Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I [*]

Vasiliki Fouka[†]

December 2016

Abstract

Can forced assimilation policies successfully integrate immigrant groups? This paper examines how a specific assimilation policy – language restrictions in elementary school – affects integration and identification with the host country later in life. After World War I, several US states barred the German language from their schools. I find that affected individuals were less likely to volunteer in WWII and more likely to marry within their ethnic group and to choose decidedly German names for their offspring. Rather than facilitating the assimilation of immigrant children, the policy instigated a backlash, heightening the sense of cultural identity among the minority.

JEL Codes: J15, Z13, N32, I28.

Keywords: Assimilation, educational and language policies, ethnicity, cultural transmission.

[†]Email: vfouka@stanford.edu

1 Introduction

From France's "burkha ban" to the politics of bilingual education in California, societies around the world grapple with the challenge of integrating ethnic minorities. Theories of nation building (Alesina and Reich, 2013) assume that policies such as imposing a national language or otherwise repressing minority culture will lead to more homogeneity. At the same time, one strand of literature has shown theoretically that identity may be strengthened in the face of policies aimed at integration (Bisin and Verdier, 2000, 2001; Bisin et al., 2011). What is unclear is when assimilation can work in practice, and what drives the risk of a backlash.

In this paper I examine the long-term effects of a particular assimilation policy: the prohibition of German in US schools after World War I. When the United States joined the war, German speakers were increasingly treated with suspicion. Before 1917, bilingual education was common in many states that were home to German immigrants — the country's largest group of migrants. Following the war, a number of states banned German as a language of instruction. I examine whether forced language integration affected the ethnic identity and actions of immigrant children. Did the ban on German lead to more assimilation, or did it contribute to a cultural backlash and greater isolation from the mainstream of American culture? Using linked census records and World War II enlistment data, I examine several outcomes for German-Americans affected by language restrictions: (i) the ethnic distinctiveness of the first names chosen for their offspring, (ii) their intermarriage rates, and (iii) their decision to volunteer for the US Army during World War II.

I exploit both within–cohort variation (comparing states with and without a German ban) and within–state variation (comparing cohorts at school with older cohorts) in a difference–in–differences model. I find a strong backlash effect for the children of German immigrants and this effect is consistent across outcomes and specifications. Treated cohorts in this group were 6–9 percentage points more likely to marry endogamously (i.e. within their ethnic group) and 7–9 percentage points less likely to volunteer in WWII. They also chose more distinctively German names for their children, with the estimated effect being equivalent to switching from a name like Ted to a name like Adelbert.

Next, I examine the mechanisms behind this reaction. The estimated backlash becomes weaker (or goes in the opposite direction) for Germans born to mixed couples. This establishes a link between the strength of the parents' ethnic identity and their offsprings' reaction to policies affecting ethnic schooling. In line with models of cultural transmission (e.g., Bisin and Verdier, 2001), the backlash is greater in counties with a smaller share of German population. This is consistent with a cultural transmission

mechanism in which parental and peer socialization are substitutes: In places where Germans constitute a smaller minority, parents try harder to shape each child's sense of ethnicity because they cannot reasonably expect that children will be socialized in their ethnic culture through peer interaction alone. The extent of the backlash was higher also in counties with a greater share of Lutherans, the predominantly German church that emphasized parochial schooling in the German language. The implication is that communities with a greater initial sense of ethnic identity react more adversely to assimilation policies.

My findings imply that linguistic immersion through the prohibition of German has no clear assimilation effect on average. Instead, and across all outcomes, a language ban leads to a robust increase in the spread between individuals of uniform and mixed German ancestry. Furthermore, the language ban has, if anything, a positive effect on years of schooling and is thus unlikely to reduce assimilation through its negative effect on education. There is, however, weak evidence that a strengthening of ethnic identity entails a penalty for individuals who become more German. German-Americans with two German parents affected by language laws earned less in the labor market. Given that schooling outcomes become better for exposed cohorts, such a drop in earnings is unlikely to be due to lower quantity or quality of education as a result of linguistic immersion. It is, however, consistent with research emphasizing the economic payoffs of assimilation (Biavaschi, Giulietti and Siddique, 2013).

The empirical setting I examine offers a number of advantages. The timing of the legislation was plausibly exogenous, as the anti-German sentiment that motivated it was not pre-existing but rather spurred by the war (Higham, 1998). Historical sources describe language campaigns of equal intensity and resistance on the part of German-Americans in most Midwestern states, with the final outcome often depending on the character of the local commissioners of education (Beck, 1965; Rippley, 1981).[1] To deal with potential unobservable confounders, I focus on the state border of four comparable states — Indiana, Ohio, Michigan, and Kentucky — and create a linked data set of individuals who lived there at the time legislation was enacted in Ohio and Indiana. Apart from increasing internal validity, this design allows me to observe long-run assimilation outcomes of German-Americans and to examine how those outcomes vary by the ethnic composition of their home town. I can thus identify conditions, such as the size and character of the minority group, that lead to a more pronounced identity backlash. Finally, the case study of German-Americans yields an interesting measure

---

[1] Lleras-Muney and Shertzer (2012) find that the only significant predictors of the passage of English-only laws were the share of immigrants and the length of compulsory education in a state.

of ethnic identification: volunteering for service in the US Army during World War II. This is a unique historical setup in which immigrants are called upon to take sides between their host country and their country of origin.

A number of theoretical studies suggest that assimilation policies can lead to a backlash of ethnic or religious identity. Bisin et al. (2011) present a mechanism for the persistence of oppositional minorities.[2] In their model, oppositional types intensify their identification with the minority culture in response to attempts at desegregation or discrimination by mainstream society. Similarly, Carvalho (2013) predicts that bans on veiling worn by Muslim women can *increase* religiosity. Carvalho and Koyama (2016) show that, when education serves as a means of transmission of the majority culture, minorities can underinvest in education as a form of cultural resistance. This paper is the first to provide empirical evidence that a backlash of identity in response to assimilation policies is more than a theoretical possibility.

My research also contributes to the literature on the economics of identity (Akerlof and Kranton, 2000). Ethnic, religious and other social identities have been shown to have a significant impact on preferences and economic behavior (Hoff and Pandey, 2006; Benjamin, Choi and Strickland, 2010; Benjamin, Choi and Fisher, 2013), but evidence on the determinants of identity formation is generally not causal in nature (Constant, Gataullina and Zimmermann, 2009; Battu and Zenou, 2010; Manning and Roy, 2010; Bisin et al., 2013). I provide evidence on a specific mechanism through which ethnic identity can be influenced: language in school and its interaction with parental socialization. In this regard, the paper most closely related to mine is Clots-Figueras and Masella (2013).[3]

This study also relates to a broad literature on immigrant assimilation. Much of this research has focused on economic assimilation and the gap between native and immigrant earnings.[4] In addition, several papers construct measures of the speed of assimilation by looking at political (Shertzer, 2013) or cultural outcome variables (Aleksynska and Algan, 2010), such as first names (Arai et al., 2009; Abramitzky, Boustan and Eriksson, 2016) or self-reported national identity (Manning and Roy, 2010). Dávila

---

[2]See also Bisin and Verdier (2000), Bisin and Verdier (2001) and Bisin, Topa and Verdier (2004).

[3]These authors find that instruction in Catalan, which was re-introduced in the schools of Catalonia in Spain after the Franco era, led to a stronger identification with the cause of Catalan independence and to a greater tendency to vote for Catalanist parties. My research addresses the reverse setup. Rather than focus on the effects of imposing a national language on the majority (as Catalan was for Catalonia), I examine the case of prohibiting a minority language.

[4]See Borjas (1985), LaLonde and Topel (1991), Hatton (1997), Minns (2000), Card (2005) and Abramitzky, Boustan and Eriksson (2014).

and Mora (2005), Neeraj, Kaestner and Reimers (2005), and Gould and Klor (2012) show how discrimination against Muslims in the United States after the 9/11 attacks reduced integration. My study contributes to this literature by identifying the effect of a specific government intervention on assimilation outcomes.

More broadly, this paper relates to a rich literature in history and the social sciences that examines the effects of education on national identity. There are many studies documenting how education and the content of the school curriculum have been used to shape preferences, homogenize societies, and "manufacture" nations (Dewey, 1916; Freire, 1970; Weber, 1976; Colley, 1992). More recently in the economics literature, Cantoni et al. (forthcoming) show how a new school curriculum in China had a measurable effect on the political attitudes of students. My study focuses more on the medium than on the content of education, but its results suggest that the purpose of assimilationist educational policies is seldom entirely achieved. The study of Friedman et al. (2011) in Kenya points in a similar direction. They find that more education in the context of a nationalist curriculum led to political alienation for school girls, and, if anything, heightened tribal identities instead of fostering national unity. Similarly, in the case of Zimbabwe, Croke et al. (forthcoming) find that more education, when provided by an authoritarian regime, decreases political participation.

The paper proceeds as follows. Section 2 discusses the historical background of German language schooling and the language restrictions imposed after WWI. Section 3 describes my data sources. Section 4 is devoted to the empirical analysis. I show that the prohibition of German in school created a backlash of ethnic identity among Americans born to German parents, as measured by ethnic name choices, endogamy rates and volunteering in World War II, and I check the sensitivity of my results along a number of dimensions. Section 5 shows how this backlash effect weakens among children of mixed couples, assesses how the response to legislation varies by a community's ethnic composition and strength of ethnic identity and examines whether language restrictions affected schooling and other outcomes later in life. Section 6 reviews my findings in the context of recent theory on cultural transmission and identity in economics. Finally, Section 7 concludes.

