically significantly different from each other.

4.2 Determinants of the Relocation Decision

To better understand which Germans migrated and what their characteristics were, we regressed the mover indicator on observables measured in 1910. The coefficients are plotted in Appendix Figure A4 and show that Germans were particularly more likely to move if they were manufacturing workers, active in the labor force in 1910, or had been in the country for a shorter period of time relative to those who had been in the U.S. for 21 years or more. Those who were literate, owned their home, were married, or had larger families were less likely to move.

While insightful, those correlations only relate to the unconditional probability of relocation. To relate the cross-county migration decision to the WWI casualty shock, we repeat the regressions from column (1) in Table 4 and further interact the indicator of having lived in a county in the top casualty rate quintile with the 1920 year dummy, as well as with other

[34] For instance, Müller would become Miller, Kruse would be changed to Cruise, etc.

observables characteristics in the following regression:

$$\begin{aligned} Pr(\text{Mover} = 1)_{ict} = \beta &\left[Q_5 \left(\text{Casualty Rate} \right)_{c,1910} \times \text{Post-WWI}_t \right] \\ &+ \delta_k \left[Q_5 \left(\text{Casualty Rate} \right)_{c,1910} \times \text{Post-WWI}_t \times G_k \right] \\ &+ \alpha_{c,1910} + \lambda_t + X'_{ict}\Gamma + \nu_{ict} \end{aligned} \tag{4}$$

where G_k includes the k interaction variables which are indicators for whether an individual reports German as their first language,[35] their first name being Wilhelm, i.e. the name of the German emperor, an indicator for having one of the 30 most common German surnames,[36] and indicators for years in the U.S. binned into 5 intervals.[37]

The results from this exercise are reported in Table 5. Reporting German as mother tongue adds around 7.5 percentage points to the probability of relocation in response to the top-quintile casualty treatment across all columns. Without additional information we cannot say whether this is because those individuals generally had a stronger accent or if this was a manifestation of their German identity. For instance, previous work has shown that speaking a language with a foreign accent has a negative effect on wages (Grogger, 2011).

Also names appear to have played an important role in the relocation decision of German-born men who lived in counties with the highest casualty rates. Having the emperor's first name of Wilhelm increased the probability of moving by 17 to 20 percentage points, while having a common and thus easily identifiable German last name increased this probability by 28 percentage points. The importance of names in the labor market has been previously studied by Biavaschi et al. (2017). They showed that there are substantial returns to name Americanization in the early 20th century and that individuals who chose to preserve their name identity were much less likely to experience occupational upgrading.

Finally, column (4) shows that those who had lived in the U.S. fewer years were more likely to move when they were in counties in 1910 from the highest casualty quintile. Relative to those who had lived there for more than 20 years, the relocation probability was 18 to 22.6

[35] Which language to report as the mother tongue is a choice in the Census. In fact, 46% of the individuals in our sample report English as their first language even though they were born in Germany. This is not accounted for by Germans moving to the U.S. at a very early age since age of arrival is 17 for those who report English as their first language and 18.5 for those who report German as their mother tongue.

[36] The 30 most common German names in the 1910 Census are: Schmidt, Meyer, Schultz, Wagner, Weber, Hoffman, Schneider, Becker, Schroeder, Mueller, Wolf, Peters, Bauer, Fischer, Koch, Klein, Zimmerman, Krueger, Keller, Beck, Kramer, Mayer, Krause, Schwartz, Hahn, Schmitt, Hartman, Lange, Schaefer, Kaiser.

[37] These bins are 0-5, 6-10, 11-15, and 16-20 years, with the comparison being 21 years or more.

percentage points higher for men living in the U.S. less than 11 years and around 5.5 percentage points higher for men living in the U.S. between 11 and 20 years.

5 Economic Effects on the Sending Communities

When the German-born left communities where social and economic discrimination increased the most, the remaining members in the community may also have experienced losses. Post-war manufacturing in the "sending" counties that saw more German out-migration might have been less successful because the German-born were 1.5 times more likely to be in manufacturing and tended to be in more skilled jobs in manufacturing than other ethnic groups. In the 1910 Census, 24.3% of German-born men aged 15 to 60 were employed in manufacturing compared to 15.9% of men of the same age in the rest of the population. Within manufacturing, 68.5% of German-born men were in operatives or craftsmen positions whereas 56.6% of non-Germans were in such occupations. Further, Germans in manufacturing were more likely to move between 1910 and 1920. The loss of a significant group of workers, particularly skilled workers, would have damaged productivity immediately and required costly and lengthy training of the new workers that eventually replaced them.

Although we would like to study a broad range of aspects of the local economy, we restrict the analysis to manufacturing because nation-wide county-level information on the same measures before and after World War I are only available for manufacturing. We focus on wages for two reasons. First is the previously mentioned concentration of Germans in this sector and the potential productivity and thus wage effects associated with losing these workers. Second, wages are our best measure of social cost given that these affect a large number of workers. If these workers are also the ones who discriminated, then the wage drop from the outflow of discriminated Germans can be interpreted as the aggregate cost of discrimination in those local economies.[38]

Theoretically, the impact of German-American outmigration on manufacturing wages is determined by whether German-Americans were substitutes or complements in production for the remaining workers in industry, which had already developed high levels of capital-skill complementarity in the early 1900s (Goldin and Katz, 1998). If substitutes, the outflow might cause increases in earnings for the remaining workers. If complements, however, the outflow

[38] In fact, there is evidence that both workers and employers were discriminating against Germans (Moser, 2012; Fouka, 2020).

would lead to a reduction in earnings for the remaining workers as they became less productive. German-Americans were more heavily employed in skilled positions during this period relative to other workers, and thus would tend to be complements to the remaining workers. The Committee on Immigration (1911b) reported that Germans in the late 19th century were "the leading race of old immigration" (Part 23, pp. 17-73) in a variety of industries, including iron and steel, meat packing, clothing, leather work, oil refining, and wool textiles, and their sons often continued in the same skilled work. In many areas they started and developed the brewing industry.[39] Given the complementarity of skilled and unskilled workers, the outflow of German-American skilled labor likely led to a reduction in earnings for many of the remaining workers in the places they left.[40]

On the extensive margin, Germans also had high rates of business ownership (see Appendix Figure A1), mostly based on small independent enterprises which disappeared due to the "ideological trauma associated with the First World War" (Zunz, 2000, p. 357).[41] Hence the move of German entrepreneurs would lead to a drop in demand for labor which would have a negative impact on wages. We provide evidence for the negative effect of German outflows on a wider array of manufacturing outcomes - firm size, number of firms, number of workers, and output per firm - in Appendix Table A6. The main analysis focuses on wages due to their consistent measurement and availability over the sample period.[42]

To estimate the impact of the German outflow on manufacturing earnings, we estimate the following equation

$$\ln \text{wages}_{ct} = \beta \left[\text{WWI German Outflow}_c \times \text{Post-WWI}_t\right] + \alpha_c + \lambda_t + X'_{ct} + \epsilon_{ct} \qquad (5)$$

where $\ln \text{wages}_{ct}$ is the natural log of average annual earnings for manufacturing workers in county c and decade t. The WWI German Outflow$_c$ measure is designed to isolate the impact of the outflow of Germans from a county c from before to after the war; therefore, it is equal to the absolute value of the change in the German-born percentage of the population between 1910

[39] In iron and steel manufacturing, for example, the Committee on Immigration (1911a) notes that for "all skilled work and work with machinery the English speaking races and Germans are generally preferred. Practically none of the other races are found in the skilled occupations." (Part 2, p. 69).

[40] For additional historical evidence on the complementarities between immigrants and natives in the early 20th century see Tabellini (2020).

[41] Germans also owned larger companies, most notably those in the brewing industry, such as Anheuser, Coors, or Yuengling.

[42] We rule out that the drop in wages is mechanically driven by the relocation of high-earning Germans in the next section.

and 1920 when the change was negative. It is equal to zero when the German-born percentage stayed the same or increased between 1910 and 1920. As result, the minimum of the measure is zero and the measure is more positive in counties where the German-born share fell more between 1910 and 1920.

All other controls in X_c are the same as in eq. (2). We also include time-varying measures of the total population size and the male population in the county to shut down any direct labor market effects that might be the direct result of suffering war casualties in the respective counties. This is especially important for the next section where we instrument the outflow of Germans from 1910 to 1920 with being in the top quintile of the WWI casualty rate. Standard errors are again clustered at the county level to account for heteroscedasticity and autocorrelation in the residuals.

The results in panel a of Table 6 show that a one percentage point increase in the outflow of German-born individuals during the war decade was associated with a 1.1 percent decline in average annual manufacturing earnings in the decades after the war as shown in column (1). This base specification only includes county and Census year fixed effects. The effect increases to a decline in average annual manufacturing earnings of 1.7 percent when including the pre-war and time-varying population controls in column (2). Once county-specific time trends are included in column (3), the decline is 7.2 percent for each one percentage point increase in the German outflow. The change in magnitude is consistent with the hypothesis that the areas that Germans left were on a higher growth trajectory in the long run than other areas and hence we would underestimate the effects of German outflows on wages. Once we control for the difference in long-term growth rates, the effect of losing German skilled workers is larger. Including the more flexible quadratic county-specific time trends in column (4) leaves the results virtually unchanged compared to the previous column. Our results are in line with findings from the forced migration literature which focuses on state-mandated expulsions and generally finds negative effects on the local economies from which the affected group was removed.[43]

To provide a different perspective on the magnitudes of these effects, consider that the average number of German-born leaving high-casualty rate counties was 215 individuals. The

[43] Testa (2018) shows that the expulsion of Germans in Western Czechoslovakia after World War II led to lasting reductions in agglomeration economies, population density, and educational attainment for the remaining population. The banking sector developed less in Italian cities that expelled their Jewish populations between 1400 and 1600 (Pascali, 2016), and the deportation of Jews in Nazi Germany also led to a communal loss in human capital that reduced education outcomes for German children (Akbulut-Yuksel and Yuksel, 2015).

average number of manufacturing workers per county was 1,300 in 1910. Even after adjusting for women and children in the count of Germans, the outflow of German-born workers accounts for 93% of the reduction in manufacturing workers in counties that experienced an outflow of Germans from 1910 to 1920.[44]

As a placebo exercise, we also estimate the model and use similar measures of the outflow of the Swedish-born and the Italian-born in Appendix Table A8. Neither of these outflows are associated with statistically significant reductions in average earnings. This is true when they are considered on their own as well as when all three outflow measures for the German-, Swedish-, and Italian-born are included in a single regression. The negative effects on the manufacturing sector as result of an outflow of workers therefore appear to be a German-specific result.

Another natural question to ask is whether counties that received Germans saw a positive impact on their economic outcomes. We repeat the estimation of (5) by replacing the outflow variable with the inflow of Germans. Appendix Table A7 reports the results from this exercise. While all estimated coefficients are positive, they are smaller and more imprecisely estimated than those in our main specification using outflows. There are two reasons why one would not expect a one-to-one transfer of gains and losses from having German manufacturing workers around. First, most Germans moved from the Midwest to the South but the manufacturing sector of the South was much less developed at the time. Second, the previous section showed that relocating Germans saw an increased probability of not working in manufacturing but in agriculture after the move, hence their skills were not always utilized in the sector in which they had a comparative advantage.

5.1 Instrumental Variables Regressions

The relationship between manufacturing earnings and the German outflow that we measure might be endogenous if Germans, in particular, migrated out due to conditions that were correlated with unmeasured factors that contributed to a decline in wages. We have explored the channel behind this finding using an instrumental variable, the indicator that the county was in the top quintile of the WWI casualty rate distribution. The results in Table 2 suggest that the indicator will be a strong instrument. The key question is whether the instrument is uncorre-

[44]This refers to column (3) of Appendix Table A6 which predicts a 6.2% reduction in manufacturing labor for a 1 percentage point outflow of Germans. Adjusting the average of 215 leaving Germans for women and children and men in non-manufacturing jobs leaves 75 German manufacturing workers, who constitute 5.77% of the 1910 manufacturing workforce.

lated with the error in the final manufacturing earnings equation. A potential violation of the exclusion restriction could occur if the high casualty rate led to a reduction in a county's labor supply or it shifted the skill composition toward lower skilled workers. We have taken multiple steps to check and eliminate these possibilities.

First, we have included controls for the potential labor supply in the final stage equation in the form of the time-varying population and male population, as opposed to our previous approach of conditioning on pre-war characteristics and interacting these with a post-war indicator. These control variables block the labor supply channel and manufacturing demand channels through which casualties would have influenced manufacturing.[45]

Second, while the casualty shock significantly affected anti-German sentiment through the propagation via newspapers, as argued earlier, its magnitude made it unlikely that it had a direct effect on the labor market. The average number of fallen soldiers in counties in the highest casualty rate quintile was 29 men, followed by 21, 19, 14, and 7 in the lower quintiles. In 1910 only 15 percent of the non-German-born workforce were in manufacturing. Thus, only 4, 3, 3, 2, and 1 of these fallen soldiers, respectively, were likely to have been in manufacturing before the war. Manufacturing employed an average of 1,300 men per county in 1910, so it seems highly unlikely that our casualty rate measure would have had more than a negligible impact on the skill distribution in manufacturing or on the wage-setting mechanisms of the industry itself.

Third, we estimated LASSO regressions of the county casualty rates on all observable county characteristics that were available in the 1910 Census standardized by population size. These including the WWI draft rate and per capita information on illiterate population, manufacturing earnings, manufacturing firms, manufacturing output, manufacturing workers, population density, different measures of home and farm owners, among others, totaling more than a hundred variables the model selected from. The LASSO results in Appendix Figure A3 show that the casualty rate was statistically significantly correlated with only the draft rate, the population in cities over 25,000, and inverse population density. Irrespective of the LASSO, we already control for such factors in all our regression analyses, which include the pre-war population size and urbanization rate, as well as the draft rate as controls, and our results are not explained away by such covariates.[46]

[45] Appendix Table A9 reports results from estimations of the OLS and IV regressions without the time-varying controls for total population and main population as robustness check. Results remain unchanged.

[46] Panel b of the same figure also shows the predictors of the draft rate for completeness.

Fourth, we re-estimated the OLS regression in (5) while including the casualty rate interacted with the post-WWI indicator as a control. The results in Appendix Table A10 show that the coefficients and standard errors of the German outflow variable does not change, and that the casualty rate coefficient is not statistically significant in any specification. All of these facts suggest that the top quintile of the casualty rate will not be correlated with the error term in the second stage.

The first stage results are reported in Appendix Table A11. In addition to reporting robustness of the first stage relation with regards to controls and county-specific time trends, we also report the δ statistic by Oster (2019) to provide evidence for robustness with respect to the unobservables.[47] The δ statistic we obtain indicates that selection on unobservables would have to be 1.1 to 1.3 times larger selection on the observables in order to explain away our result. Typically, $\delta \geq 1$ is considered a robust result. This increases the confidence in our previous argument that there are no unobservable characteristics that are driving our first stage result but that are potentially left in the error term of the second stage which would lead to a violation of the exclusion restriction.

The IV results are reported in panel b of Table 6. The IV coefficients are more negative than the OLS coefficients in panel a with statistically significant coefficients. A one percentage points increase in the outflow of German-born individuals from 1910 to 1920 was associated with a 5 percent reduction in average annual wages per worker in manufacturing in the baseline specification in column (1). When adding controls, including the time-varying population controls that shut down potential direct labor market effects of the instrument, the effect size increases to negative 7.1 percent in column (2). As in the OLS results, the inclusion of county-specific time trends again increases the effect size with a one percentage points outflow of Germans being associated with an approximate 10.6 percent reduction in our manufacturing wage measure in column (3). Again there is no difference when using the more flexible quadratic county-specific trends in column (4).

The differences in magnitude between the OLS and IV estimates likely arise because the component of the German-born outflow that is identified by the instrument is the component associated with the most virulent anti-German discrimination, which would raise the costs of staying high enough that a larger share of the skilled Germans would have been likely to leave

[47] This test uses the R^2 and coefficient movement of the variable of interest in the regressions with and without controls. Assuming a maximum $R^2 = 1.3$, Oster (2019) defines a selection relationship between observables and unobservables captured by δ, which is the degree to which selection on unobservables would have to be larger than selection on the observables in order for β to be equal to zero.

the county. This implies that IV estimates a local average treatment effect which estimates the effect of the German outflow on average annual manufacturing wages for the counties where Germans were compelled to leave due to the casualty shock and who would have remained otherwise.

5.2 Robustness and Dynamic Treatment Effects

To capture dynamics in the impact of the German outflow, we re-estimated eq. (5) while replacing the Post-WWI dummy with a vector of year fixed effects that omits 1910 as the base year. The coefficient plot in Figure 6 show that the coefficient of the out-flow variable interacted with the 1900 dummy was slightly negative and not statistically significant.[48] The outflow effect in 1920 implies a 5 percent reduction in average annual earnings in the first census year after the war. The negative impact declines to a statistically significant 1.9 percent in 1930 and has gone away by 1940. The most intense anti-German sentiment built up from 1916 when Germany's threats against U.S. shipping increased and even more during the war years of 1917 and 1918 that saw direct involvement of the U.S. It makes sense that the strongest negative effects on earnings of the outflow occurred in 1920 just after the War ended because the firms likely had the highest costs associated with losing skilled workers and retraining them. The effect weakened greatly during the 1920s after this initial period of adjustment to the less negative relationship seen in 1930 and was completely gone by 1940. This adjustment period of at least ten years stands in comparison to the duration of anti-German sentiment of only four years, and the intense discrimination of two years during the war years with U.S. involvement. This suggests that even short-term discriminatory shocks can generate disproportionally longer negative effects in the discriminating economy.

Appendix Table A12 repeats this exercise with the OLS and IV setting. In particular, we estimate equation (5) using only the pre-war periods and the data from each separate post-war Census year, i.e. we separately estimate the effect of going from the pre-war period to 1920, the effect of going from the pre-war period to 1930, and the effect of going from the pre-war period to 1940 both with the OLS and IV estimators. The results of each regression are reported in the corresponding table rows. The IV estimates are again larger than the OLS estimates in 1920 but the overall pattern of the strong negative wage effect in 1920 that mostly had faded away but not completely by 1930.

[48] If this were an actual pre-trend that is just noisily estimated, it would actually make us understate our main result.

As a final robustness check, we test for the spatial correlation of the German outflow measure, log wages in manufacturing per worker, and the WWI casualty rate using the Getis-Ord statistic for positive and negative spatial correlation provided by Kondo (2016). The results from this test are reported in Appendix Table A2. All variables display significant spatial correlation in the 1920 cross section with wages displaying the highest and casualty rates the lowest spatial correlation patterns. Columns (4)-(6) then take these variables, partial out state fixed effects and the pre-war county controls used in sections 3 and 5, and re-estimate the spatial correlation statistic using the residuals. In all cases the spatial correlation reduces significantly. After partialling out these controls, more than 91% of all counties do not show a significant spatial correlation in the WWI casualty rates meaning that, conditional on coviariates, war casualties are as good as randomly assigned in space.

To adjust the previous estimators for potential spatial autocorrelation, Appendix Table A13 reports the OLS and IV results from estimating equation (5) and the corresponding first stage using the standard error correction proposed by Conley (1999). The distance cutoffs are 50, 100, and 200km. The significance of all results survives the correction of standard errors for spatial autocorrelation.

5.3 Evidence for American Born Workers from Linked Census Data

To mirror the exercise in section 4, we also linked American-born individuals from 1910 to 1920 to provide further evidence on the effect of German outflows on natives in the local communities. We then regress

$$\begin{aligned} y_{ict} = \beta \left[\text{WWI German Outflow}_c \times \text{Post-WWI}_t\right] \\ + \alpha_{c,1910} + \lambda_t + X'_{ict}\Gamma + \nu_{ict} \end{aligned} \tag{6}$$

where the notation is the same as in equation 3. The outcomes of the linked American-born individuals we consider are an indicator for whether an individual changed their county of residence between 1910 and 1920, an indicator for working in a semi-skilled job in 1920, an indicator for working in the manufacturing sector in 1920, and the natural logarithm of one plus an individual's occupational income score. Again, this is the closest available substitute for an income measure prior to 1940.[49] We consider three samples of American-born individuals: i)

[49]The occupational income score assigns to each occupation its median wage in 1950. Adding the one before taking logs does not drive the results as these are robust to using the inverse hyperbolic sine transformation as

the full sample of linked American-born men, ii) the sub-sample of those who resided in the same county in both 1910 and 1920, and iii) the sub-sample of individuals who resided in the same county in both Census years and who were employed in manufacturing in 1910.[50]

Table 7 presents the results from regressing equation 6 using the three samples of linked American-born individuals. When considering the full sample, American-born individuals were 1.9 percentage points less likely to move away from their 1910 county of residence for every one percentage points outflow of German-born individuals from before to after the war. This speaks against an argument that Germans would have left counties due to poor performance of the local economies as otherwise also the native-born would have chosen to leave. Instead, they were more likely to stay.

The effect of the German outflow on the other labor market outcomes is remarkably stable across all three samples. This outflow is associated with an increased probability of U.S.-born workers in the county experiencing an occupational downgrade between 1910 and 1920. A one percentage point increase in the German outflow was associated with statistically significant increases in the probability of an occupational downgrade of 0.1 percentage points for the full sample and 0.2 percentage points for the stayers. The effect was even larger at 0.3 percentage points for American-born stayers who had worked in manufacturing in 1910 but the coefficient is imprecisely estimated. Relative to the baseline average of occupational downgrading from 1910 to 1920 of 0.043, this corresponds to a 2.3% increase in the unconditional downgrading probability. While this seems modest, for 2 out of 100 American stayers such a downgrading was a highly undesirable outcome. Dropping from a high-skilled to a semi-skilled job, or from a semi-skilled to a low-skilled job implied a decline in the occupational income score of 25.5% and 51.1%, respectively. Even though few Americans had to pay this cost, it was substantial for those who were affected due to the outflow of Germans from their local economy.

Column (3) shows that the probability of working in a semi-skilled job in 1920, as opposed to working in a low-skilled occupation, dropped by 0.5 percentage points for each percentage point increase in the outflow of Germans. Again this is only statistically significant in the full sample and the sample of stayers, but not for those stayers who had already been employed in manufacturing in 1910. The outmigration of Germans, who typically worked in manufacturing jobs before the war, did not create any positive employment gains for Americans. In fact, the

well. The additional results are available on request.

[50]Note that when estimating equation 6 on the two stayer samples, the county fixed effects are absorbed by the individual fixed effects as residence is then time-invariant within individual. Likewise, the mover outcome is not available for those samples by construction.

estimated coefficients are even negative but not significant in any of the samples. Finally, we find that a one percent point outflow of Germans implied a reduction in Americans' occupational income scores of 4.1 to 4.3%. While this is not a direct income measure, its magnitude is strikingly similar to the effect estimated in the first part of section 5 using the county-level manufacturing data. The linked individual data show that such an effect also exists for those who did not work in manufacturing alone as the estimated coefficient is similar across all three samples.[51]

6 Conclusion

How does a relatively short-lived discrimination shock against a certain group impact their relocation decisions, their future economic outcomes, and the outcomes of the local communities that chose to discriminate against said group? These are the questions we sought to address in this paper by studying the case of Germans in the U.S. during and after World War I. Germans, an economically well-integrated, large, and respected minority group, became hugely unpopular in the U.S. for the duration of World War I, and especially once the United States themselves entered the conflict. Previous work on this topic has studied the adverse effects of this discrimination shock on the Germans themselves including their labor market outcomes at the New York stock exchange (Moser, 2012), the response of Germans to prohibition of their language in schools and general assimilation Fouka (2019, 2020), or compulsory licensing via the Trading with the Enemy Act (Moser and Voena, 2012; Baten et al., 2017).

We add to this literature by studying the internal migration of Germans within the U.S. in response to increased local discrimination and anti-German sentiment during the war. For this purpose, we propose a new measure of an exogenous shifter of anti-German sentiment which is the county-level WWI casualty rate. The main idea is that war casualties suffered abroad by local communities increases the animus against Germans in those communities. We show that higher casualty rates are significantly correlated with Germans being mentioned as *enemies*, or *huns*, i.e. the derogatory term for Germans during the war, as well as a higher frequency of reported tarrings and featherings of Germans. Unlike the newspaper data, WWI casualties can be fairly accurately measured for all counties.

[51] Appendix Table A14 probes for the robustness of our results to the quality of the generated Census links by considering only individuals who had a single unique match which had a birth year difference of at most plus or minus one year.

Using county level data from the Census, we show in difference-in-differences regressions that Germans moved away from counties in the top quintile of the WWI casualty rate. Counties in the bottom quintile saw an increase in the share of German population from before to after the war. Interestingly, Germans moved away from areas in which they were a historically salient minority group such as the Midwest, and towards areas with low shares of Germans such as the South. The relationship between WWI casualties and changes in the share of German-born individuals does not exist before the war, and it does not exist for other groups such as Swedes or Italians. This provides further confidence in the interpretation of our casualty rate as a measure of shifting anti-German sentiment.

We then link German-born individuals using the full-count Census files of 1910 to 1920 to understand the post-war outcomes of cross-county migrants as well as the characteristics that determined their relocation decision. Germans who lived in a county in 1910 that would end up being in the top quintile of the casualty rate were again significantly more likely to move, especially if they were working in manufacturing. Movers had an increased probability of living in the South in 1920, not having naturalized, having experienced an occupational downgrading compared to their 1910 profession, and they were more likely to be farmers after the war. A potential reason is the lack of manufacturing in the South and farming as possible route to escape discrimination of Germans by moving to more remote locations. As in the literature on physical characteristics and labor market outcomes (Hamermesh and Biddle, 1994; Biavaschi et al., 2017), observable characteristics significantly affected the relocation decision of Germans in response to the casualty shock. Stating their mother tongue as German in the Census, having the first name of the German emperor Wilhelm, a common German surname, or having been in the country for a shorter period of time strongly increase the probability of moving county from 1910-20.

The final contribution of the paper is the estimation of the effect of German outmigration on the sending economies, i.e. the counties that essentially discriminated away parts of their German-born population. This relates to the literature on forced migration which has focused on the relocation effects on migrants themselves as well as the receiving economies, but where less is known about the effects on the sending communities (Becker and Ferrara, 2019). Using county-level Census data from 1900-40 in a difference-in-differences setting, we show that counties that saw a larger outflow of Germans from 1910-20 experienced a drop in average manufacturing wages per worker. Since Germans were disproportionally employed in the man-

ufacturing sector before the war, losing these workers appears to have reduced productivity and thus wages in the manufacturing sector in the post-war period. Instrumenting the German outflow with being in the top quintile of the WWI casualty rate confirms this result. Furthermore the dynamic wage effect is such that the largest drop in wages due to a reduction in the share of German-born population during the war is strongest in the immediate post-war period and then begins to fade out.

Using individual linked Census data for American-born workers, we do not find any positive employment gains in manufacturing. In fact, Americans were less likely to move and more likely to experience an occupational downgrading and reduced occupational income scores from before to after the war in counties where more Germans left. We therefore conclude that counties which chose to satisfy a short-run anti-group sentiment by discriminating away their German population paid for this in the longer term with reduced wages in the manufacturing sector and employment opportunities for native workers. Even though the discrimination shock was relatively short-lived over a period of 4 years, the wage drop persisted for more than a decade.

References

Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos, "Marrying Up: The Role of Sex Ratio in Assortative Matching," *American Economic Journal: Applied Economics*, 2011, *3* (3), 124–157.

_ , Leah Boustan, and Katherine Eriksson, "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, 2014, *122* (3), 467–506.

Akbulut-Yuksel, Mevlude and Mutlu Yuksel, "The Long-Term Direct and External Effects of Jewish Expulsions in Nazi Germany," *American Economic Journal: Economic Policy*, 2015, *7* (3), 58–85.

Alesina, Alberto, Johann Harnoss, and Hillel Rapoport, "Birthplace diversity and economic prosperity," *Journal of Economic Growth*, 2016, *21*, 101–138.

Alsan, Marcella and Marianne Wanamaker, "Tuskegee and the Health of Black Men," *Quarterly Journal of Economics*, 2018, *133* (1), 407–455.

_ , Owen Garrick, and Grant Graziani, "Does Diversity Matter for Health? Experimental Evidence from Oakland," *American Economic Review*, 2019, *109* (12), 4071–4111.

Baten, Joerg, Nicola Bianchi, and Petra Moser, "Compulsory licensing and innovation - Historical evidence from German patents after WWI," *Journal of Development Economics*, 2017, *126*, 231–242.

Becker, Gary, *The Economics of Discrimination*, University of Chicago Press, Chicago, IL, 1957.

Becker, Sascha O. and Andreas Ferrara, "Consequences of Forced Migration: A Survey of Recent Findings," *Labour Economics*, 2019, *59*, 1–16.

_ , Irena Grosfeld, Pauline Grosjean, Nico Voigtländer, and Ekaterina Zhuravskaya, "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers," *American Economic Review, forthcoming.*, 2020.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economics Perspectives*, 2014, *28* (2), 29–50.

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 2004, *94* (4), 991–1013.

Biavaschi, Costanza, Corrado Giulietti, and Zahra Siddique, "The Economic Payoff of Name Americanization," *Journal of Labor Economics*, 2017, *35* (4), 1089–1116.