2   Historical background

This section outlines the history of the German language in US schools until the early 20th century. It also discusses the reasons that led to the restriction of German as a language of instruction during and after World War I.

2.1 Germans in the United States and the German language in schools

Germans were the single largest foreign group that migrated to the post-colonial United States until at least the 1970s. German immigration started in the 17th century, increased after the failed revolutions of 1848, and peaked in the 1890s, when economic migrants replaced political refugees in the arriving immigrant cohorts. Between 1880 and 1920, Germans constituted the largest element among the foreign-born in the United States; in 1900, the first and second generation of Germans together accounted for more than 10% of the total US population (Conzen, 1980).

As the dominant non–English speaking group, Germans established a large network of private (mainly religious) schools, in which the German language was taught and used as a medium of instruction. They also succeeded in introducing German instruction to the public schools of districts with a large German population. In cities such as Cincinnati and Indianapolis, designated German-English schools provided a form of bilingual education that included half-day instruction in German (Schlossman, 1983; Zimmerman, 2002). Such bilingual programs were favored by German parents and supported by school officials as a way of drawing first- and second-generation German children away from private schools, which were perceived to perpetuate exclusive ethnic communities and to endanger the linguistic and cultural homogenizing function of the public school. Some proponents of dual German-English instruction pointed out its assimilating function not just for the children of German immigrants but also for their parents. According to the Milwaukee Association of Collegiate Alumnae: "Foreign mothers, who are busy all day in their homes, have but one opportunity to acquire the language of their adopted country, and that is from their children, who bring English home from the schools" (Schlossman, 1983).

Although there is no comprehensive census of private schools and their instruction practices, individual state census records reveal the prevalence of German in parochial schools prior to World War I. According to the 1917 Minnesota Educational Census, the state counted 308 parochial schools with a total enrollment of 38,853 pupils; more than two thirds of these schools used both German and English as a medium of instruction (Rippley, 1981).[5] Official statistics aside, a number of sources confirm the unofficial use of German by teachers in the classroom as a natural way of introducing first- and second-generation children of German parents to English (Schlossman,

---

[5]In the early 20th century, 35 out of 48 states taught some form of German in school mostly in the form of a foreign language in secondary education (Wüstenbecker, 2007).

[1983](#)). For parochial schools that employed German-born teachers and were located in predominantly German rural communities, this practice was the norm.

Despite this ethnic group's large network of schools, the prevalence of using German and the importance placed by German-Americans on conserving their culture and a sense of *Deutschtum*, by the early 1900s Germans were fairly well assimilated — in both socioeconomic and cultural terms. In the words of Higham (1998), "public opinion had come to accept the Germans as one of the most assimilable and reputable of immigrant groups. Repeatedly, older Americans praised them as law-abiding, speedily assimilated, and strongly patriotic."

2.2 WWI, anti-Germanism, and language restrictions

The outbreak of the First World War made the large German community the focus of American patriotic reaction. The growing anti-Germanism of the early war years, which was further agitated by the insistence of the German-American press on strict American neutrality, found its expression in a series of both spontaneous and organized acts of harrasment and persecution once the United States entered the war in 1917. Numerous German-Americans were arrested as spies or forced to demonstrate their loyalty by buying liberty bonds under the threat of vandalism or tarring and feathering. The hanging of Robert Prager in Collinsville, Illinois, was the most well known in a series of lynching attacks against German-Americans (Luebke, 1974). Berlin, Michigan, was renamed to Marne in honor of the American soldiers who fought in the Second Battle of Marne. Hamburgers became "liberty steaks"[6] and sauerkraut consumption fell by 75% in the period 1914–1918 (*New York Times*, 25 April 1918). Moser (2012) shows that the number of German-language operas staged at the New York Metropolitan Opera fell dramatically during the war years.

The German language also came under attack. At the federal level, the 1917 Trading With The Enemy Act and also the Espionage Act required all foreign language publications to translate into English any news referring or related to the war. At the state and local level, various restrictions were placed on the use of German. The state of Iowa prohibited, among other things, the use of German over the telephone. Iowa state governor William Lloyd Harding stated in the *New York Times* in June 1918 that "English should and must be the only medium of instruction in public, private, denominational and other similar schools. Conversation in public places, on trains, and

---

[6]An interesting parallel is the renaming of french fries to "freedom fries", after France opposed the US invasion of Iraq in 2003 (Michaels and Zhi, 2010).

over the telephone should be in the English language. Let those who cannot speak or understand the English language conduct their religious worship in their home" (Baron, 1990).

This political climate encouraged support for language restrictions in the schools. Since the war's outbreak, nationalist organizations had propagandized against the instruction of German. A 1915 pamphlet of the American Defense League, one of the largest nationalist political groups of the time, reads as follows: "Any language which produces a people of ruthless conquistadores [*sic*] such as now exists in Germany, is not fit to teach clean and pure American boys and girls." This propaganda merged with a pre-existing nativist movement that originated in the 19th century, but had strengthened in the early 1900s in response to the unprecedented flow of immigration to the United States (Kazal, 2004). During and after the war years, these attitudes were enshrined in legislation restricting foreign languages in a number of states.

Until that time, the legislative framework regulating the language of instruction in schools was heterogeneous. By 1914, 22 states had some sort of provision requiring the use of English. As documented in Edwards (1923), English had been the language of instruction in the public or common schools of some states since the end of the 19th century; in other states, such as New York and Rhode Island, English was recognized later on as the official school language to meet requirements of the compulsory schooling law. In many states, however, provisions regarding the use of foreign languages were permissive; for example, Colorado permitted German or Spanish to be taught when requested by the parents of 20 or more pupils (Luebke, 1999). The state of Ohio in 1903 allowed for German instruction in the public schools upon the demand of "75 freeholders resident in the district", making such instruction optional "and auxiliary to the English language" in 1913 (Leibowitz, 1971).

World War I marks a clear break in the pre-existing trends of English language legislation; in the period 1917–1923, there were 23 states that prohibited the use of foreign languages as a medium of instruction or as a separate subject in elementary grades (Knowlton Flanders, 1925). Though not always explicitly targeted against German, these laws are generally viewed by legal scholars as resulting from anti-German sentiment during the war years (Van Alstyne, 1990; Bennett Woodhouse, 1992). Their main difference from previous legislation is that they applied to all schools — whether public, private, or parochial.[7] Since English was already the main (and most often the only) language of instruction in public schools, the laws were mainly aimed at private

---

[7]Similar in spirit was the 1889 Bennett Law of Wisconsin, which was fiercely opposed by the state's Lutheran and Catholic population and repealed in 1891.

schools and at German-Americans, the ethnic group with the largest and oldest system of private schools in the country.

In 1923, the US Supreme Court repealed the 1919 Nebraska law — and with it all legislation that restricted foreign-language education in the private schools — as a violation of the Fourteenth Amendment. Despite this ruling, most parochial schools did not re-introduce instruction in German and the number of high school students studying German, which dropped precipitously during the war years, never returned to its pre-war levels (Schlossman, 1983; Wüstenbecker, 2007).

3   Data

My analysis focuses on Indiana and Ohio, the only two states that passed legislation targeted specifically against the German language. Both of these states had permissive provisions on language use in schools prior to 1919, and both provided dual language instruction programs in the public schools of their main cities, Indianapolis and Cincinnati (Schlossman, 1983). During the period in question, their neighboring states (Michigan and Kentucky) neither introduced nor had in place any language laws. I first construct a unique data set of individuals living at the border of these states (and of their neighbors) at the time legislation was enacted and then link this data over time to later census years so as to observe choices of first names for children and intermarriage outcomes. Subsequently, to investigate whether exposure to legislation affected the national identity and patriotism of Germans later in life — as proxied by their decision to volunteer or not for service in the Second World War — I use the World War II Army Enlistment Records digitized by the National Archives. I link a subset of this data to the 1930 census in order to obtain information on the ethnic background of enlisted men.

3.1   Laws

Both Indiana and Ohio explicitly singled out German as a language to be prohibited in elementary school grades in 1919.[8] The law in Ohio reads as follows:

That all subjects and branches taught in the elementary schools of the state of

---

[8]The only other state that explicitly prohibited German in its schools was Louisiana in 1918. This prohibition was part of a legislative package known as Act 114, which was enacted as an expedited war measure and also prohibited the use of German in public and over the phone. It was repealed by the US Supreme Court in 1921.

Ohio below the eighth grade shall be taught in the English language only...Provided, that the German language shall not be taught below the eighth grade in any of the elementary schools of this state. (108 Ohio Laws, 614, 1919)

The wording was almost identical in Indiana:

All private and parochial schools...shall be taught in the English language only...provided, that the German language shall not be taught in any such schools within this state. (School Laws of Indiana, 1919)

I combine data on English-only laws with information on the age range of compulsory schooling from Goldin and Katz (2008). Because the legislation I am considering was passed in 1919, cohorts exposed to it were those that should — according to the compulsory schooling law of their respective states — be in school at the time a law was in effect. The compulsory schooling age in the period was 7–16 in Indiana, Michigan and Kentucky and 8–16 in Ohio.

3.2   Indiana and Ohio borders

I use the newly digitized full count of the 1920 census to construct a unique data set of all native-born males in the 1880–1916 birth cohorts who had parents born in Germany and who lived in a county on either side of the border of Indiana and Ohio with Michigan and Kentucky in 1920 — the census year closest to the introduction of these anti-German laws (see Figure 1, where the border counties are shaded).

Focusing attention on state borders is meant to increase the comparability of affected and non affected Germans in dimensions other than language restrictions. Using 1910 county-level data from ICPSR and the Census of Religious Bodies, Table 1 shows that this is not enough to eliminate all pre-treatment differences between states: Indiana and Ohio are more urban, and have a higher share of foreign and German-born population. However, most of these differences are largely attributable to the presence of Cincinnati, the largest urban agglomeration of the region, and become substantially smaller once it is excluded for the sample. Excluding Cincinnati from any subsequent analysis does not affect the results.

Using the procedure and criteria just described, I begin with a data set of 30,987 males observed in 1920. I am interested in how exposure to language restrictions affected the later assimilation outcomes of these individuals. To compile these outcomes, I use the complete-count 1930 and 1940 US censuses to link records over time. Following standard census-linking procedures (Ferrie, 1996; Abramitzky, Boustan and Eriksson, 2014), I start by using the phonetic equivalent of first and last name, the birthplace,

and the year of birth (allowing for a two-year band around the recorded year) to locate an individual in a later census. One of the drawbacks of this procedure is that it yields a large number of records with multiple matches. Discarding these multiple matches results in loss of information. I therefore extend this process by computing the string distance between first and last names in the original data and the target census year. I use the Jaro-Winkler algorithm (Mill, 2012), which yields a measure that takes values from 0 to 1, with 1 implying that two strings are identical. I sum up the Jaro-Winkler measures for first and last name and filter multiple matches by keeping only those with the smallest value in this composite Jaro-Winkler index.

The Jaro-Winkler distance allows for further refinement of the matched data set by providing a way to discard names that are sufficiently different in origin and target census years, and thus getting rid of false positive matches. The higher the value of the JW distance chosen as a threshold, the larger the share of the initially matched data that is discarded. I choose as this threshold the JW value for which the change in the share of matched data dropped is maximized. Intuitively, I increase the JW – and thus the precision of the match – until the point where increasing it further would imply losing too many observations. This procedure (further detailed in the Appendix) leaves me with a total of 14,593 unique matched records in either 1930 or 1940. Table 2 provides summary statistics for the linked data sets. Figure 2 shows the locations in 1920 of all individuals successfully linked in either 1930 or 1940.

First names. I use the names that individuals in my sample choose for their children as a proxy for ethnic identity. Names have an indisputable cultural component and to a great extent reflect the parents' racial, ethnic, and social background and preferences (Lieberson, 2000; Fryer and Levitt, 2004; Head and Mayer, 2008; Cook, Logan and Parman, 2013). As such, the choice of first names for their offspring is indicative of parental tastes and, for immigrants, of assimilation into the host society (Abramitzky, Boustan and Eriksson, 2016). In particular, if cohorts affected by an anti-German law choose to give their offspring names that are less German-sounding and more common among natives, then that would indicate an assimilation effect of language restrictions.[9]

In order to measure a name's ethnic content, I follow Fryer and Levitt (2004) in

---

[9]Algan, Mayer and Thoenig (2013) show that the economic penalty associated with culturally distinctive names is an additional important determinant of parents' naming decisions. In the current setup, there is no clear reason to believe that local labor market conditions faced by children differ depending on their parents having been or not affected by language laws in elementary school. If greater discrimination in states with a language law persisted to the children's generation, this should in fact have led parents to give less and not more German names to their children.

constructing an empirical index of German name distinctiveness, by using census data on first names and ethnic origin. This *German name index* (GNI) captures how much more frequent a name is among the population of German origin compared with the rest of the population. A name found only among German ethnics would have index value 100, whereas a name given to no individuals of German origin would have index value 0.

Table 3 provides an overview of what this index captures in the 1930 5% IPUMS sample. The left panel shows the 10 names with the highest value of the name index that were given to more than 1,000 individuals in 1930; all are distinctively German-sounding. Not all distinctive names are common among Germans, but many of these names, including Herman and Christian, are also on the list of most popular names among German immigrants. The right panel of Table 3 lists the 10 most popular names with a zero GNI value. Names such as Clyde, Russell, and Melvin are characteristically un-German in that they had been given to no German-born individuals in the 1930 IPUMS sample.

Because the GNI is computed relatively to names of foreign-born Germans, many names in my sample of the second generation have a GNI value of 0. I take the logarithm of the GNI to deal with this skewed distribution and to allow for an intuitive interpretation of results in percentage terms.[10] In the main empirical analysis, I will use both the logarithm of the average GNI of all children and the logarithm of the GNI of the first son as outcome variables that proxy for ethnic identity.[11]

Intermarriage. Intermarriage has been characterized as "the final stage of assimilation" (Gordon, 1964). Unlike first names, it is not a pure choice, but a general equilibrium outcome, determined by others' preferences and by the constraints of the marriage market. However, it is arguably a good indicator of immigrant integration in the host country, as it reflects acceptance of the host culture on the part of the immigrans and vice versa. I investigate the extent to which being exposed to restrictive legislation at school affects the probability that second-generation German-Americans end up marrying within their own ethnic group.

How can marriage decisions be affected by the language of instruction in school? The choice of a spouse involves an important preference component (Fisman et al., 2008; Banerjee et al., 2009), and US society has historically been characterized by marriage segregation along racial, religious, and ethnic lines (Pagnini and Morgan,

---

[10]I use log(GNI+x), where x is a small positive number, to avoid loss of data where GNI=0.

[11]Male names continue to be more traditional than female ones even in modern-day Germany (Gerhards, 2005).

1990; Fryer, 2007). Bisin, Topa and Verdier (2004) show theoretically how parents seeking to socialize their children into their culture will marry homogamously, and they demonstrate that US patterns of religious endogamy are in line with this prediction. To the extent that the language of instruction in school affects the ethnic preferences of second-generation immigrants, we can expect to see changes in marriage choices later in life as one response to language restrictions. In particular, if removing German from the curriculum had the effect suggested by proponents of the policy, then English-only instruction should lead to greater assimilation as reflected in higher intermarriage rates. That might happen because, in the first place, children would no longer be indoctrinated "with the German language, customs, and prejudices of the Fatherland ... against the social and religious customs of the American communities in which they claim citizenship."[12] Greater familiarization with the American language and culture, as the Americanization movement aimed to inculcate, would make these children prefer American spouses later in life. Second, to the extent that such Americanization would make these offspring more receptive to social environments other than their closed ethnic communities, the market for marriage partners would contain more non-ethnic members and thus would increase the likelihood of intermarriage.

The earlier US censuses pose some difficulties for determining an individual's ethnic background. In 1940, the question on parental birthplaces was posed to only 5% of the universe. This means that I can observe the ethnic background of the spouse of a native person only in 1930. In this census year, younger cohorts are observed at an age when they have likely not yet completed their marriage spells (ages 18–27). Other than leaving us with a small number of observations, comparison of these cohorts between states with and without a language law should still yield unbiased estimates, though they are not likely to be representative of the general population of German-Americans.[13]

3.3   World War II enlistment records

Data on men who enlisted in the US Army during World War II are from the *Army Serial Number Electronic File, ca. 1938 –1946*. The database is the end product of digitizing the original WWII draft computer punch cards by the National Archives and Records Administration. The complete database comprises of nearly 9 million

---

[12] "The German language school question", *The Outlook* (26 February 1919).

[13] For example, Chiswick and Houseworth (2011) document a higher likelihood of endogamous marriages among individuals who marry young.

records of enlistments in the Army, the Enlisted Reserve Corps, and the Women's Army Auxiliary Corps. Each entry provides information on enlistment details (Army serial number, enlistment date and place, enlistment term and Army component), and also on several demographic and socioeconomic characteristics of the enlistee (nativity, race, civil status, birth year, birthplace, education, and occupation).

From this universe, I restrict my attention to individuals born in Indiana, Ohio, Michigan, or Kentucky during the period 1880–1916, whom I match to legislation based on their state of birth. Because the enlistment database does not contain information on the birthplace of an individual's parents, I perform a procedure, similar to the one described in Section 3.2, that links enlistees to the 1930 census and determines their ethnic origin. This is not the census year closest in time to the enlistment date range, but it is the closest one for which I can obtain information on parental nativity (since this variable is not generally recorded in the 1940 census).

Volunteers. After Japan attacked Pearl Harbor in early December 1941, Nazi Germany declared war on the United States. Following their country's entry to World War II, thousands of American men volunteered for service. The decision to volunteer is motivated by patriotism and, in the case of first- or second-generation immigrants, it clearly signifies a strong identification with their host country. Especially for second generation Germans, who would be called to fight against the country of their parents, a decision to volunteer is an unmistakable indicator of assimilation.