Boehnke, Jörn and Victor Gay, "The Missing Men: World War I and Female Labor Force Participation," *TSE Working Paper 20-1064*, 2020.

Breen, William J., *Uncle Sam at Home: Civilian Mobilization, Wartime Federalism, and the Council of National Defense, 1917-1919*, Greenwood Press, Westport, CT, 1984.

Card, David and Alan B. Krueger, "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics*, 1992, *107* (1), 151–200.

Cavalcanti, Tiago and Jose Tavares, "The Output Cost of Gender Discrimination: A Model-based Macroeconomics Estimate," *Economic Journal*, 2016, *126* (590), 109–134.

Charles, Kerwin and Jonathan Guryan, "Prejudice and wages: an empirical assessment of Becker's The economics of discrimination," *Journal of Political Economy*, 2008, *116* (5), 773–809.

Committee on Immigration, *Reports of the Immigration Commission - Immigrants in Industries, Part 2: Iron and Steel Manufacturing*, Government Printing Office, Washington, DC, 1911.

_ , *Reports of the Immigration Commission - Immigrants in Industries, Part 23: Summary Report of Immigrants in Manufacturing and Mining*, Government Printing Office, Washington, DC, 1911.

Conley, Timothy G., "GMM estimation with cross sectional dependence," *Journal of Econometrics*,

1999, *92* (1), 1–45.

Cook, Lisa D., "Violence and Economic Activity: Evidence from African American Patents, 1870-1940," *Journal of Economic Growth*, 2014, *19* (2), 221–257.

Crowder, Enoch H., *Final Report of the Provost Marshal General to the Secretary of War on the Operations of the Selective Service System to July 15, 1919*, Government Printing Office, Washington, DC, 1920.

D'Acunto, Francesco, Marcel Prokopczuk, and Michael Weber, "Historical Antisemitism, Ethnic Specialization, and Financial Development," *Review of Economic Studies*, 2019, *86* (3), 1170–1206.

Feigenbaum, James, Soumyajit Mazumder, and Cory Smith, "When Coercive Economies Fail: The Political Economy of the U.S. South After the Boll Weevil," *Working Paper*, 2019.

Ferrara, Andreas, "World War II and African American Socioeconomic Progress," *CAGE Warwick working paper 387*, 2018.

Fouka, Vasiliki, "How do immigrants respond to discrimination? The case of Germans in the US during World War I," *American Political Science Review*, 2019, *113* (2), 405–422.

_ , "Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I," *Review of Economic Studies*, 2020, *87* (1), 204–239.

Goldin, Claudia and Lawrence F. Katz, "The Origins of Technology-Skill Complementarity," *Quarterly Journal of Economics*, 1998, *113* (3), 693–732.

Grogger, Jeffrey, "Speech Patterns and Racial Wage Inequality," *Journal of Human Resources*, 2011, *46* (1), 1–25.

Haines, Michael R., "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002," *Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. https://doi.org/10.3886/ICPSR02896.v3*, 2010.

Hamermesh, Daniel S. and Jeff E. Biddle, "Beauty and the Labor Market," *American Economic Review*, 1994, *84* (5), 1174–1194.

Haulsee, William M., Frank G. Howe, and Alfred C. Doyle, *Soldiers of the Great War, Vols. 1-3*, Soldiers Record Publishing Association, 1920.

Heedegaard, Morten S. and J.-R. Tyran, "The Price of Prejudice," *American Economic Journal: Applied Economics*, 2018, *10* (1), 40–63.

Hegi, Benjamin P., "Old Time Good Germans: German-Americans in Cooke County, Texas, during World War I," *Southwestern Historical Quarterly*, 2005, *109* (2), 234–257.

Higham, John, *Strangers in the Land*, Rutgers University Press, 1998.

Hornung, Erik, "Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia," *American Economic Review*, 2014, *104* (1), 84–122.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow, "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Kendrick, John W., *Productivity Trends in the United States*, Princeton University Press, 1961.

Kondo, Keisuke, "Hot and cold spot analysis using Stata," *Stata Journal*, 2016, *16* (3), 612–631.

Lang, Kevin and Jee-Yeon K. Lehmann, "Racial Discrimination in the Labor Market: Theory and Empirics," *Journal of Economic Literature*, 2012, *50* (4), 959–1006.

_ **and Michael Manove**, "Education and Labor Market Discrimination," *American Economic Review*, 2011, *101* (4), 1467–1496.

_ **,** _ **, and William T. Dickens**, "Education and Labor Market Discrimination," *American Economic Review*, 2005, *95* (4), 1327–1340.

Lee, Jongkwan, Giovanni Peri, and Vasil Yasenov, "The Labor Market Effects of Mexican Repatriations: Longitudinal Evidence from the 1930s," *IZA Discussion Paper No. 12689*, 2019.

Lleras-Muney, Adriana and Allison M. Shertzer, "Did the Americanization Movement Succeed? An Evaluation of the Effect of English-Only and Compulsory Schooling Laws on Immigrants," *American Economic Journal: Economic Policy*, 2015, *7* (3), 258–290.

Lübke, Frederick C., *Bonds of Loyalty*, Northern Illinois University Press, 1974.

_ , *Germans in the New World: Essays in the History of Immigration*, University of Illinois Press, 1999.

Matheis, Mike, "Local Economic Impacts of Coal Mining in the United States 1870 to 1970," *Journal of Economic History*, 2016, *76* (4), 1152–1181.

Mazumder, Soumyajit, "Becoming White: How Mass Warfare Turned Immigrants into Americans," *mimeo*, 2018.

Meyer, Bruce D. and Nikolas Mittag, "Misclassification in binary choice models," *Journal of Econometrics*, 2017, *200* (2), 295–311.

Miller, Thomas W., *Report of the Alien Property Custodian*, Government Printing Office, 1922.

Moser, Petra, "Taste-based discrimination evidence from a shift in ethnic preferences after WWI," *Explorations in Economic History*, 2012, *49* (2), 167–188.

_ **and Alessandra Voena**, "Compulsory Licensing: Evidence from the Trading with the Enemy Act," *American Economic Review*, 2012, *102* (1), 396–427.

Neumark, David, "Experimental Research on Labor Market Discrimination," *Journal of Economic Literature*, 2018, *56* (3), 799–866.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, 2019, *37* (2), 187–204.

Pascali, Luigi, "Banks and Development: Jewish Communities in the Italian Renaissance and Current Economic Performance," *Review of Economics and Statistics*, 2016, *98* (1), 140–158.

Phelps, Edmund, "The Statistical Theory of Racism and Sexism," *American Economic Review*, 1972, *62* (4), 659–661.

Roberts, Evan and Alexandra Burda, "Correlates and Consequences of American War Casualties in World War I," *University of Minnesota Working Paper*, 2018, *2018-3*.

Schmick, Ethan J. and Allison M. Shertzer, "The Impact of Early Investments in Urban School Systems in the United States," *NBER Working Paper No. 25663*, 2020.

Tabellini, Marco, "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration," *Review of Economic Studies*, 2020, *87* (1), 454–486.

Testa, Patrick, "The Economic Legacy of Expulsion: Lessons from Postwar Czechoslovakia," *SSRN Working Paper*, 2018, *https://ssrn.com/abstract=3251541*.

Washington, Thomas, *Officers and Enlisted Men of the United States Navy Who Lost Their Lives during the World War*, Government Printing Office, 1920.

Zunz, Olivier, *The Changing Face of Inequality: Urbanization, Industrial Development, and Immigrants in Detroit, 1880-1920*, University of Chicago Press, 2000.

Tables

Table 1: WWI Casualty Rates and Anti-German Reporting in the News

	Enemy	Huns	Tarring and feathering
	(1)	(2)	(3)
WWI Casualty Rate$_c$	4.745**	2.370**	0.674**
	(2.113)	(1.182)	(0.317)
Outcome mean	18.389	4.886	1.794
Observations	2,199	2,199	2,199
Adj. R^2	0.141	0.127	0.078

Note: Cross-sectional county-level regressions of newspaper based anti-German sentiment measures on the WWI casualty rate. Outcomes express the share of articles mentioning Germans as enemy or huns or that report tarring and feathering relative to all articles mentioning Germans in any context between the war years of 1914 and 1918 in county c. Newspaper outlets are geo-located at the county level. If a county did not have a newspaper, the closest newspaper was assigned and weighted by distance to the nearest outlet. Controls include the WWI draft rate, the total number of articles published in the county over the given time period and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, as well as state fixed effects. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: County Data Summary Statistics

	Mean	St. Dev.	Min	Max
Outcomes				
Percent German-born individuals	1.370	2.187	0	21.861
Reduction in % German Population, 1910-20	0.213	0.684	0	8.767
Log Per Capita Wages in Manufacturing	1.332	0.502	-0.583	2.931
WWI Variables				
WWI Casualty Rate	0.264	0.138	0.017	2.911
Dummy Top Casualty Quintile	0.143	0.350	0	1
Casualty Rate in the Top Quintile	0.469	0.160	0.362	2.911
WWI Draft Rate	10.072	4.594	1.783	100
Pre-War Controls				
Pre-WWI German Share	1.948	3.339	0	26.954
Pre-WWI Population	15,036	14,927	0	110,368
Pre-WWI Urbanization Rate	1.022	5.926	0	73.839
Pre-WWI Male-to-Female Ratio	0.899	0.541	0	6.897
Pre-WWI Share of Manufacturing Employment	2.090	2.991	0	26.247
Newspaper Measures				
% Articles Mentioning Germans as Enemies	18.365	12.435	0	100
% Articles Mentioning Germans as Huns	4.858	6.877	0	100
% Articles Reporting Tarring and Feathering	1.807	2.194	0	20
Number of Articles per County (1917-18)	962.926	2516.143	1	51,186
Observations	10,474			

Note: County level characteristics for U.S. counties in the estimation sample from 1900 to 1940. Wages in manufacturing are deflated and have 1910 as base year. The share of newspaper articles mentioning Germans as Huns or enemies, or that report tarring and feathering of Germans are relative to the total number of articles mentioning Germans in any context in 1917 and 1918, i.e. the years of U.S. involvement in WWI.

Table 3: Effect of WWI Casualties on % German Population

	Outcome: Share of German Population (pre-war mean = 1.815)			
	(1)	(2)	(3)	(4)
Q_1(Casualty Rate)$_c$ × Post-WWI$_t$	0.377***	0.367***	0.152***	0.154***
	(0.075)	(0.074)	(0.048)	(0.041)
Q_2(Casualty Rate)$_c$ × Post-WWI$_t$	0.128	0.158*	0.063	0.063
	(0.087)	(0.085)	(0.046)	(0.048)
Q_4(Casualty Rate)$_c$ × Post-WWI$_t$	-0.068	-0.087	-0.087	-0.087
	(0.099)	(0.098)	(0.077)	(0.060)
Q_5(Casualty Rate)$_c$ × Post-WWI$_t$	-0.302***	-0.314***	-0.195*	-0.195***
	(0.105)	(0.106)	(0.110)	(0.065)
Observations	10,367	10,367	10,367	10,367
Counties	2,230	2,230	2,230	2,230
Adj. R^2	0.804	0.811	0.980	0.968
Controls		Yes	Yes	Yes
Linear county time trends			Yes	
Quadratic county time trends				Yes

Note: Difference-in-differences regressions of the share of Germans in county c in decade t from 1900-40, interacting quintiles of the WWI casualty rate with a post-war indicator. Q_τ indicates the τ^{th} quintile of the casualty rate distribution. The omitted comparison quintile is quintile three. The pre-war outcome mean is measured in 1910. All regressions include county and decade fixed effects. The sample period is 1900-40. If included, controls contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: WWI Casualties and Outcomes of Germans in Linked Census Data

	Mover	Distance moved	South	Not naturalized	Occupational downgrade	Farmer
	(1)	(2)	(3)	(4)	(5)	(6)
Q_5(Casualty Rate)$_c$ ×	0.634***	584.565***	0.034***	0.043*	0.026**	0.200***
Post-WWI$_t$	(0.172)	(41.679)	(0.008)	(0.025)	(0.011)	(0.021)
Outcome mean	0.247	301.5	0.0250	0.231	0.155	0.059
Observations	290,488	121,720	143,314	143,314	143,314	143,314
Adj. R^2	0.384	0.444	0.074	0.215	0.579	0.109
Mover sample		yes	yes	yes	yes	yes

Note: Difference-in-differences regressions of individual's outcomes on an indicator for whether county c is in the top quintile of the WWI casualty rate distribution which is interacted with a post-war indicator. The sample is a two period panel and contains German-born individuals who were linked from the 1910 to 1920 full count Census files. Mover is an indicator for whether an individual left their county of residence between 1910-20. Distance is the distance in miles between a person's residence in 1910 and 1920 conditional on having moved between the two Census years. All other outcomes are indicators in 1920 for living in the South, having obtained citizen status by 1920, having experienced an occupational downgrade (i.e. from high- to semi-/low-skilled, or from semi to low-skilled jobs using the 1950 occupational definitions of the U.S. Census Bureau), and for working in agriculture in 1920 when the individual's occupation in 1910 was not in agriculture. The mover sample only includes individuals who moved county between 1910-20. All regressions include county fixed effects, a year indicator for 1920, birth year and place of birth fixed effects, as well as baseline controls measured in 1910 and interacted with the 1920 indicator. The baseline controls include the following measures from the 1910 Census: indicators for urban status, eight skill groups, farm status, employment status, marital status, years since entry to the U.S. in bins (0-5, 6-10, 11-15, and 16-20 years), school attendance, labor force participation, and count measures for family size and the number of weeks spent in unemployment in 1909. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Determinants of the Individual Relocation Decision

	Outcome: Moved county between 1910-20 (mean = 0.247)			
	(1)	(2)	(3)	(4)
Q_5(Casualty Rate)$_c$ × Post-WWI$_t$	0.557***	0.557***	0.523***	0.503***
	(0.172)	(0.172)	(0.166)	(0.164)
Q_5(Casualty Rate)$_c$ × Post-WWI$_t$ interacted with:				
First language is German	0.077***	0.077***	0.075***	0.076***
	(0.029)	(0.029)	(0.029)	(0.029)
First name is Wilhelm		0.161**	0.189***	0.175**
		(0.073)	(0.072)	(0.073)
Common German last name			0.288***	0.278***
			(0.012)	(0.011)
In U.S. for 0-5 years				0.184***
				(0.034)
In U.S. for 6-10 years				0.230***
				(0.018)
In U.S. for 11-15 years				0.059***
				(0.014)
In U.S. for 16-20 years				0.059***
				(0.013)
Observations	290,488	290,488	290,488	290,488
Adj. R^2	0.384	0.384	0.392	0.394

Note: Difference-in-differences regressions of individual's outcomes on an indicator for whether county c is in the top quintile of the WWI casualty rate distribution which is interacted with a post-war indicator. The sample is a two period panel and contains German-born individuals who were linked from the 1910 to 1920 full count Census files. Mover is an indicator for whether an individual left their county of residence between 1910-20. The top casualty quintile and 1920 dummy interaction is further interacted with the following variables. The first language is German indicator equals one if an individual reports German as their native language (even though the person may speak English). The indicator for the first name being Wilhelm captures effects of having the name of the German Emperor at the time. Common German last name is an indicator for having one of the 30 most common German surnames in the 1910 Census (Schmidt, Meyer, Schultz, Wagner, Weber, Hoffman, Schneider, Becker, Schroeder, Mueller, Wolf, Peters, Bauer, Fischer, Koch, Klein, Zimmerman, Krueger, Keller, Beck, Kramer, Mayer, Krause, Schwartz, Hahn, Schmitt, Hartman, Lange, Schaefer, Kaiser). All regressions include county fixed effects, a year indicator for 1920, birth year and place of birth fixed effects, as well as baseline controls measured in 1910 and interacted with the 1920 indicator. The baseline controls include the following measures from the 1910 Census: indicators for urban status, eight skill groups, farm status, employment status, marital status, years since entry to the U.S. in bins (0-5, 6-10, 11-15, and 16-20 years), school attendance, labor force participation, and count measures for family size and the number of weeks spent in unemployment in 1909. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: OLS and IV Results for German Outflows and Log Wages in Manufacturing

	Outcome: log manufacturing wage per capita (pre-war mean = 1.905)			
	Panel a: OLS			
	(1)	(2)	(3)	(4)
WWI German Outflow$_c\times$	-0.011**	-0.017***	-0.072***	-0.070***
Post-WWI$_t$	(0.004)	(0.004)	(0.025)	(0.025)
Observations	10,609	10,474	10,474	10,474
Counties	2,302	2,258	2,258	2,258
Adj. R^2	0.874	0.878	0.911	0.857
Controls		Yes	Yes	Yes
Linear county time trends			Yes	
Quadratic county time trends				Yes
	Panel b: IV			
	(1)	(2)	(3)	(4)
WWI German Outflow$_c\times$	-0.050*	-0.091**	-0.114**	-0.115**
Post-WWI$_t$	(0.029)	(0.038)	(0.053)	(0.053)
Observations	10,367	10,367	10,367	10,367
Counties	2,274	2,230	2,230	2,230
K-P F-stat	29.052	30.170	22.223	22.204
Controls		Yes	Yes	Yes
Linear county time trends			Yes	
Quadratic county time trends				Yes

Note: OLS and IV regressions of the log per capita wage in manufacturing in county c and decade t on the % outflow of German population (measured from 1910-20) interacted with a post-WWI indicator. In panel b, the outflow of Germans is instrumented with the WWI casualty rate. The sample period is 1900-40. All regressions include county and year fixed effects. If included, controls contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. The regressions also include time varying measures for the total population size and the size of the male population. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: German Outflows and Economic Outcomes of Americans

	Panel a: full sample				
	(1)	(2)	(3)	(4)	(5)
	Mover	Occupational downgrade	Semi-Skilled	Manufacturing	log occ. score
WWI German Outflow$_c$×	-0.019***	0.001**	-0.005***	-0.001	-0.041***
Post-WWI$_t$	(0.006)	(0.000)	(0.001)	(0.001)	(0.005)
Outcome mean	0.186	0.043	0.142	0.059	0.891
Observations	4,696,376	4,696,376	4,696,376	4,696,376	4,696,326
Adj. R^2	0.245	0.475	0.508	0.234	0.621
	Panel b: stayers				
		(2)	(3)	(4)	(5)
		Occupational downgrade	Semi-Skilled	Manufacturing	log occ. score
WWI German Outflow$_c$×		0.002**	-0.005***	-0.000	-0.043***
Post-WWI$_t$		(0.001)	(0.001)	(0.001)	(0.006)
Outcome mean		0.036	0.121	0.048	0.876
Observations		2,952,508	2,952,508	2,952,508	2,952,472
Adj. R^2		0.433	0.555	0.298	0.642
	Panel c: stayers working in manufacturing				
		(2)	(3)	(4)	(5)
		Occupational downgrade	Semi-Skilled	Manufacturing	log occ. score
WWI German Outflow$_c$×		0.003	-0.005	-0.011	-0.042***
Post-WWI$_t$		(0.004)	(0.007)	(0.007)	(0.016)
Outcome mean		0.147	0.515	0.682	1.225
Observations		115,788	115,788	115,788	115,788
Adj. R^2		0.147	0.551	0.182	1.225

Note: Difference-in-differences regressions of individuals' outcomes on the percentage decline in the German population in their 1910 county of residence from 1910 to 1920 interacted with a post-war indicator. The sample is a two period panel and contains white American-born individuals who were linked from the 1910 to 1920 full count Census files. Mover is an indicator for whether an individual left their county of residence between 1910-20. Occupational downgrade is an indicator for having experienced an occupational downgrade from 1910 to 1920 (i.e. from high- to semi-/low-skilled, or from semi to low-skilled jobs using the 1950 occupational definitions of the U.S. Census Bureau). Semi-skilled is an indicator for whether an individual worked in a semi-skilled occupation in 1920. Manufacturing is an indicator for working in the manufacturing sector in 1920, and log occ. score is the natural logarithm of one plus an individual's occupational income score in 1920. All regressions include county fixed effects, a year indicator for 1920, birth year and place of birth fixed effects, as well as baseline controls measured in 1910 and interacted with the 1920 indicator. The baseline controls include the following measures from the 1910 Census: indicators for urban status, eight skill groups, farm status, employment status, marital status, years since entry to the U.S. in bins (0-5, 6-10, 11-15, and 16-20 years), school attendance, labor force participation, and count measures for family size and the number of weeks spent in unemployment in 1909. The stayer sample consists of individuals who lived in the same county in both 1910 and 1920. The sample of stayers in panel c consists of individuals who were employed in manufacturing in 1910. County fixed effects are absorbed by the individual fixed effects in the stayer samples. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

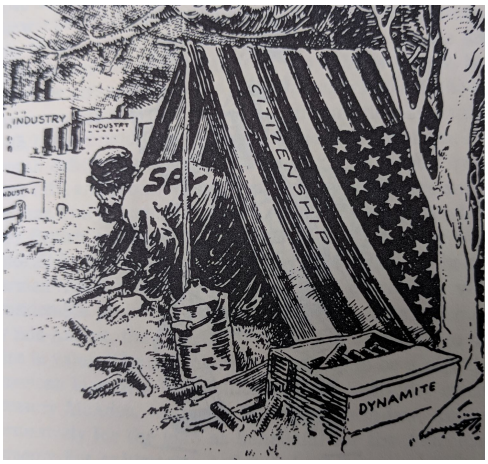
Figures

Figure 1: Anti-German Posters and Prints during WWI

(a) War Bond Advertisement

(b) A German Soldier as Child Murderer

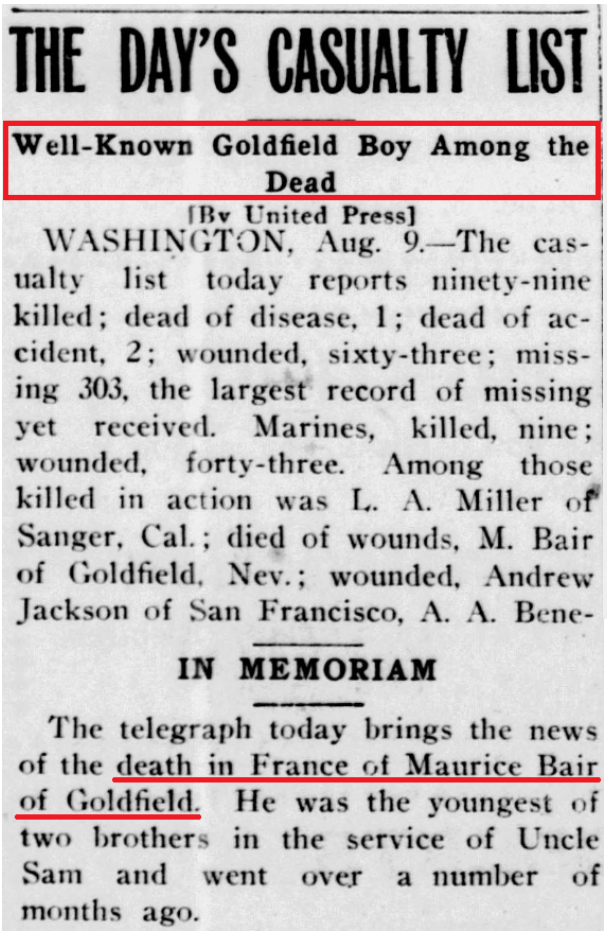
(c) Defamation of German-Americans



Note: Examples of Anti-German propaganda during WWI. Panel a) shows a German soldier as rapist who is stopped by an American soldier to promote the purchase of war bonds. Panel b) depicts a German soldier as child murderer. Panel c) discredits German-Americans who allegedly hide under the cover of U.S. citizenship to act as spies who target the U.S. industry shown in the background which they sabotage with the dynamite in the front of the image. Image source: Lübke (1974).

Figure 2: Anti-German Sentiment and War Casualties in Newspapers

(a) Local Casualty Reports



THE DAY'S CASUALTY LIST

Well-Known Goldfield Boy Among the Dead

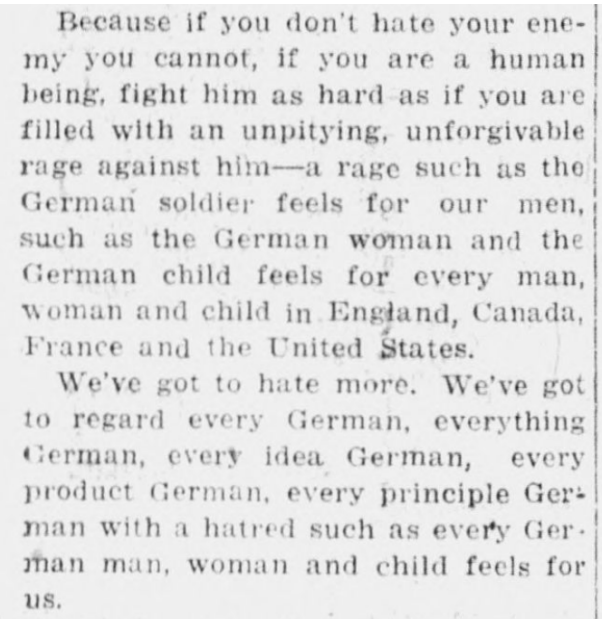
[By United Press]

WASHINGTON, Aug. 9.—The casualty list today reports ninety-nine killed; dead of disease, 1; dead of accident, 2; wounded, sixty-three; missing 303, the largest record of missing yet received. Marines, killed, nine; wounded, forty-three. Among those killed in action was L. A. Miller of Sanger, Cal.; died of wounds, M. Bair of Goldfield, Nev.; wounded, Andrew Jackson of San Francisco, A. A. Bene-

IN MEMORIAM

The telegraph today brings the news of the death in France of Maurice Bair of Goldfield. He was the youngest of two brothers in the service of Uncle Sam and went over a number of months ago.

(b) Anti-German Sentiment



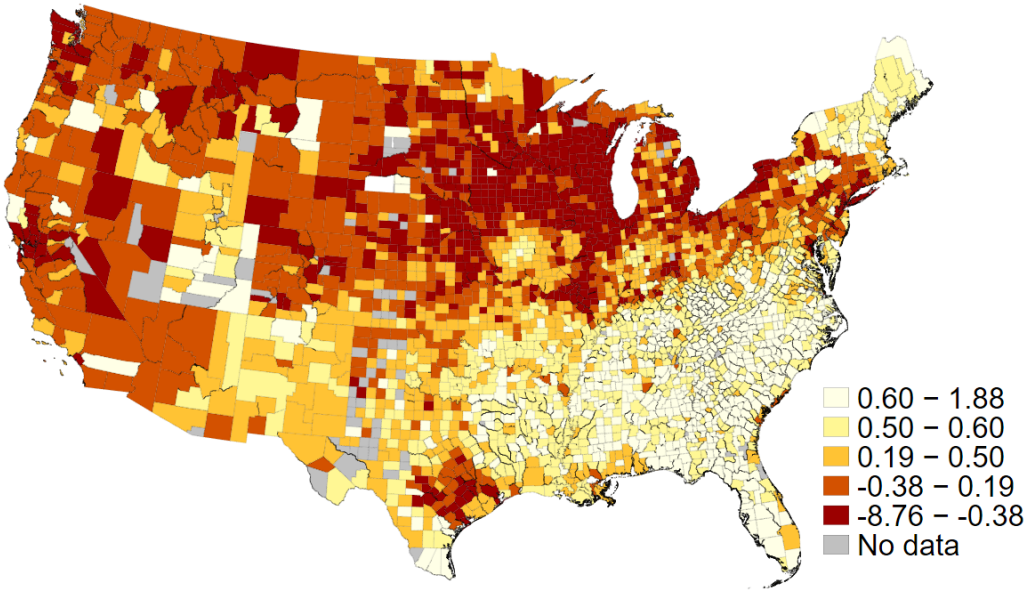
Because if you don't hate your enemy you cannot, if you are a human being, fight him as hard as if you are filled with an unpitying, unforgivable rage against him—a rage such as the German soldier feels for our men, such as the German woman and the German child feels for every man, woman and child in England, Canada, France and the United States.

We've got to hate more. We've got to regard every German, everything German, every idea German, every product German, every principle German with a hatred such as every German man, woman and child feels for us.

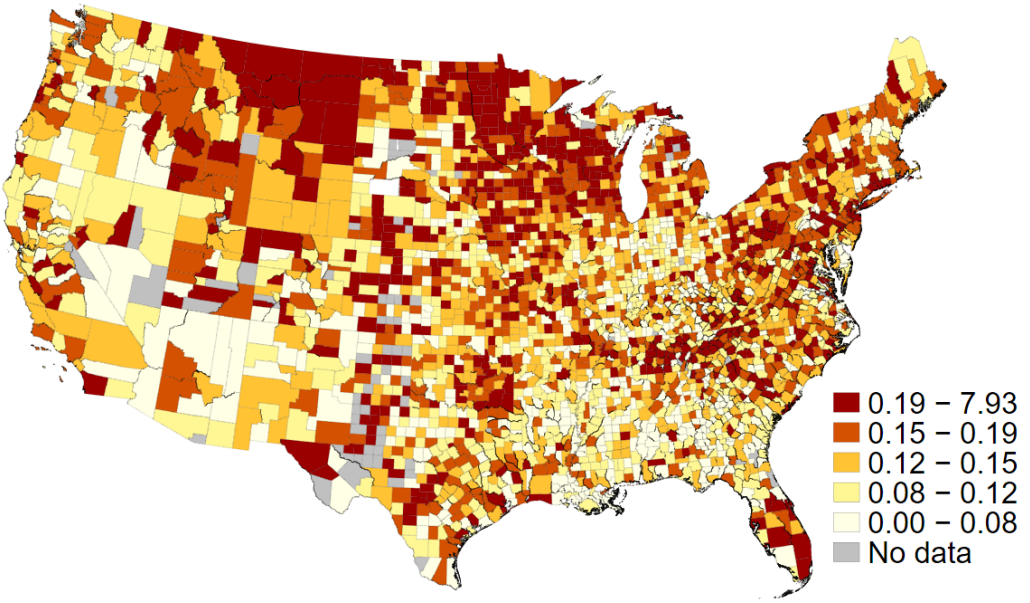
Note: Figure (a) shows an example of publicly communicated anti-German sentiment in the Albuquerque Morning Journal in the paper on May 13th 1918. Panel (b) displays reporting on the day's casualty list and the highlighting of local war deaths in the Carson City Daily paper of August 9th 1918.