It is not straightforward to determine whether a person volunteered for the Army or was conscripted. According to the draft classification, enlisted men are those members of the Armed Forces of the United States who volunteered for service. These individuals can be identified by their serial numbers, which belong to the 11 through 19 million series.[14] However, it was possible for a drafted man to enlist in the regular army as a volunteer prior to his induction; doing so gave him more say in the choice of unit and conditions of service. This possibility introduces measurement error when serial numbers are used as a method to identify volunteers, yet the estimation procedure will not be biased provided this error does not differ systematically across cohorts and states. Voluntary enlistment was ended by presidential executive order in 1942, so I restrict my attention to men enlisted between 1940 and 1942. Summary statistics for the linked sample are provided in Table 4.

---

[14]Army Regulation no. 615-30, 1942.

4 Empirical analysis

My identification strategy is a difference-in-differences approach that is based on comparing cohorts of school age and cohorts too old to be at school between states with and without a language law. My main specification takes the form:

$$Y_{isc} = \alpha + \beta T_{cs} + \lambda_c + \theta_s + \delta \mathbf{Z}_{isc} + \varepsilon_{isc} \tag{1}$$

where T_{cs} is an indicator for individuals living in a state with a law and who were within the age range for compulsory schooling at the time that law was in place. The terms λ_c and θ_s signify cohort and state of birth/residence fixed effects. $\mathbf{Z}_{isc}$ is a vector of name string properties that affect the probability of a record being matched in a later census.[15] The coefficient of interest is β: the estimated average effect of legislation on exposed cohorts.

In the above specification, state and cohort fixed effects account for average differences in the outcome variable across states and cohorts. As with every DiD approach, the identifying assumption is that there exists no omitted time-varying and state-specific factor correlated with the passage of language laws. Because it is difficult to completely rule out this concern in an observational setting, I will report specifications that include interactions of state-level variables recorded before the enactment of the law and that are plausibly correlated with its passage (most notably the share of the German element in a state's population) with cohort fixed effects.

$$Y_{isc} = \alpha + \beta T_{cs} + \lambda_c + \theta_s + \delta \mathbf{Z}_{isc} + \sum_{c=1}^{37} \gamma_c \times \text{German share}_s \times \lambda_c + \varepsilon_{isc} \tag{2}$$

It is worth remarking that I do not know precisely which children of German origin attended schools where German was actually used as a language of instruction. This lack of sharp variation across cohorts in terms of language used in school will likely bias all estimates toward zero, since children in non-German schools will be either unaffected by the ban, or – in the case of spillovers across schools – less affected by it than children who actually experienced a change in the language regime. In any case, the DiD coefficient should be an unbiased estimate of the intention-to-treat, that captures the effect of the law on the entire population of Germans in relevant cohorts (including non compliers).

---

[15]These include the length and commonness of the first and last name. Commonness is computed as the share of people in the 1920 census with the same first or last name.

Discrimination The main DiD identifying assumption will be violated if legislation is endogenous to factors that directly affect assimilation outcomes. A plausible scenario is that Indiana and Ohio introduced restrictive laws because those states were characterized by relatively more anti-German sentiment. In that case there should be greater discrimination against Germans, which would affect some outcomes (such as intermarriage) directly and not through any mechanism related to language used in school. This scenario is unlikely for two main reasons. First, in order for differences in the intensity of discrimination to have a differential effect on the younger cohorts exposed to school laws, these differences would have to be increasing over time. Yet we expect the opposite to be true because anti-Germanism peaked during and shortly after the war years and began to subside thereafter. In particular for endogamy, it is equally (if not more) likely that discrimination would affect marriage outcomes for the control cohorts born 1890–1900 — who would be at a marriageable age exactly during the war years — than the treated cohorts born after 1903. Second, sources point to all states conducting a campaign of similar intensity against German during and after the war. Beck (1965) reports that both Ohio and Michigan had many proponents of a language ban, and language restrictions in both states faced militant opposition from Catholic and Lutheran churches. That German was banned in Ohio, but not Michigan, was due largely to idiosyncratic factors.

4.1   Main estimates

The main results for naming patterns are illustrated graphically in Figure 3. The figure plots the density function of the log GNI of the first son, for treatment and control cohorts. While, for older cohorts, the GNI distribution is practically identical between states with and without a law, the younger cohort in Indiana and Ohio experiences a marked shift in the density to the right.

Table 5 reports estimated coefficients derived from a regression of equation (1)'s form. The dependent variable is the the log average German name index of all children in Panel A and the logarithm of the index for the first son in Panel B. The estimated effect of the law is positive in all specifications. Column [1] controls only for properties of the name string. Column [2] inserts state controls measured pre-legislation in 1910 interacted with birth cohort dummies; hence these regressions partially account for the effects of time-varying state-specific unobservables. These controls are, for year 1910, the share of first- and second- generation Germans and an indicator for the border segment. Columns [3] and [4] include county fixed effects and linear state-specific trends respectively. While the magnitude of the log average GNI does not change significantly, the linear trend captures part of the increase in the log of the GNI of

the first son. Still, taken together these results suggest a backlash effect resulting from exposure to a German ban. The magnitude of the coefficient for log(GNI) is meaningful: in my data, it implies that exposure to a language law leads fathers to switch from an Anglo-Saxon name like Andrew to a Germanic name like Fritz or Kurt, or from a name like Ted, to a name like Adelbert.

Panel C reports estimates with endogamy as the dependent variable. The baseline effect of the language ban reported in column [1] suggests that exposure to the law increases intermarriage. There are two factors that affect endogamy rates: one is preferences for ethnic mixing and the other is the size of the marriage market. These two factors do not necessarily move in the same direction. For example, while it is likely that intermarriage across ethnic lines is less desirable in smaller, and potentially more traditional communities, it is also the case that the lack of potential partners inside one's ethnic group makes intermarriage *more* likely in those places. It therefore seems crucial to control for the size of the marriage market in order to isolate the effect of ethnic preferences (Banerjee et al., 2009; Voigtländer and Voth, 2013). This is done in column [2] which includes controls for the (log) population of German origin (first and second generation), as well as for the (log) population in the state in 1910. Because there exists no time-varying measure of these size proxies, I include them interacted with birth cohort indicators. Inclusion of these controls results in a positive coefficient, the magnitude of which is little affected by inclusion of county fixed effects and linear state trends in columns [3] and [4]. The magnitude implies that exposure to a ban of German in school increases the likelihood of endogamous marriage by 7 to 9 percentage points.

Column [5] includes an additional robustness check. As revealed by Table 1, much of the difference in baseline characteristics across border counties is driven by Cincinnati. When Cincinnati is excluded from the sample the estimated effects remain strongly positive and nearly identical in magnitude.

This analysis suggests that the removal of German from elementary schools led to a backlash, a significantly *lower* assimilation along all dimensions for children of German couples. I next turn my attention to WWII volunteering; this is a novel and informative proxy of ethnic identity that has the additional benefit of capturing a clear individual decision.

The results for volunteering rates are presented graphically in Figure 4. As expected, volunteering rates are lower for older cohorts and increase for younger ones. This increase, however, is 7.1 percentage points higher in Indiana and Ohio than in Michigan and Kentucky. While the difference in volunteering rates between states with and without a law is roughly 2 percentage points for cohorts unaffected by the language ban, it is 9.6 percentage points for affected cohorts.

Table 6 reports regression results for volunteering rates. The estimated coefficient suggests that exposure to language laws decreases the likelihood of volunteering by 7 percentage points. Considering that the average volunteering rate among younger cohorts is 13%, this effect is very large. Column [2] incorporates an enlistment year dummy and two additional control variables: one indicator for married individuals and another one for dependent family members. Each of these factors reduces the probability of volunteering in the US Army, but their inclusion has little effect on the magnitude of my estimated coefficient. Column [3] introduces interaction effects of birth cohort fixed effects with the share of Germans in the state in 1910; this increases the effect of laws on volunteering rates, and the coefficient remains significant at the 1% level. Controlling for linear state trends (column [4]) reduces the magnitude of the estimated coefficient, but the effect still corresponds to nearly half the average volunteering rate among younger cohorts.

4.2   Robustness

Table 5 has already demonstrated the robustness of results to time-varying state-level differences, linear state trends, as well as to the exclusion of Cincinnati from the sample. Here, I report a number of falsification tests, which show that no backlash effect is present in different periods or for groups of immigrants other than Germans.

Figure 5 depicts graphically the absence of differential trends in outcomes for older cohorts. It plots the interaction coefficients of twelve birth cohort bins with a dummy for states with a language ban in regressions which include state and birth cohort fixed effects. The upper panel shows that the Germanness of children's names was not trending upwards for individuals in Indiana and Ohio too old to have been affected by the language ban. With the exception of two outlying cohorts before 1885, the difference in the log GNI is not significantly different from zero for any of the cohorts born before 1903, the year marking the first affected cohort of German-American students. Furthermore, the magnitude of the estimated coefficient peaks for the cohorts born between 1907 and 1912, the ones with the maximum exposure to the ban before the law's repeal in 1923. Effects are similar for endogamy and volunteering rates, though more noisily estimated due to the inherent limitations of the data related to these two outcomes: treated cohorts are too young to be married in 1930, and the subset of matched enlisted men with German parents is small. Despite the noise, there is no indication of pre-trends in either case.