Figure 3: Spatial Distribution of German Population Flows and WWI Casualty Rate

(a) Change % German-born 1910-20



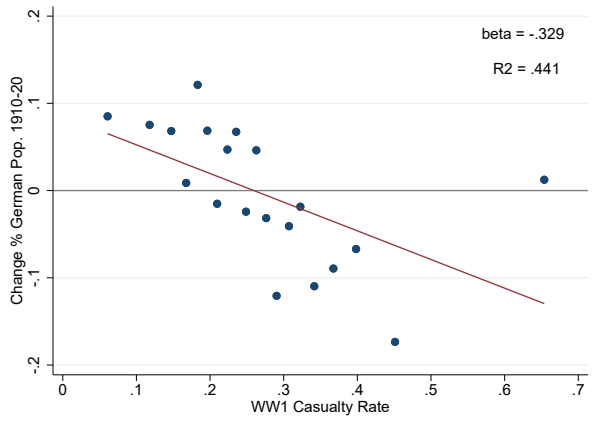
(b) WWI Casualty Rate



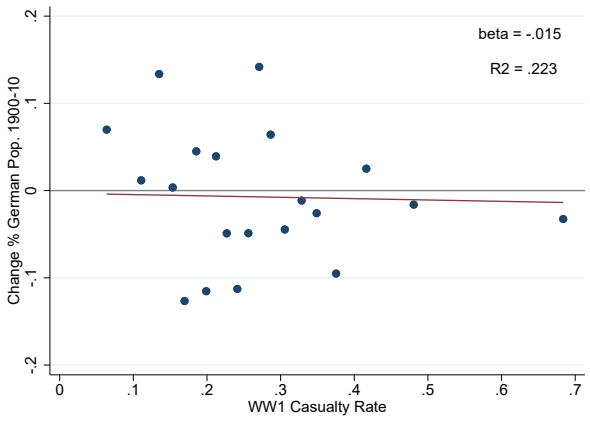
Note: Panel (a) maps the quintiles of the change in the county-level share of the German-born population from 1910 to 1920. Total population changes have been residualized out to avoid confounding changes in the share by an influx of other immigrant groups, for instance. Panel (b) maps the county-level WWI casualty rate which is defined as the total number of WWI deaths over the male population of service eligible age in 1910 times one hundred.

Figure 4: Population Changes and War Casualties by Immigrant Group

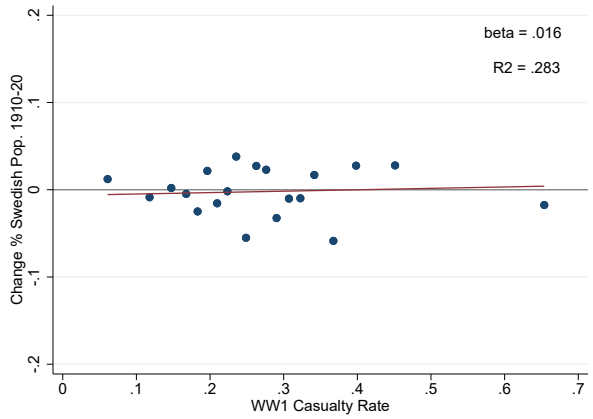
(a) Δ German share, 1910-20



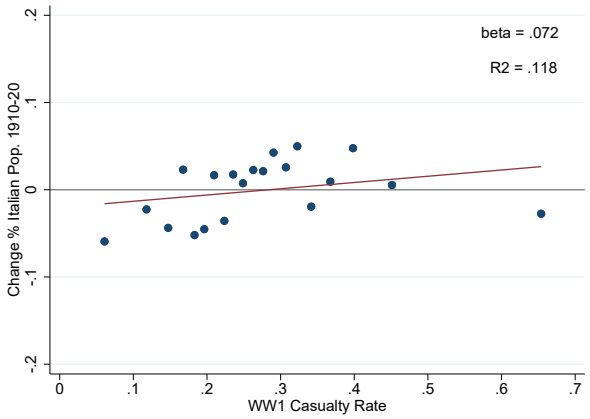
(b) Δ German share, 1900-10



(c) Δ Swedes share, 1910-20

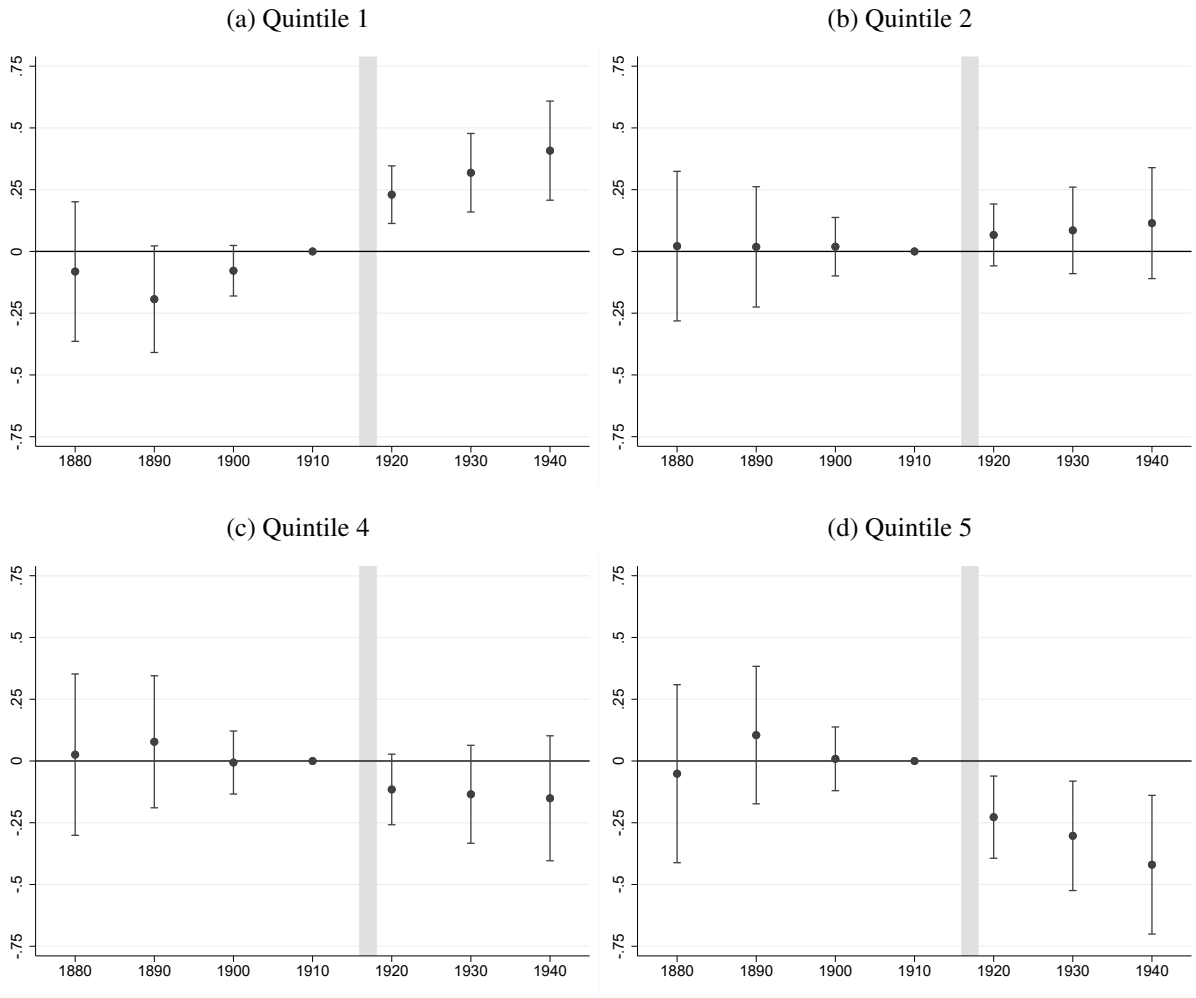


(d) Δ Italian share, 1910-20



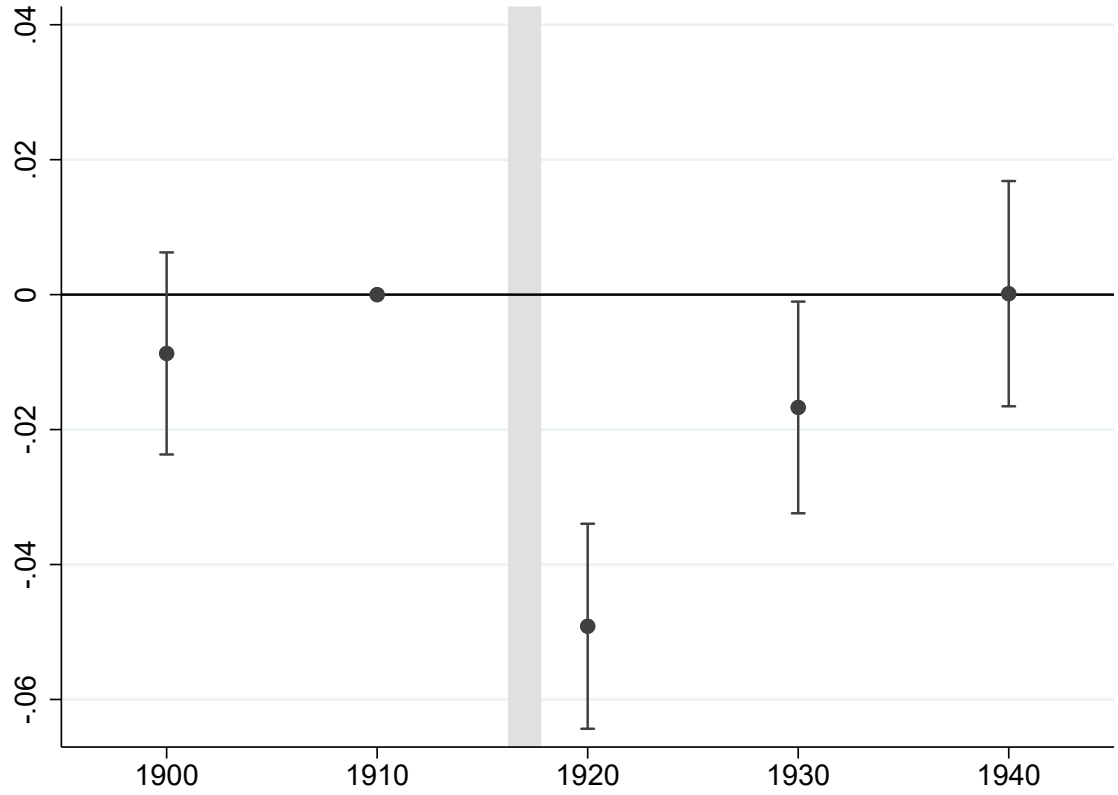
Note: Binned scatter plots for the relation between the change (Δ) in the population share of a given group (individuals born in Germany, Italy, Sweden) and the WWI casualty rate within U.S. states. Measures in the top right corner display the slope coefficient (beta) and fit (R^2) of the regression line in each plot.

Figure 5: Effect of WWI Casualty Rates on the Share of Germans by Casualty Quintile



Note: Difference-in-differences coefficient plots from county-level regressions of % German-born on the WWI casualty rate interacted with time fixed effects (base = 1910). All regressions include county and year fixed effects and controls (WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate). Standard errors are clustered at the county level and represented as 95% error bars. The shaded area marks the years of U.S. involvement in the war. Panel (a) shows the impact of the casualty rate on the county share of Germans in counties with the lowest (bottom quintile) casualty rates. Panel (b) shows the same for counties in the top quintile of the casualty rate distribution.

Figure 6: The Dynamic Effect of German Outflows on Manufacturing Wages



Note: Coefficient plot from a regression of log manufacturing wages per capita on the % decline in the share of German population from 1910-20 interacted with year fixed effects (base = 1910). The regression includes county and time fixed effects and controls which contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate. Standard errors are clustered at the county-level and reported as 95% error bars around the point estimates.

Appendix

Table A1: Average Economic and Social Characteristics by Group in 1910

	Germans	Swedish	English	Italians	Americans
% urban	0.655	0.578	0.705	0.724	0.363
% farmers	0.209	0.232	0.102	0.027	0.377
% home owners	0.499	0.475	0.364	0.172	0.455
% naturalized	0.815	0.793	0.739	0.247	
% literate	0.957	0.976	0.987	0.677	0.911
% speak English	0.825	0.903	0.967	0.470	0.964
% business owner	0.147	0.116	0.082	0.038	0.144
Earnings score	111.404	94.190	108.083	87.319	115.557
Education score	77.154	59.783	68.500	49.871	92.013
Observations	1,198,372	347,935	404,200	820,743	17,474,027

Note: Average characteristics of immigrants and Americans (U.S.-born with both parents born in the United States) from the 1910 Census. Literacy refers to both reading and writing. Business owners refers to those whose employment status is assigned as *employer* in the Census. Occupational earnings scores were constructed from 1950 data to compute the median earnings of each occupation. The occupational education scores measures the percentage of individuals per occupation with one or more years of college education in 1950.

Table A2: Getis-Ord $G_i^*(d)$ Test for Spatial Clustering

	Raw Variables			Net of Covariates		
Getis-Ord $G_i^*(d)$	outflow	ln(wages)	casualty rates	outflow	ln(wages)	casualty rates
z-score interval	(1)	(2)	(3)	(4)	(5)	(6)
$z \leq$ -2.58	801	764	406	56	52	2
-2.58 $<$ z $\leq$ -1.96	160	83	116	59	136	40
-1.96 $<$ z $<$ 1.96	780	619	1,191	1,972	1,841	2,008
1.96 $\leq$ z $<$ 2.58	43	139	110	60	99	57
2.58 $\leq$ z	419	598	357	56	75	73
Observations	2,203	2,203	2,203	2,203	2,203	2,203
State FE and controls				Yes	Yes	Yes

Note: Getis-Ord $G_i^*(d)$ test for local spatial independence in the cross sectional variables in 1920 in a 200km radius with a binary spatial weight matrix. Local spatial independence is given when the z-score on the corresponding test statistic lies within -1.96 $<$ z $<$ 1.96. Spatial clusters of unusually low/high variable values (cold/hot spots) are found for counties with z-scores of z $\leq$ -1.96 (cold spots) and 1.96 $\leq$ z (hot spots). The number of counties in each z-score bin is provided in the rows of the table. Each county is identified by the latitude and longitude of its centroid. The first three columns use the raw variables. The last three columns partial out controls (WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator) and state fixed effects.