The enlistment data set further allows me to compare the behavior of second generation Germans with that of other immigrant groups and of the general population. The removal of German from school curricula should affect German-American children who

were formerly taught in this language but should not affect other immigrants or natives. Column [5] of Table 6 shows that the law had indeed no effect on enlisted native-born individuals of Italian ancestry. While Italians were the other large immigrant group of the time, they were much less educated and did not have an organized network of ethnic schools, like the Germans. Despite the larger sample size, the coefficient for Italian-Americans is imprecisely estimated and positive. More broadly, I can compare the difference in volunteering rates of second-generation Germans across states and cohorts with the respective difference for the rest of the sample. This approach gives rise to a triple-differences specification of the form

$$Y_{isc} = \alpha + \lambda_c + \theta_s + \beta_1 T_{cs} + \gamma_1 G_{isc} + \sum_{c=1}^{37} \gamma_{2c} G_{isc} \times \lambda_c$$

$$+ \sum_{s=1}^{4} \gamma_{3c} G_{isc} \times \theta_s + \beta_2 T_{cs} \times G_{isc} + \delta \mathbf{Z}_{isc} + \varepsilon_{isc} \qquad (3)$$

where G_{isc} is an indicator for individuals with German parents. The coefficient β_2 now identifies the average effect of legislation on affected cohorts of German-origin individuals. Column [6] of Table 6 reports estimates from this specification. While the effect of the law is negative and significant at the 1% level for German-Americans, it is one third in magnitude and imprecisely estimated for potentially treated cohorts of non-German origin.

A general concern with statistical inference is the small number of states in my data. All reported significance levels are based on standard errors clustered at the state of residence in 1920 × cohort cell, the level at which the effect of the law varies. A more conservative clustering at the level of the state, which allows for arbitrary patterns of autocorrelation within states across time (Bertrand, Duflo and Mullainathan, 2004), yields consistent estimates of the standard errors only with a sufficiently large number of clusters.[16] To avoid relying on asymptotic assumptions for such a small sample, I use randomization inference to non-parametrically compute p-values. I randomly assign the enactment of language laws to state×cohort cells, keeping the share of cells with a language law equal to the empirically observed proportion. I then compare the distribution of the estimated interaction coefficients from 10,000 random assignments to the actual effect of the law (Gerber and Green, 2012; Young, 2015). Figure 6 plots these distributions for each of my main outcome variables and reports p-values computed

---

[16]The significance of estimated coefficients in the entire analysis is little affected when clustering at the state level. Results available upon request.

as the share of coefficients with value larger (in absolute value) than the estimated treatment effect. With the exception of endogamy, all results are similar to those obtained by conventional inference.

Taken together, the results presented in this section suggest that removing a child's home language from the school need not lead to more assimilation and can, in fact, have the exact opposite effect on ethnic preferences. The purpose of the next section is to shed more light on the channels through which language in the school affects assimilation outcomes later in life.

5 Mechanisms

Here, I first test whether language restrictions are more likely to succeed in assimilating immigrants when immigrants are themselves relatively more assimilated into mainstream society. I then show that the backlash effect depends in a similar way on the ethnic character of the community, and that it is not driven by lower educational achievement among affected cohorts.

5.1 Parents' ethnic background

Why would we expect an intervention that alters the ethnic character of education to have different effects on different groups of immigrants? When schooling is a substitute for parental investment in the ethnic preferences of children, a decrease in the ethnic content of education will increase the investment of parents with a strong ethnic identity but have the opposite effect on the investment of more assimilated parents. Common sense and the history of bilingual programs both suggest a similar dynamic. Allowing for the use of a minority language as an aide in early school years can actually help children assimilate, by allowing them to transition smoothly from the language of home and their parents to English. In the extreme case — when German language instruction is no longer an option at school — those parents with a strong preference for socializing their children to German culture will make a greater effort to instill that culture at home.

Here I investigate how the effects of language policies differ along one important dimension of heterogeneity in parents' ethnic identity: ethnic intermarriage. Toward this end, I extend my data set to include individuals born to mixed couples (German father and non-German mother). Ethnic identity is expected to be stronger when both parents are German, not only because the child then has two German role models in the family instead of one, but also because within-group marriage is the endogenous decision of individuals who care relatively more about their ethnic identity and its

transmission to their offspring. Such individuals choose to marry someone from their own ethnic group precisely because doing so increases the likelihood that children will inherit the parents' culture (Bisin, Topa and Verdier, 2004).[17]

Table 7 repeats the baseline analysis, this time examining individuals whose father is German but whose mother is not. The effect of legislation on the GNI is now negative and imprecisely estimated. Endogamy still increases in response to the law, but less than in the baseline. Only for volunteering are results similar to those in the sample of homogamous couples. Volunteering rates decline by up to 9.5 p.p. for treated cohorts, indicating an equally pronounced strengthening of ethnic identification in response to language laws.

Having one non-German parent makes it more likely for language restrictions to succeed in assimilating German-Americans. To the extent that mixed couples have a less pronounced sense of German identity, the finding is compatible with a theoretical mechanism in which the effort of enculturating children is increasing in the initial sense of identity. Particularly in the case of first names, estimates suggest an assimilating effect of legislation which is, however, smaller in magnitude than the backlash observed in the group born to two German parents. Taken together with Section 4.1, these results suggest that language laws increase the variance in outcomes within the German group.

5.2 Ethnic composition and strength of identity

Community size. Does the language ban's effect depend on the share of Germans in the community? The answer to this question is not clear a priori. On the one hand, communities with more Germans might be better organized and hence better able to react against efforts to suppress their culture.[18] On the other hand, smaller communities tend to be more cohesive in their actions. Models of cultural transmission (e.g. Bisin and Verdier 2001) predict a stronger backlash among smaller minorities. Because a child is more likely to be assimilated when part of a small minority, parents are more incentivized to invest heavily in that child's identity. Thus smaller minorities have a stronger sense of ethnic identity.

I examine how the backlash effect of the law for individuals with German parents depends on the share of Germans in their county. For this I employ a triple-differences specification in which the treatment dummy is interacted with the share S_{sj} of first-

---

[17]Several studies document lower ethnic attachment among the offspring of interethnic marriages (Alba, 1990; Waters, 1990; Perlmann and Waters, 2007).

[18]In a model of language assimilation emphasizing a trade-related mechanism, Lazear (1999) shows that smaller minorities are more likely to assimilate.

and second- generation Germans in the county in 1910:

$$Y_{isjc} = \alpha + \lambda_c + \theta_s + z_{sj} + \beta_1 T_{sc} + \gamma_1 S_{sj} + \sum_{c=1}^{37} \gamma_{2c} S_{sj} \times \lambda_c$$

$$+ \sum_{s=1}^{4} \gamma_{3c} S_{sj} \times \theta_s + \beta_2 T_{cs} \times S_{sj} + \delta \mathbf{Z}_{isjc} + \varepsilon_{isjc} \qquad (4)$$

where j denotes counties and z_{sj} is a county fixed effect. The left panel of Figure 7 plots the coefficient β_2 and 90% confidence intervals against the share of Germans, S_{sj}. The dependent variable is the logarithm of the GNI of the first son.[19] The magnitude of the coefficient is decreasing in the share of Germans, indicating a greater reaction in counties where Germans constitute a smaller minority.

Strength of identity. Theories of cultural transmission suggest that a reaction to language restrictions should be increasing in the strength of initial ethnic identity. Section 5.1 provides evidence for this by showing that the observed backlash weakens substantially or disappears among individuals with only one German parent. I use the share of Lutherans in a county as an additional measure of German identity. Although most German-Americans in the United States at the start of the 20th century were Catholics, it was Lutheranism that had the most German members (Wüstenbecker, 2007). The Lutheran religion was also the one most strongly emphasizing conservation of the German language as a medium for transmitting the faith. Lutheran churches could follow this language policy more independently than could German Catholic churches, which were guided not by Germany but rather by the Pope in Rome (Rippley, 1985; Wüstenbecker, 2007). The Catholic Church was multiethnic but dominated by the Irish and Polish, which caused concern among prominent German-Americans that Catholic parishes were losing their German character (Viereck, 1903). German Lutherans were — among all old-church Protestants — the denomination with the highest commitment to parochial schooling (Kraushaar, 1972).

With Lutheranism as a proxy for German ethnic identity, I use a specification identical to equation (4) but with S_{sj} now denoting the share of members of Lutheran churches in a county in 1906. Data for calculating this share are from the 1906 Census of Religious Bodies. The right panel of Figure 7 plots the triple-interaction coefficient against the share of Lutherans for the GNI of the first son for those individuals in the border data set who have two German parents. The magnitude of the reaction is indeed

---

[19]Here and in the rest of Section 5.2, results are similar qualitatively when the dependent variable is instead the log average GNI of all children.

increasing with the share of Lutherans suggesting a stronger backlash in places with a greater sense of Germanness.

5.3 Effects on educational achievement

One way through which language in elementary school can affect life outcomes such as intermarriage, is via ethnic preferences. One channel in particular is parental investment, with German parents reacting to language restrictions at school by adjusting their own investment in their children's ethnic identity. Yet it is also possible that banning the German language from elementary schools has direct effects on the content and quality of education of German-American children.[20] Such effects could in turn impact language proficiency, mobility rates, or cultural adaptability and thereby intermarriage (Bleakley and Chin, 2010; Wozniak, 2010; Furtado and Theodoropoulos, 2010). For many children, especially those born to homogamous German couples, instruction in German at school may contribute to their smooth transition from the language spoken at home to the language of society. In the absence of this auxiliary language, the schooling outcomes of these children might be worse. Conversely, for children of already assimilated families, German instruction might constitute an impediment to their progress in English language courses (Chin, Daysal and Imberman, 2013).

Table 8 tests these notions, by examining how the German language ban affected the years of schooling completed by individuals in the border data set. Columns [1] and [2] report heterogeneous effects by nativity of the mother. The language ban increases schooling for both children of homogamous and of mixed couples. This positive effect is larger for the latter group, but differences are small, as are the respective magnitudes. These results suggest that the observed backlash effect is not due to lower quantity of education and lend support to the claim that observed effects resulted mainly from ethnic preferences and the parents' socialization efforts.[21]

Backlash costs. Is the strengthening of ethnic identity in response to language laws, costly for exposed cohorts later in life? Studies on intermarriage (for a review, see Furtado and Trejo 2013) — and on other assimilation decisions of immigrants, such as Americanizing surnames (Biavaschi, Giulietti and Siddique, 2013) — indicate

---

[20]Eriksson (2014) and Ramachandran (2013) demonstrate that mother tongue instruction in primary school has positive effects on years of schooling, literacy and wages in South Africa and Ethiopia respectively. In a related study, Dee and Penner (2016) show that "culturally relevant" curricula which include ethnic studies courses increase both school attendance and educational attainment.

[21]That the German language ban increases schooling, but decreases integration later in life accords with the findings of Friedman et al. (2011) and Croke et al. (forthcoming) that more education in an authoritarian context reduces political participation.

that, notwithstanding the possibility of immigrants' self-selection, assimilation entails a premium in the labor market.[22] Conversely, it is conceivable that strongly adhering to one's own ethnicity implies a cost (Battu and Zenou, 2010). Individuals who marry endogamously lose access to valuable networks outside their ethnic community and thus may be sacrificing mobility by retaining strong ties with their communities.

Columns [3]–[4] in Table 8 is only a weak indication that such costs might apply in the case of German-Americans. The estimated effect of the German language ban on the log of yearly wage earnings of individuals with German parents is large, but not statistically significant: exposure to that law implies an imprecisely estimated 12% reduction in yearly wage income for this group. There is no clear indication of an assimilation premium for the mixed group; the estimated coefficient is negative and insignificant, and the magnitude is small. Given that the reaction of this group was not uniform in terms of all outcomes examined, and also often moved in the direction of a backlash, this latter result makes sense.

6   Discussion

The findings reported here suggest that restrictions on immigrants' native language increase ethnic identity and that this response is more common among children of parents with a strong attachment to their ethnicity. This pattern can be accounted for by a model of cultural transmission of ethnic identity in which ethnic schooling and parental investment in identity are substitutes. When the school's function of socializing children to their parents' preferred culture is weakened, parents respond by increasing their own investment at home. While these efforts alone may be insufficient to counteract the assimilating effect of an education that is less ethnic in character, any additional psychological benefit derived from social interactions with other oppositional types can act as a multiplier of parental investment. This peer effect channel can induce investment that is high enough to result in a reversal of the policy's effects. Section A.1 in the Appendix presents a formalization of this intuition, building on the framework of Bisin and Verdier (2001) and Bisin et al. (2011). The substitution of a German language curriculum with increased investment in other forms of German enculturation need not only take place at home. Lutheran schools could be responding to language restrictions by modifying their curriculum along other dimensions emphasizing German

---

[22]A related body of work on blacks adopting a white racial identity in 19th and 20th century America (Mill and Stein, 2015; Nix and Qian, 2015) uncovers a substantial positive effect of "passing" on economic outcomes.

education, and churches could be increasing their efforts to inculcate German culture through sermons or Sunday schools in the absence of linguistic means.

A cultural transmission framework can account for a stronger backlash in communities with a greater sense of Germanness, where parental investment in ethnic identity is initially higher. This channel may be complementary to others: localities with a stronger German identity may facilitate parental investment outside the school through German clubs and associations, or may achieve a less strict enforcement of the language ban in the classrooms. The stronger reaction in places where Germans are a smaller share of the total population can also be accounted for in terms of strength of identity. Intuitively, in places where the German minority is smaller, parents put more effort in shaping each child's sense of ethnicity because reliance on peer interaction is not guaranteed to transmit their culture. Such places have a stronger initial sense of ethnic identification and consequently react more to any attempted assimilation. Other theories can produce similar results. Jia and Persson (2016) also find evidence of a negative correlation between the response to external incentives for assimilation and the size of the group. In the context of policies favoring minorities in China, they show that material benefits for changing the identity of one's children have a smaller impact on parents' decisions when the size of the group sharing that identity is small. Building on Bénabou and Tirole (2011b), they suggest that social considerations and intrinsic motivations are more likely to crowd out any external incentives in smaller communities.

A reaction to attempted assimilation is broadly compatible with a number of theories. Applying their seminal framework of identity on education, Akerlof and Kranton (2002) show how schools which promote a single social category or educational ideal can alienate students whose background is too distant from the behaviors that this ideal prescribes. Their model can explain the clash between immigrant students and Americanizing schools of the early 20th century — interestingly, those less assimilated would be more likely to distance themselves from the behaviors prescribed by the school. In the framework of Bisin et al. (2011), children are allowed to choose their own identity. When the share of oppositional types in society is reduced because of an assimilation attempt by the majority, the remaining oppositional individuals have an incentive to strengthen their identity and thus to reduce their costs of interacting with people who are different from them. In this model, a language ban would lead to fewer but more intensely oppositional types. In Bénabou and Tirole (2011a), identity is an asset built with investment over time. Increases in the salience of identity or in the uncertainty of one's type, such as might well be sustained by second-generation immigrants discriminated by the majority population, can lead to costly investments in identity if the initial ethnic identification (here the sense of Germanness among the parents' generation) is

strong.

7   Conclusion

Can cultural assimilation be engineered through government policies? I examine the prohibition of German in US elementary schools and its effects on the assimilation of German children. Using both linked census records and information on WWII volunteers, I show that the policy had a negative effect on assimilation outcomes, particularly for individuals of more homogeneous German background. This effect is larger in areas where there were fewer Germans. This strongly suggests that parents overcompensate, investing in their child's identity all the more as horizontal socialization declines. Effects are larger in areas with more Lutherans suggesting that an ethnic community's initial degree of identity determines the magnitude of its reaction to assimilation efforts.

Can the historical case study of US Germans inform modern-day language and integration policies? The debate about language restrictions is very much alive in immigrant receiving states and countries, such as California and Germany.[23] This suggests that modern day societies face many of the same questions. Furthermore, the finding of a backlash in a well-integrated prosperous immigrant group such as the Germans in the US[24] implies that negative consequences of assimilation policies may be even more likely amongst poor marginalized groups — such as Muslims in Europe.

One of the implications of this paper is that policies favoring linguistic and cultural autonomy may actually increase social cohesion — both by facilitating assimilation for the least integrated minority members and by decreasing the variance within the minority group.[25] My findings thus highlight a dimension that is complementary to educational achievement and that should be considered when debating bilingual education and linguistic immersion policies.

---

[23]In 2006, the Herbert Hoover School (a low-track secondary school in Berlin) implemented a ban on Turkish and other foreign languages on its premises, a policy that earned it the German National Prize and $94,000 from the National German Foundation. The school's director, Jutta Steinkamp, explained that "this ban [has been introduced] to enable our students to take part in German society through speaking and understanding the language properly" and that "knowing the language is a precondition for successful integration" (Crutchfield, 2007).

[24]They had the highest rates of naturalization among the foreign-born (Rippley, 1985).

[25]This evidence from history accords with studies reporting positive effects of contemporary multiculturalist policies on immigrant integration (Wright and Bloemraad, 2012).

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy*, 122(3): 467–717.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2016. "Cultural Assimilation during the Age of Mass Migration." NBER Working Paper 22381.

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *The Quarterly Journal of Economics*, 115(3): 715–753.

Akerlof, George A., and Rachel E. Kranton. 2002. "Identity and Schooling: Some Lessons for the Economics of Education." *Journal of Economic Literature*, XL: 1167–1201.

Alba, Richard D. 1990. *Ethnic Identity: The Transformation of White America.* New Haven, CT:Yale University Press.

Aleksynska, Mariya, and Yann Algan. 2010. "Assimilation and Integration of Immigrants in Europe." IZA Discussion Paper 5185.

Alesina, Alberto, and Bryony Reich. 2013. "Nation Building." NBER Working Paper 18839.

Algan, Yann, Thierry Mayer, and Mathias Thoenig. 2013. "The Economic Incentives of Cultural Transmission: Spatial Evidence from Naming Patterns across France." CEPR Discussion Papers 9416.

Arai, Mahmood, Damien Besancenot, Kim Huynh, and Ali Skalli. 2009. "Children's First Names and Immigration Background in France." Stockholm University, Department of Economics Research Paper in Economics 13.