Table A3: Linked Individual Data Summary Statistics in 1910

	Mean	St. Dev.	Min	Max
Demographics				
Moved county form 1910-20	0.494	0.500	0	1
Age	41.696	10.985	15	60
Married	0.758	0.428	0	1
First language is English	0.880	0.325	0	1
First language is German	0.094	0.292	0	1
First name is Wilhelm	0.001	0.027	0	1
Common German surname	0.098	0.297	0	1
Lives on a farm	0.259	0.438	0	1
Lives in a city	0.606	0.489	0	1
In the U.S. for 0 to 5 years	0.086	0.280	0	1
In the U.S. for 6 to 10 years	0.071	0.257	0	1
In the U.S. for 11 to 15 years	0.049	0.217	0	1
In the U.S. for 16 to 20 years	0.155	0.362	0	1
Economic characteristics				
Participates in the labor force	0.804	0.397	0	1
Manufacturing worker	0.200	0.400	0	1
Owns house	0.544	0.498	0	1
Self-employed	0.185	0.388	0	1
Observations				

Note: Summary statistics of the 1910 baseline characteristics of the linked German-born individuals. Individuals were linked from the 1910 to 1920 full count U.S. Census. The omitted category for years lived in the U.S. is non-response. The common German surname variable is defined as indicator for whether the individual's surname was among the 30 most common German surnames in the 1910 Census (Schmidt, Meyer, Schultz, Wagner, Weber, Hoffman, Schneider, Becker, Schroeder, Mueller, Wolf, Peters, Bauer, Fischer, Koch, Klein, Zimmerman, Krueger, Keller, Beck, Kramer, Mayer, Krause, Schwartz, Hahn, Schmitt, Hartman, Lange, Schaefer, Kaiser). The first language refers to the language reported as first language in the Census, i.e. individuals often speak both languages but choose which of them they report as their first language. The indicator for the first name being Wilhelm refers to the name of the German Emperor at the time.

Table A4: Relocation Response by Industry Group

	Mover	Mover	Mover	Distance moved	Distance moved	Distance moved
	(1)	(2)	(3)	(4)	(5)	(6)
Q_5(Casualty Rate)$_c$ ×	0.504***	0.741***	0.699***	743.769***	612.308***	652.615***
Post-WWI$_t$	(0.178)	(0.177)	(0.186)	(120.164)	(57.320)	(136.667)
Outcome mean	0.240	0.248	0.243	276.5	295.7	329.5
Observations	75,552	55,882	33,222	32,108	23,010	12,296
Adj. R^2	0.370	0.390	0.382	0.467	0.407	0.436
Agriculture	yes			yes		
Manufacturing		yes			yes	
Retail			yes			yes
Mover sample				yes	yes	yes

Note: Difference-in-differences regressions using the linked sample of German born individuals from the 1910 and 1920 full count Census files. Mover is an indicator for whether an individual left their county of residence between 1910-20. Distance is the distance in miles between a person's residence in 1910 and 1920 conditional on having moved between the two Census years. The mover sample only includes individuals who moved county between 1910-20. The agriculture, manufacturing, and retail sub-samples include all individuals who have worked in these industries in 1910. Those three industries account for 57% of all German employment in 1910. All regressions include county fixed effects, a year indicator for 1920, birth year and place of birth fixed effects, as well as baseline controls measured in 1910 and interacted with the 1920 indicator. The baseline controls include the following measures from the 1910 Census: indicators for urban status, eight skill groups, farm status, employment status, marital status, years since entry to the U.S. in bins (0-5, 6-10, 11-15, and 16-20 years), school attendance, labor force participation, and count measures for family size and the number of weeks spent in unemployment in 1909. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Robustness to Match Quality in the Linked Census Data

	Mover	Distance moved	South	Not naturalized	Occupational downgrade	Farmer
	(1)	(2)	(3)	(4)	(5)	(6)
Q_5(Casualty Rate)$_c$ ×	0.516***	547.546***	0.040***	0.061**	0.033**	0.199***
Post-WWI$_t$	(0.159)	(43.863)	(0.011)	(0.029)	(0.016)	(0.022)
Outcome mean	0.213	283.0	0.0255	0.240	0.151	0.057
Observations	216,096	77,192	91,512	91,512	91,512	91,512
Adj. R^2	0.341	0.405	0.103	0.247	0.560	0.110
Mover sample		yes	yes	yes	yes	yes

Note: Difference-in-differences regressions using the linked sample of German born individuals from the 1910 and 1920 full count Census files. The sample only uses high quality matches from the linking procedure, which are those who have received a unique match based on name, birth place, and year of birth. Mover is an indicator for whether an individual left their county of residence between 1910-20. All other outcomes are indicators in 1920 for living in the South, having obtained citizen status, having experienced an occupational downgrade (i.e. from high- to semi-/low-skilled, or from semi to low-skilled jobs using the 1950 occupational definitions of the U.S. Census Bureau), and for working in agriculture in 1920 when the individual's occupation in 1910 was not agriculture. The mover sample only includes individuals who moved county between 1910-20. All regressions include county fixed effects, a year indicator for 1920, birth year and place of birth fixed effects, as well as baseline controls measured in 1910 and interacted with the 1920 indicator. The baseline controls include the following measures from the 1910 Census: indicators for urban status, eight skill groups, farm status, employment status, marital status, years since entry to the U.S. in bins (0-5, 6-10, 11-15, and 16-20 years), school attendance, labor force participation, and count measures for family size and the number of weeks spent in unemployment in 1909. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: OLS Results for German Outflows and Outcomes in Manufacturing

	(1) ln(firm size)	(2) ln(firms)	(3) ln(labor)	(4) ln(output per firm)
WWI German Outflow$_c\times$	-0.040***	-0.026**	-0.062***	-0.019**
Post-WWI$_t$	(0.013)	(0.013)	(0.017)	(0.009)
Observations	10,474	10,474	10,474	10,474
Counties	2,258	2,258	2,258	2,258
Adj. R^2	0.844	0.885	0.895	0.736

Note: OLS regressions of different manufacturing outcomes on the % outflow of German population (measured from 1910-20) interacted with a post-WWI indicator. Outcomes are the natural logs of the average size of manufacturing firms (i.e. workers per firm), number of manufacturing establishments, number of manufacturing workers, and manufacturing output per firm. The sample period is 1900-40. All regressions include county and year fixed effects. Controls contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Manufacturing Outcomes in Counties that received Germans

	(1) ln(mfg wage)	(2) ln(firm size)	(3) ln(firms)	(4) ln(labor)	(5) ln(output per firm)
WWI German Inflow$_c\times$	0.025	0.016	0.097	0.105	0.027
Post-WWI$_t$	(0.021)	(0.075)	(0.060)	(0.100)	(0.055)
Observations	10,388	10,388	10,388	10,388	10,388
Counties	2,223	2,223	2,223	2,223	2,223
Adj. R^2	0.880	0.845	0.898	0.900	0.666

Note: OLS regressions of different manufacturing outcomes on the % inflow of German population (measured from 1910-20) interacted with a post-WWI indicator. An inflow is positive for counties that saw an increase in their German population share between 1910 and 1920 and is zero otherwise. Outcomes are the natural logs of the average wage in manufacturing, average size of manufacturing firms (i.e. workers per firm), number of manufacturing establishments, number of manufacturing workers, and manufacturing output per firm. The sample period is 1900-40. All regressions include county and year fixed effects. Controls contain the WWI draft rate, population size, and pre-war measures of the share of Germans, male-to-female ratio, share of manufacturing employment, population size, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Placebo Check Using Outflows of Other Immigrant Groups

	Outcome: log manufacturing wage per capita (pre-war mean = 1.905)			
	(1)	(2)	(3)	(4)
WWI German Outflow$_c$	-0.016***			-0.017***
	(0.004)			(0.004)
WWI Swedes Outflow$_c$		-0.009		-0.012
		(0.014)		(0.013)
WWI Italian Outflow$_c$			0.014	0.014
			(0.011)	(0.011)
Observations	10,474	10,474	10,474	10,474
Counties	2,258	2,258	2,258	2,258
Adj. R^2	0.876	0.876	0.876	0.877

Note: Difference-in-differences regressions of the log per capita wage in manufacturing on the % outflow of different immigrant groups (Germans, Swedes, and Italians, measured from 1910-20) interacted with a post-WWI indicator. The sample period is 1900-40. All regressions include county and year fixed effects. If included, controls contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: OLS and IV Results for German Outflows and Log Wages in Manufacturing

	Outcome: log manufacturing wage per capita (pre-war mean = 1.905)			
	Panel a: OLS			
	(1)	(2)	(3)	(4)
WWI German Outflow$_c\times$	-0.011**	-0.016***	-0.072***	-0.071***
Post-WWI$_t$	(0.004)	(0.004)	(0.022)	(0.025)
Observations	10,474	10,474	10,474	10,474
Counties	2,258	2,258	2,258	2,258
Adj. R^2	0.875	0.876	0.911	0.856
Controls		Yes	Yes	Yes
Linear county time trends			Yes	
Quadratic county time trends				Yes
	Panel b: IV			
	(1)	(2)	(3)	(4)
WWI German Outflow$_c\times$	-0.050*	-0.071*	-0.106**	-0.106**
Post-WWI$_t$	(0.029)	(0.036)	(0.052)	(0.052)
Observations	10,367	10,367	10,367	10,367
Counties	2,230	2,230	2,230	2,230
K-P F-stat	29.040	30.853	22.736	22.715
Controls		Yes	Yes	Yes
Linear county time trends			Yes	
Quadratic county time trends				Yes

Note: OLS and IV regressions of the log per capita wage in manufacturing on the % outflow of German population (measured from 1910-20) interacted with a post-WWI indicator. In panel b, the outflow of Germans is instrumented with the WWI casualty rate. The sample period is 1900-40. All regressions include county and year fixed effects. If included, controls contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. This specification excludes time-varying controls for total population and male population as compared to the main specification. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Testing for a Potential Direct Casualty Rate Effect

Outcome: log per capita wage in manufacturing (pre-war mean = 1.905)			
	(1)	(2)	(3)
WWI German Outflow$_c$ × Post-WWI$_t$	-0.016***		-0.016***
	(0.004)		(0.004)
WWI Casualty Rate$_c$ × Post-WWI$_t$		-0.002	0.014
		(0.029)	(0.029)
Observations	10,474	10,474	10,474
Counties	2,258	2,258	2,258
Adj. R^2	0.876	0.876	0.876

Note: County-level regressions of log average manufacturing wages on the % outflow of German-born population from 1910-20, and the World War I casualty rate, both interacted with a post-war indicator. The sample period is 1900-40. All regressions include county and year fixed effects. Controls contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: First Stage Effect of Being in the Top WWI Casualty Quintile on German Outflows

	Outcome: Share of German Population (pre-war mean = 1.815)			
	(1)	(2)	(3)	(4)
Q_5(Casualty Rate)$_c$ ×	0.329***	0.258***	0.241***	0.241***
Post-WWI$_t$	(0.069)	(0.053)	(0.045)	(0.045)
Observations	10,390	10,390	10,390	10,390
Counties	2,253	2,253	2,230	2,230
Adj. R^2	0.527	0.666	0.895	0.830
Oster's δ	1.053	1.180	1.316	1.314
Controls		Yes	Yes	Yes
Linear county time trends			Yes	
Quadratic county time trends				Yes

Note: County-level difference-in-differences regressions of the outflow of Germans (1910-20) in county c in decade t on an indicator for being in the top quintile of the WWI casualty rate distribution interacted with a post-war indicator. All regressions include county and year fixed effects. The sample period is 1900-40. If included, controls contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. The δ statistic by Oster (2019) reports the degree of selection on the unobservables (relative to the observables) that would be required to explain away the top-quintile casualty coefficient. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: OLS, First Stage, and IV Results with Conley Standard Errors

	ln(wages) (OLS) (1)	Post-war German outflow (OLS) (2)	ln(wages) (IV) (3)
WWI German Outflow	-0.016		-0.069
(50km)	[0.004]***		[0.031]**
(100km)	[0.005]***		[0.032]**
(200km)	[0.007]**		[0.034]**
Q_5(Casualty Rate)$_c$ × Post-WWI$_t$		0.259	
(50km)		[0.027]***	
(100km)		[0.032]***	
(200km)		[0.043]***	
Observations	2,203	2,203	2,203

Note: OLS, first stage, and IV results with standard errors adjusted for cross-sectional spatial dependence and panel-specific serial correlation. The spatial kernel uses a linear Bartlett window with spatial correlation cutoff points at 50, 100, and 200km. The cutoffs are displayed on the left below the main explanatory variable for each regression. All regressions include controls (WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate), as well as county and Census year fixed effects for 2,203 counties between 1900 and 1940. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Dynamic OLS and IV Effects

Outcome: log per capita wage in manufacturing (pre-war mean = 1.905)		
	(OLS)	(IV)
WWI German Outflow (Effect in 1920)	-0.035***	-0.121***
	(0.005)	(0.040)
WWI German Outflow (Effect in 1930)	-0.011**	-0.041
	(0.005)	(0.046)
WWI German Outflow (Effect in 1940)	-0.001	-0.018
	(0.006)	(0.051)