Austen-Smith, David, and Roland G. Fryer. 2005. "An Economic Analysis of "Acting White"." *The Quarterly Journal of Economics*, 120(2): 551–583.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune. 2009. "Marry for What: Caste and Mate Selection in Modern India." NBER Working Paper 14958.

Baron, Dennis. 1990. *The English-Only Question: An Official Language for Americans?* New Haven:Yale University Press.

Battu, Harminder, and Yves Zenou. 2010. "Oppositional Identities and Employment for Ethnic Minorities. Evidence for England." *Economic Journal*, 524(120): F52–F71.

Beck, Walter Herman. 1965. *Lutheran Elementary Schools in the United States: A History of the Development of Parochial Schools and Synodical Educational Policies and Programs.* St. Louis:Concordia Publishing House.

Bénabou, Roland, and Jean Tirole. 2011*a*. "Identity, Morals, and Taboos: Beliefs as Assets." *The Quarterly Journal of Economics*, 126(2): 805–855.

Bénabou, Roland, and Jean Tirole. 2011*b*. "Laws and Norms." NBER Working Paper 17579.

Benjamin, Daniel J., James J. Choi, and A. Joshua Strickland. 2010. "Social Identity and Preferences." *American Economic Review*, 100(4): 1913–28.

Benjamin, Daniel J., James J. Choi, and Geoffrey Fisher. 2013. "Religious Identity and Economic Behavior." Cornell University Mimeo.

Bennett Woodhouse, Barbara. 1992. ""Who Owns the Child?": Meyer and Pierce and the Child as Property." *William and Mary Law Review*, 33(4).

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should we Trust Differences-In-Differences Estimates." *The Quarterly Journal of Economics*, 119(1): 249–275.

Biavaschi, Costanza, Corrado Giulietti, and Zahra Siddique. 2013. "The Economic Payoff of Name Americanization." IZA Discussion Paper 7725.

Bisin, Alberto, and Thierry Verdier. 2000. ""Beyond the Melting Pot": Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits." *The Quarterly Journal of Economics*, 115(3): 955–988.

Bisin, Alberto, and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory*, 97(2): 298–319.

Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, and Yves Zenou. 2011. "Formation and Persistence of Oppositional Identities." *European Economic Review*, 55(8): 1046–1071.

Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, and Yves Zenou. 2013. "Bend It Like Beckham: Ethnic Identity and Integration." New York University Mimeo.

Bisin, Alberto, Giorgio Topa, and Thierry Verdier. 2004. "Religious Intermarriage and Socialization in the United States." *Journal of Political Economy*, 112(3): 615–664.

Bleakley, Hoyt, and Aimee Chin. 2010. "Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants." *American Economic Journal: Applied Economics*, 2(1): 165–92.

Borjas, George J. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *Journal of Labor Economics*, 3(4): 463–489.

Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang. forthcoming. "Curriculum and Ideology." *Journal of Political Economy*.

Card, David. 2005. "Is the New Immigration Really so Bad?" *Economic Journal*, 115(507): F300–F323.

Carvalho, Jean-Paul. 2013. "Veiling." *The Quarterly Journal of Economics*, 128(1): 337–370.

Carvalho, Jean-Paul, and Mark Koyama. 2016. "Resisting Education." GMU Working Paper in Economics 13-18.

Chin, Aimee, N. Meltem Daysal, and Scott A. Imberman. 2013. "Impact of Bilingual Education Programs on Limited English Proficient Students and their Peers: Regression Discontinuity Evidence from Texas." *Journal of Public Economics*, 107(C): 63–78.

Chiswick, Barry, and Christina Houseworth. 2011. "Ethnic Intermarriage Among Immigrants: Human Capital and Assortative Mating." *Review of Economics of the Household*, 9(2): 149–180.

Clots-Figueras, Irma, and Paolo Masella. 2013. "Education, Language and Identity." *Economic Journal*, 123(570): F332–F357.

Colley, Linda. 1992. *Britons: Forging the Nation, 1707-1837.* New Haven and London:Yale University Press.

Collins, William J., and Marianne H. Wanamaker. 2013. "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data." NBER Working Paper 19124.

Constant, Amelie, Liliya Gataullina, and Klaus F. Zimmermann. 2009. "Eth-

nosizing Immigrants." *Journal of Economic Behavior and Organization*, 69: 274–287.

Conzen, Cathleen Neils. 1980. "Germans." In *Harvard Encyclopedia of American Ethnic Groups.* , ed. Stephan Thernstrom, Ann Orlov and Oscar Handlin. Cambridge, Massachusetts:Belknap Press of Harvard University Press.

Cook, Lisa D, Trevon D Logan, and John M Parman. 2013. "Distinctively Black Names in the American Past." NBER Working Paper 18802.

Croke, Kevin, Guy Grossman, Horacio Larreguy, and John Marshall. forthcoming. "Deliberate Disengagement: How Education Can Decrease Political Participation in Electoral Authoritarian Regimes." *American Political Science Review.*

Crutchfield, Anne. 2007. "Language Minority Education Policy: Turkish Immigrant Pupils in Germany." University of Hawaii Mimeo.

Dávila, Alberto, and Marie Mora. 2005. "Changes in the Earnings of Arab Men in the US between 2000 and 2002." *Journal of Population Economics*, 18(4): 587–601.

Dee, Thomas, and Emily Penner. 2016. "The Causal Effects of Cultural Relevance: Evidence from an Ethnic Studies Curriculum." NBER Working Paper 21865.

Dewey, John. 1916. *Democracy and Education: An Introduction to the Philosophy of Education.* New York, N.Y.:Macmillan.

Edwards, I.N. 1923. "The Legal Status of Foreign Languages in the Schools." *The Elementary School Journal*, 24(4): 270–278.

Eriksson, Katherine. 2014. "Does the Language of Instruction in Primary School Affect Later Labor Market Outcomes? Evidence from South Africa." *Economic History of Developing Regions*, 29(2): 311–335.

Feigenbaum, James. 2015. "Intergenerational Mobility during the Great Depression." Harvard mimeo.

Ferrie, Joseph. 1996. "A New Sample of Males Linked from the Public Use Micro Sample of the 1850 U.S. Federal Census of Population to the 1860 U.S. Federal Census Manuscript Schedules." *Historical Methods*, 29: 141–156.

Fisman, Ray, Sheena S. Iyengar, Emir Kamenica, and Itamar Simonson. 2008. "Racial Preferences in Dating." *Review of Economic Studies*, 75(1): 117–132.

Freire, Paulo. 1970. *Pedagogy of the Oppressed.* New York, N.Y.:Continuum.

Friedman, Willa, Michael Kremer, Edward Miguel, and Rebecca Thornton. 2011. "Education as Liberation?" NBER Working Paper 16939.

Fryer, Roland G. 2007. "Guess Who's Been Coming to Dinner? Trends in Interracial Marriage over the 20th Century." *Journal of Economic Perspectives*, 21(2): 71–90.

Fryer, Roland G., and Steven D. Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *The Quarterly Journal of Economics*, 119(3): 767–805.

Furtado, Delia, and Nikolaos Theodoropoulos. 2010. "Why Does Intermarriage Increase Immigrant Employment? The Role of Networks." *The B.E. Journal of Economic Analysis & Policy*, 10(1): 1–33.

Furtado, Delia, and Stephen J. Trejo. 2013. "Interethnic Marriages and their Economic Effects." In *International Handbook on the Economics of Migration.* , ed. Amelie F. Constant and Klaus F. Zimmermann, 276–292. UK:Edward Elgar Publishing.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpreta- tion.* New York:W. W. Norton.

Gerhards, Jürgen. 2005. *The Name Game: Cultural Modernization & First Names.* Transaction Publishers.

Goldin, Claudia, and Lawrence F. Katz. 2008. "Mass Secondary Schooling and the State: The Role of State Compulsion in the High School Movement." In *Understanding Long-Run Economic Growth: Geography, Institutions, and the Knowledge Economy.* 275–310. NBER.

Gordon, Milton A. 1964. *Assimilation in American Life: The Role of Race, Religion, and National Origins.* New York:Oxford University Press.

Gould, Eric D, and Esteban F Klor. 2012. "The Long-Run Effect of 9/11: Terrorism, Backlash, and the Assimilation of Muslim Immigrants in the West." CEPR Discussion Paper 8797.

Haines, Michael R., and Inter-university Consortium for Political and Social Research. 2010. *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 [Computer file]. ICPSR02896-v3.* Ann Arbor, MI:Inter-university Consortium for Political and Social Research [distributor].

Hatton, Timothy J. 1997. "The Immigrant Assimilation Puzzle in Late Nineteenth-Centuty America." *The Journal of Economic History*, 57(01): 34–62.

Head, Keith, and Thierry Mayer. 2008. "Detection of Local Interactions from the Spatial Pattern of Names in France." *Journal of Regional Science*, 48(1): 67–95.

Higham, John. 1998. *Strangers in the Land: Patterns of American Nativism, 1860-1925.* New Brunswick:Rutgers University Press.

Hoff, Karla, and Priyanka Pandey. 2006. "Discrimination, Social Identity, and Durable Inequalities." *American Economic Review*, 96(2): 206–211.

Hood, William R. 1920. *State Laws Relating to Education Enacted in 1918 and 1919.* United States Department of the Interior. Bureau of Education.