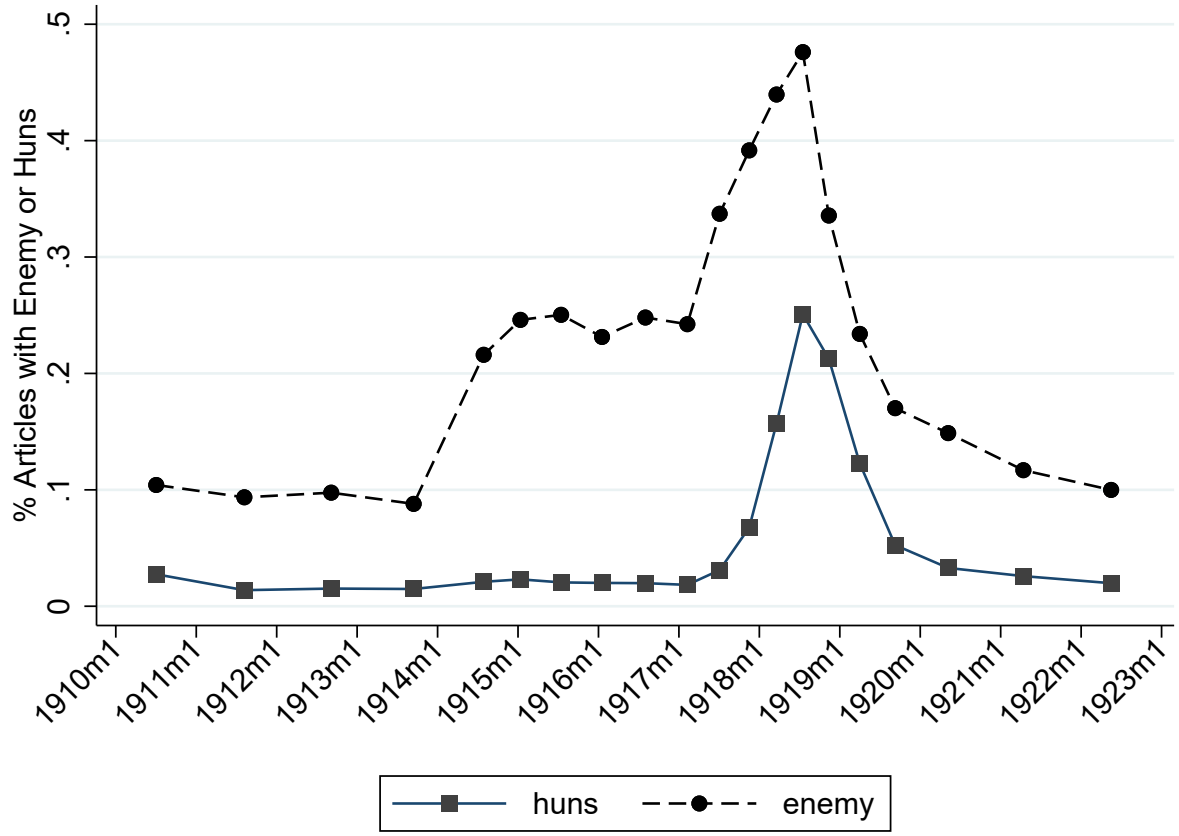
Note: OLS and instrumental variables regressions of the log per capita wage in manufacturing on the % outflow of German population (measured from 1910-20) interacted with a post-WWI indicator. The outflow of Germans is instrumented with the WWI casualty rate in the IV regressions. The sample period is 1900-40. Each cell is a separate regression with the pre-treatment periods being 1900 and 1910, and the post-treatment period being 1920 (first row), 1930 (second row), and 1940 (third row). All regressions include county and year fixed effects as well as controls which contain the WWI draft rate and pre-war measures of population size, share of Germans, male-to-female ratio, share of manufacturing employment, and urbanization rate, all of which are measured in 1910 and interacted with a post-treatment indicator. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: American's Economic Outcomes and Robustness to Link Quality

	Panel a: full sample				
	(1)	(2)	(3)	(4)	(5)
	Mover	Occupational downgrade	Semi-Skilled	Manufacturing	log occ. score
WWI German Outflow$_c$×	-0.018***	0.001*	-0.006***	-0.001	-0.038***
Post-WWI$_t$	(0.006)	(0.001)	(0.001)	(0.001)	(0.005)
Outcome mean	0.169	0.0418	0.137	0.0558	0.885
Observations	3,893,738	3,893,738	3,893,738	3,893,738	3,893,688
Adj. R^2	0.220	0.463	0.515	0.245	0.622
	Panel b: stayers				
		(2)	(3)	(4)	(5)
		Occupational downgrade	Semi-Skilled	Manufacturing	log occ. score
WWI German Outflow$_c$×		0.001**	-0.004***	-0.000	-0.040***
Post-WWI$_t$		(0.001)	(0.002)	(0.001)	(0.006)
Outcome mean		0.037	0.120	0.047	0.877
Observations		2,579,210	2,579,210	2,579,210	2,579,210
Adj. R^2		0.433	0.556	0.298	0.642
	Panel c: stayers working in manufacturing				
		(2)	(3)	(4)	(5)
		Occupational downgrade	Semi-Skilled	Manufacturing	log occ. score
WWI German Outflow$_c$×		0.001	-0.001	-0.010	-0.035**
Post-WWI$_t$		(0.004)	(0.005)	(0.007)	(0.016)
Outcome mean		0.149	0.515	0.681	1.221
Observations		99,552	99,552	99,552	99,552
Adj. R^2		0.357	0.551	0.482	0.596

Note: Difference-in-differences regressions of individuals' outcomes on the percentage decline in the German population in their 1910 county of residence from 1910 to 1920 interacted with a post-war indicator. The sample is a two period panel and contains white American-born individuals who were linked from the 1910 to 1920 full count Census files. Linked individuals were kept if they were uniquely matched and a birth year difference of at most plus/minus 1 years. Mover is an indicator for whether an individual left their county of residence between 1910-20. Occupational downgrade is an indicator for having experienced an occupational downgrade from 1910 to 1920 (i.e. from high- to semi-/low-skilled, or from semi to low-skilled jobs using the 1950 occupational definitions of the U.S. Census Bureau). Semi-skilled is an indicator for whether an individual worked in a semi-skilled occupation in 1920. Manufacturing is an indicator for working in the manufacturing sector in 1920, and log occ. score is the natural logarithm of one plus an individual's occupational income score in 1920. All regressions include county fixed effects, a year indicator for 1920, birth year and place of birth fixed effects, as well as baseline controls measured in 1910 and interacted with the 1920 indicator. The baseline controls include the following measures from the 1910 Census: indicators for urban status, eight skill groups, farm status, employment status, marital status, years since entry to the U.S. in bins (0-5, 6-10, 11-15, and 16-20 years), school attendance, labor force participation, and count measures for family size and the number of weeks spent in unemployment in 1909. The stayer sample consists of individuals who lived in the same county in both 1910 and 1920. The sample of stayers in panel c consists of individuals who were employed in manufacturing in 1910. County fixed effects are absorbed by the individual fixed effects in the stayer samples. Standard errors are clustered at the county-level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

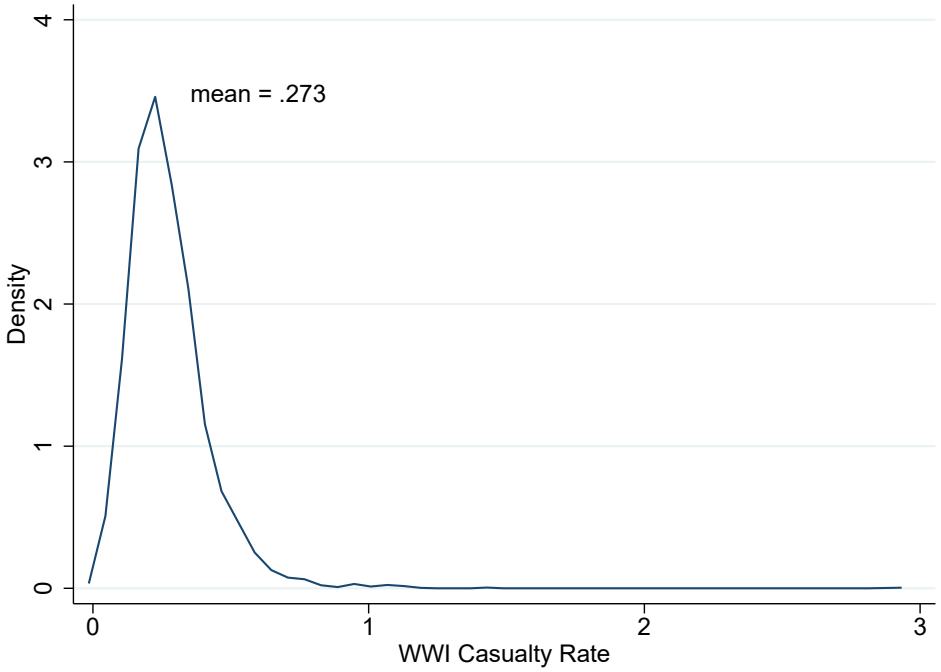
Figure A1: Share of Newspaper Articles on Germans Mentioning the Words *Enemy* or *Huns*



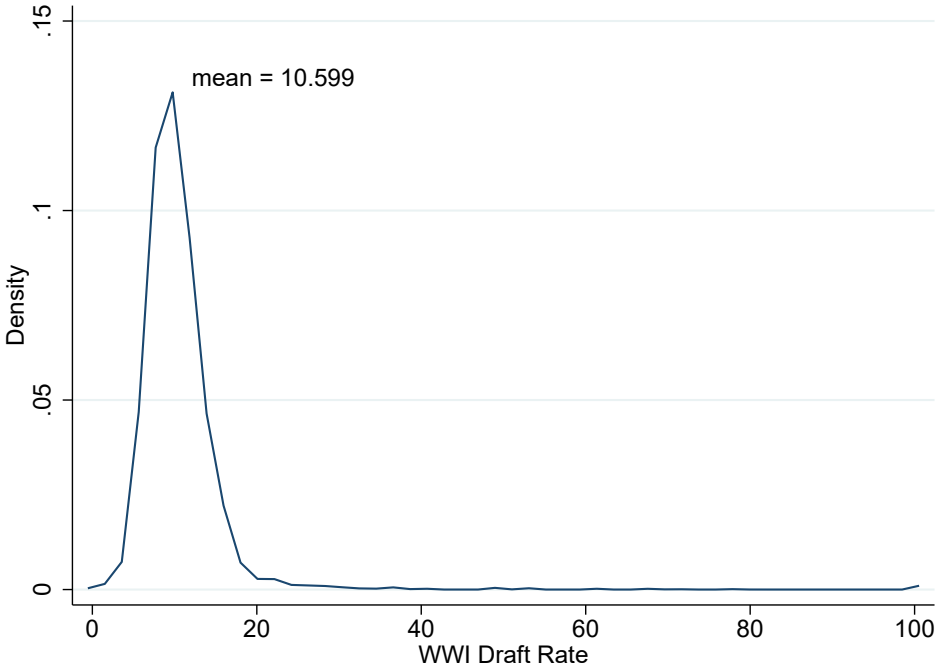
Note: Binned scatter plots with connecting lines for data on newspaper articles per month and year mentioning Germans as *enemy* or *huns*. The density of dots represents the frequency of publishing in a given time interval. The figure shows how Germans are referred to as *enemy* (enemy, enemies, foe) from the start of the war, however, this spikes together with the use of the derogatory word *huns* once the U.S. enter the war in the first half of 1917. The denominator of the share is the number of all articles mentioning Germans in any context.

Figure A2: Casualty and Draft Rate Density Plots

(a) Casualty Rate Kernel Density



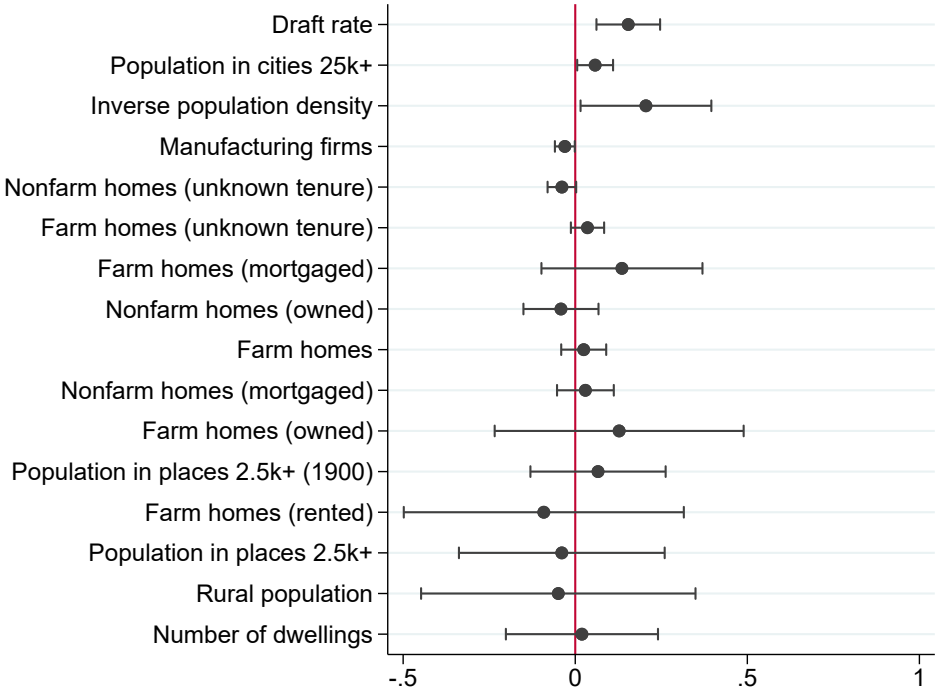
(b) Draft Rate Kernel Density



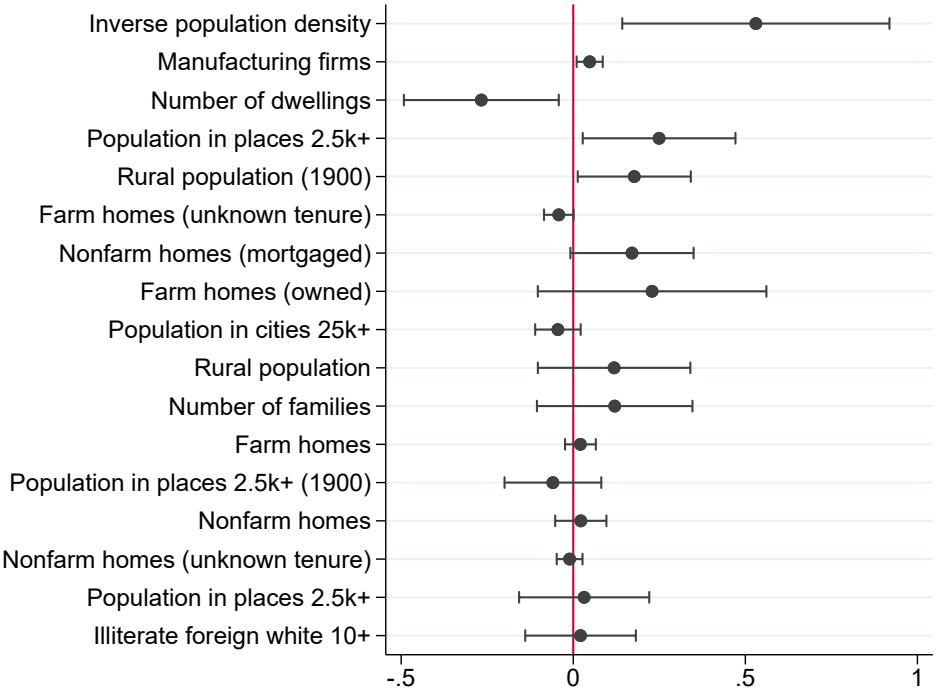
Note: Kernel density plots of the distributions of World War I casualty and draft rates at the county level relative to the voting age male population in each county in 1910.

Figure A3: LASSO Selection of 1910 Variables to Predict Draft and Casualty Rates

(a) WWI Casualty Rate Determinants

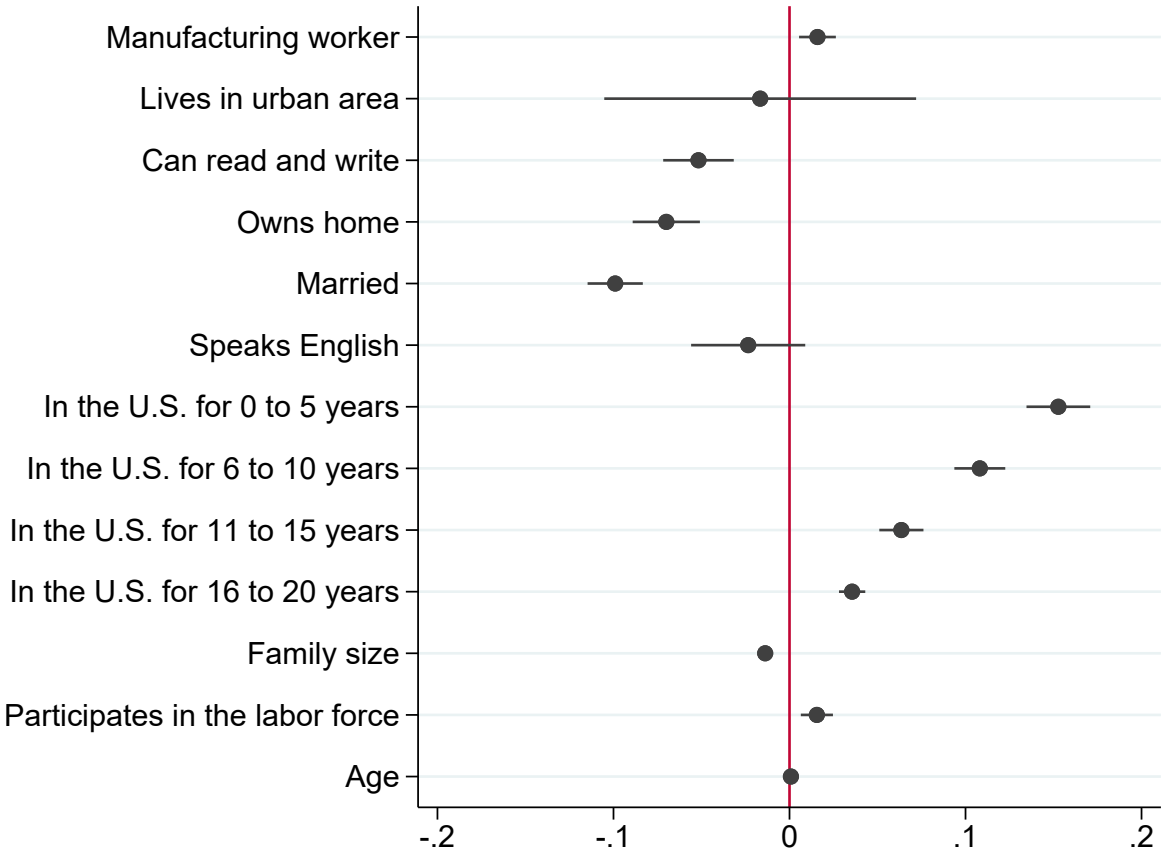


(b) WWI Draft Rate Determinants



Note: Panels (a) and (b) show the selected coefficients from a cross-sectional LASSO regression of the WWI casualty and draft rate, respectively, on all variables related to population, manufacturing, and agricultural characteristics in the 1910 county-level Census. All explanatory variables are expressed in per capita terms and standardized to have mean zero and variance one, i.e. each coefficient estimates the impact on the outcome for a standard deviation increase in the explanatory variable. Coefficients are ranked by their t statistics from high to low. Each LASSO regression also included state fixed effects.

Figure A4: Characteristics of Movers in 1910



Note: Individual baseline characteristics in 1910 that predict moving county from 1910 to 1920 in the sample of linked German born individuals. The regression of the mover indicator on the individual characteristics includes county fixed effects and standard errors are clustered at the county level and shown as 95% error bars. Family size is the number of members in the individual's household. All explanatory variables are binary except for age and family size.

Data Appendix

Census data (county): The county-level Census files were taken from the ICPSR data by Haines (2010). We supplement these data with the information on WWI casualties and enlistments described below, as well as with information on manufacturing which was generously provided by Mike Matheis to fill missing information in the ICPSR data. The Census data provide variables which we use to construct the main controls. These include pre-war county characteristics such as the average population share of Germans, population size, male-to-female ratio, share of employment in manufacturing, and urbanization rate. Monetary values are deflated to the base year of 1910.[52] Manufacturing data was kindly shared by Matheis (2016).

Census data (individual): To link individuals of German ancestry or who were born in Germany (or the German lands) from 1910 to 1920, we follow the procedure by Abramitzky et al. (2014). To identify Germans and Americans of German ancestry, we first keep individuals in the 1910 and 1920 Census files who state that they and/or either of their parents were born in Germany. In the 1910 file, we keep individuals aged 15 to 60 and in 1920 we keep individuals aged 25 to 70. We then matched strictly on first and last name, and place of birth. All unique matches were kept. For cases that did not have a unique match in the 1920 Census, we resolved ties by selecting the match which was the closest in terms of birth year in an interval of plus/minus 3 years. If this did not resolve the tie, we kept the match in which both parents were reported to have the same place of birth in both Census years. Observations were discarded if they still did not have a unique match along the birth year and parental birth place dimensions. We also dropped a small number of individuals who switched race between Census years.

Newspaper data: The newspaper data was downloaded from Chronicling America.[53] We downloaded all newspaper articles from January 1st, 1910, to December 31st, 1922. We then kept all articles mentioning *German(s)* or *Germany* and we excluded all information that was not related to advertisements for beer or local church services and events. For all discrimination measures we used in this paper, the denominator is the number of all articles mentioning Germans in any way (except for the advertisements mentioned). The discrimination measures are constructed as the share of articles mentioning Germans as *enemies* (enemy, enemies, foe), or *huns* (hun, huns, Attila), or that mention the tarring and feathering of Germans or supporters of

[52] The deflator is the historical CPI data provided by the Federal Reserve Bank of Minneapolis at: https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1800-

[53] The data can be accessed at: https://chroniclingamerica.loc.gov/

the German Empire during the years of U.S. involvement in the war. We geo-coded all newspaper outlets to their respective states and counties. If a county was not home to a newspaper outlet, we assigned the discrimination measures of the nearest county with a newspaper and weighed the observation down by the inverse linear distance to that county in regressions.

Casualty data: We digitized the WWI casualty data from Haulsee, Howe and Doyle (1920) for the Army and Washington (1920) for the Navy. Both sources cover ca. 80,000 of the 110,000 total U.S. war deaths. The Army published residence information of the fallen soldiers together with their full name, rank, and cause of death. An example is provided in figure A5. The Navy published the residence information of a soldier's next of kin. In most cases this would be the spouse or their parents. Using this information, we geo-coded all residences to the county level using the 1910 county border definitions. The corresponding county FIPS codes then allowed us to link the casualty information with the aggregate county level information from the Census as well as with the individual-level Census data.

The most common cause of death were directly related to combat. 42.82% of soldiers were killed in action, 35.03% died of disease, another 17.11% of wounds or injuries, and 5.04% were killed in accidents. Roberts and Burda (2018) provide an analysis of the correlates of WWI casualties and socioeconomic characteristics at the county level. They show that Northern counties with higher war mortality rates were predominantly rural and had higher illiteracy rates, while in the South this relation was reversed where more urban counties saw higher casualty rates. The proportion of African American men had no effect on mortality rates.

Draft data: We also digitized information on the number of enlisted men by county from table 20 in the Final Report of the Provost Marshal General (Crowder, 1920). The report lists the total number of soldiers called for service, those who were eventually inducted, accepted, and rejected, and those whose draft was ultimately canceled, for instance because they received an occupational deferment. The numbers are reported for each local draft board with most boards being responsible for a single county. For larger counties and cities, multiple boards were responsible for the drafting such as in Wilmington, Delaware, as shown in panel (b) of figure A5. Multiple boards for a single county were aggregated into one observation and given the FIPS code of the corresponding county. If a board specifically served a city, e.g. Birmingham, Alabama, then we aggregated the information as before and assigned the FIPS code of the city's county (e.g. Jefferson for Birmingham, AL).

Figure A5: WWI Casualty and Draft Lists

(a) Casualty List

CALIFORNIA

KILLED IN ACTION

Lieutenant Colonels

CRAIG, J. M., San Francisco.
HOLLIDAY, William E., Santa Monica.

Majors

BEASLEY, Shadworth O., San Francisco.
MILLER, Oscar F., Los Angeles.
SMYTH, Roy Melvin, Alameda.

Captains

MacPHERSON, Harry H., San Francisco.
SMITH, Clarence F., Los Angeles.
VARNEY, Kit Roberts, San Francisco.

Lieutenants

BABCOCK, Robert C., San Francisco.
BARRY, David M., Santa Barbara.
BEACH, Egbert William, Piedmont.
BELL, Kenneth C., Pasadena.
BOYER, Herbert, San Francisco.
CROWELL, Fleming M., Los Angeles.
DAVIDSON, Gilford C., San Francisco.
ELAM, Edwin M., Berkeley.
FULTON, Hugh, Oakland.
GARD, Frank J., Glendora.
HAMMER, Earl M., San Francisco.
HANLY, William J., Oakland.
HARDING, Stacy Ludden, Antioch.
HARTER, Clifford C., Santa Barbara.
HITCHCOCK, Roger W., Los Angeles.

Sergeants—Continued

JONES, James M., San Francisco.
LAKE, Thomas J., Los Angeles.
LARSEN, Peter W., San Miguel.
LUY, Richard L., San Gabriel.
MacPHERSON, William M., Madera.
MANDEVILLE, John L., San Diego.
McCAUSLAND, Clinton, Ripon.
McFALL, Hope, Manteca.
McKINNON, Elwyn Charles, Los Angeles.
McMILLAN, Laning R., Corona.
MESTROVITCH, James L., Fresno.
PATTERSON, Frederick H., Los Angeles.
PETERSON, Peter N., Neroman.
POWELL, Ballard B., Sacramento.
ROBBINS, George W., Los Angeles.
ROSS, George W., Oakland.
ROSS, Karl E., Stockton.
SHEEHY, Norman R., Los Angeles.
SIMMONS, Melvin K., Fairfield.
STEVENS, Edward J., San Francisco.
SULLIVAN, John Q., Lost Hills.
SWEETNAM, John M., Sebastopol.
THOMPSON, Charles H. ? ? ?
WALTERS, Charles, San Diego.
WHITE, Thomas R., Sacramento.
WHITNEY, William E., Oakland.
WILLIAMS, Charles V. G., Chino.

Corporals

ADAMS, Herbert H., Oakdale.
AGGELER, Jerrold J., Stockton.
ALTMAN, Henry, San Francisco.
BAHNEY, John W., Sacramento.
BALLARD, Blackburn W., Colura.
BATCHELOR, Louis W., San Fran-

Corporals—Continued

MILLER, Harry A., Oakland.
MORRIS, Fred L., Los Angeles.
NEEDHAM, Clyde W., Lodi.
NUNES, Alfred, Centerville.
PALMERLEE, Chester C., Long Beach.
PASSERINI, Frank, San Francisco.
PEDRIOLI, Louis, Modesto.
PERRY, William S., Jr., Berkeley.
PETERSON, Arthur L., Long Beach.
POORE, Raymond, Pasadena.
RICHESON, Franklin Carter, Dinuba.
ROBERTS, Harold William, San Francisco.
ROBINSON, Glen H., Pescadero.
RUBIDOUX, Mack J., Riverside.
SAXEY, Harry, Willow Creek.
SCHMALZ, John W., San Francisco.
SCHNEIDER, Harry N., Morgan Hill.
SHANKLAND, Claude G., Bakersfield.
SIEVERS, Maxwell H., Salinas.
SOUZA, Manuel, Jr., Cambria.
SOWELL, Vernon L., Lemoore.
SPARGO, John, San Francisco.
STAPLES, Guy W., Linden.
SWEET, Ora A., Crockett.
TROMBLY, Charles H., Pasadena.
VINTHER, Claudius F., Berkeley.
WALL, Earnest W., Sacramento.
WEYLANDT, Lester L., Peters.
WIENS, Gary, Los Angeles.
WILKINS, James H., Jr., San Rafael.
WILSON, Robert H., Los Angeles.
WOODWARD, Earl, Lathrop.

(b) Draft List

TABLE 20.—*Calls, inductions, acceptances, and rejections, by local boards*—Continued.

DELAWARE.

Local board.	Total called.	Total inducted.	Total accepted.	Total rejected.	Total rejected. Cancellation of draft.
Kent	434	417	386	30	1
New Castle	946	1,002	919	78	5
Sussex	670	682	649	31	2
WILMINGTON No. 1	947	1,001	945	48	8
WILMINGTON No. 2	789	819	755	55	9
WILMINGTON No. 3	648	688	635	48	5
WILMINGTON No. 4	825	918	810	103	5
Total	5,259	5,527	5,099	393	35

DISTRICT OF COLUMBIA.

No. 1	1,170	1,189	1,011	52	126
No. 2	1,495	1,466	1,224	79	163
No. 3	1,127	1,154	1,069	64	21
No. 4	1,037	1,008	925	66	17
No. 5	891	906	847	36	23
No. 6	1,047	1,008	935	43	30
No. 7	709	704	648	42	14
No. 8	1,346	1,335	1,250	73	22
No. 9	919	903	857	27	19
No. 10	989	878	827	23	28
No. 11	454	478	441	26	11
Total	11,184	11,029	10,024	531	474

Note: Panel (a) shows an example of the casualty records from Haulsee et al. (1920), page 139, for the state of California. Casualties are ordered by cause of death, rank, and alphabet. Soldier-level information includes state, rank, first, middle, and surnames, as well as the city or county of residence. Causes of death are killed in action (42.82%), disease (35.03%), wounds and injuries (17.11%), and accidents (5.04%). Army casualties total almost 80,000 of the overall 110,000 war deaths sustained by the United States during World War I. Panel (b) displays the draft records for each local draft board and state from Crowder (1920). The records include information on the total number of examined, inducted, accepted, and rejected men, as well as those whose draft was canceled, e.g. due to a deferment or change in health status that prevented deployment.