Jia, Ruixue, and Torsten Persson. 2016. "Individual vs. Social Motives in Identity Choice: Theory and Evidence from China." Stockholm University mimeo.

Kazal, Russell. 2004. *Becoming Old Stock: The Paradox of German-American Identity.* Princeton, N.J.:Princeton University Press.

Knowlton Flanders, Jesse. 1925. *Legislative Control of the Elementary Curriculum.* New York City:Teachers College, Columbia University.

Kraushaar, Otto F. 1972. *American Non-Public Schools: Patterns of Diversity.* Baltimore:Johns Hopkins University Press.

LaLonde, Robert J., and Robert H. Topel. 1991. "Immigrants in the American Labor Market: Quality, Assimilation, and Distributional Effects." *The American Economic Review Papers and Proceedings*, 81(2): 297–302.

Lazear, Edward P. 1999. "Culture and Language." *Journal of Political Economy*, 107(S6): S95–S126.

Leibowitz, Arnold H. 1971. *Educational Policy and Political Acceptance: The Imposition of English as the Language of Instruction in American Schools.* ERIC Clearinghouse for Linguistics, Center for Applied Linguistics.

Lieberson, Stanley. 2000. *A Matter of Taste: How Names, Fashions, and Culture Change.* New Haven:Yale University Press.

Lleras-Muney, Adriana, and Allison Shertzer. 2012. "Did the Americanization Movement Succeed? An Evaluation of the Effect of English-Only and Compulsory School Laws on Immigrants' Education." NBER Working Paper 18302.

Long, Jason, and Joseph Ferrie. 2013. "Intergenerational Occupational Mobility in Great Britain and the United States Since 1850." *American Economic Review*, 103(4): 1109–1137.

Luebke, Frederick C. 1974. *Bonds of Loyalty: German-Americans and World War I.* Urbana:Northern Illinois University Press.

Luebke, Frederick C. 1999. *Germans in the New World: Essays in the History of Immigration.* Urbana:University of Illinois Press.

Manning, Alan, and Sanchari Roy. 2010. "Culture Clash or Culture Club? National Identity in Britain." *The Economic Journal*, 120(542).

Michaels, Guy, and Xiaojia Zhi. 2010. "Freedom Fries." *American Economic Journal: Applied Economics*, 2(3): 256–81.

Mill, Roy. 2012. "Assessing Individual-Level Record Linkage Between Historical Datasets." Stanford University Mimeo.

Mill, Roy, and Luke C.D. Stein. 2015. "Race, Skin Color, and Economic Outcomes in Early Twentieth-Century America." Arizona State University Mimeo.

Minns, Chris. 2000. "Income, Cohort Effects and Occupational Mobility: A New Look at Immigration to the United States at the Turn of the 20th Century." *Explorations in Economic History*, 37: 326–350.

Moser, Petra. 2012. "Taste-Based Discrimination. Evidence from a Shift in Ethnic Preferences after WWI." *Explorations in Economic History*, 49(2): 167–188.

Neeraj, Kaushal, Robert Kaestner, and Cordelia Reimers. 2005. "Labor Market Effects of September 11th on Arab and Muslim Residents of the United States.." *Journal of Population Economics*, 18(4): 587–601.

Nix, Emily, and Nancy Qian. 2015. "The Fluidity of Race: "Passing" in the United States, 1880-1940." NBER Working Paper 20828.

Pagnini, Deanna L., and S. Philip Morgan. 1990. "Intermarriage and Social Distance Among U.S. Immigrants at the Turn of the Century." *American Journal of Sociology*, 96(2): 405–432.

Parman, John. 2015. "Childhood Health and Human Capital: New Evidence from Genetic Brothers in Arms." *The Journal of Economic History*, 75(1): 30–64.

Parman, John M. 2011. "American Mobility and the Expansion of Public Education." *The Journal of Economic History*, 71(1): 105–132.

Perlmann, Joel, and Mary C. Waters. 2007. "Intermarriage and Multiple Identities." In *The New Americans: A Guide to Immigration since 1965.* , ed. Mary C. Waters and Reed Udea, 110–123. Cambridge, MA:Harvard University Press.

Ramachandran, Rajesh. 2013. "Language Use in Education and Primary Schooling Attainment: Evidence from a Natural Experiment in Ethiopia." Universitat Autonoma de Barcelona Mimeo.

Rippley, La Vern J. 1981. "Conflict in the Classroom: Anti-Germanism in Minnesota Schools, 1917-19." *Minnesota History*, 47(5).

Rippley, La Vern J. 1985. "Ameliorated Americanization: The Effect of World War I on German-Americans in the 1920s." In *America and the Germans: An Assessment of a Three-Hundred-Year History.* , ed. Frank Trommler and Joseph McVeigh. Philadelphia:University of Pennsylvania Press.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0 [Machine-Readable Database].* University of Minnesota.

Ruppenthal, J.C. 1919. "The Legal Status of the English Language in the American School System." *School and Society*, 10: 658–66.

Schlossman, Steven L. 1983. "Is There an American Tradition of Bilingual Educa-

tion? German in the Public Elementary Schools, 1840-1919." *American Journal of Education*, 91(2): 139–186.

Shertzer, Allison. 2013. "Immigrant Group Size and Political Mobilization: Evidence from European Migration to the United States." NBER Working Paper 18827.

United States Department of Commerce. Bureau of the Census. 1984. *Census of Religious Bodies, 1906.* Inter-university Consortium for Political and Social Research [distributor].

Van Alstyne, William W. 1990. "Academic Freedom and the First Amendment in the Supreme Court of the United States: An Unhurried Historical Review." William and Mary Law School Faculty Publications 748.

Viereck, Louis. 1903. *Zwei Jahrhunderte Deutschen Unterrichts in den Vereinigten Staaten.* Braunschweig:F. Vieweg und Sohn.

Voigtländer, Nico, and Hans-Joachim Voth. 2013. "Married to Intolerance: Attitudes toward Intermarriage in Germany, 1900-2006." *American Economic Review, Papers and Proceedings*, 103(3): 79–85.

Waters, Mary C. 1990. *Ethnic Options: Choosing Identities in America.* Berkeley, California:University of California Press.

Weber, Eugen. 1976. *Peasants Into Frenchmen: The Modernization of Rural France.* Stanford, California:Stanford University Press.

Wozniak, Abigail. 2010. "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources*, 45(4): 944–970.

Wright, Matthew, and Irene Bloemraad. 2012. "Is There a Trade-off between Multiculturalism and Socio-Political Integration? Policy Regimes and Immigrant Incorporation in Comparative Perspective." *Perspectives on Politics*, 10: 77–95.

Wüstenbecker, Katja. 2007. *Deutsch-Amerikaner im Ersten Weltkrieg: US-Politik und Nationale Identitäten im Mittleren Westen.* Stuttgart:Steiner.

Young, Alwyn. 2015. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." LSE Mimeo.

Zimmerman, Jonathan. 2002. "Ethnics against Ethnicity: European Immigrants and Foreign-Language Instruction, 1890-1940." *The Journal of American History*, 88(4): 1383–1404.

Figures and Tables

Figure 1. Counties on the northern and southern borders of Indiana and Ohio

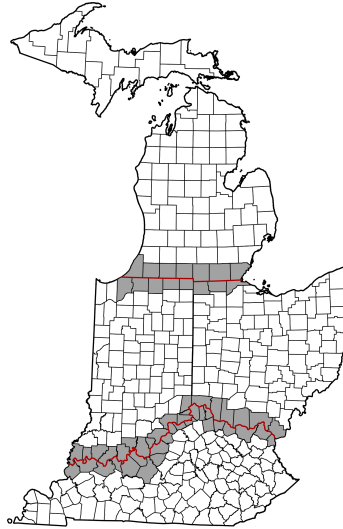
Figure 2. Locations of linked data set in 1920
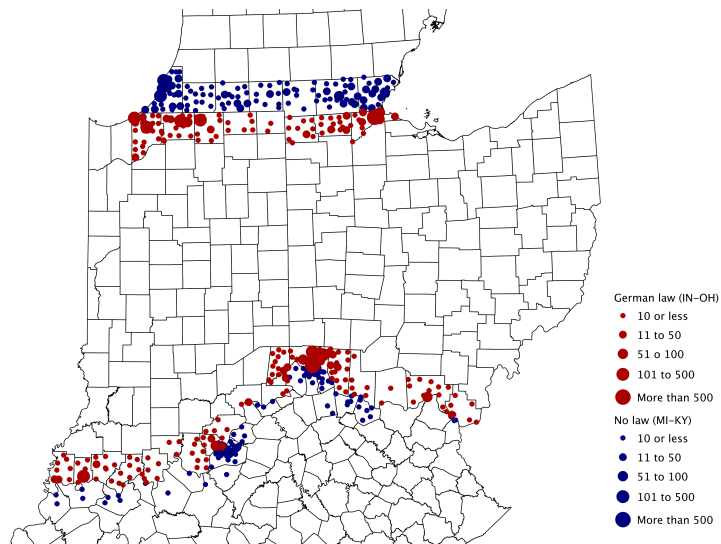
Notes: The map shows the town-level location of all males, who were born 1880–1916 to German parents, were living in a border county in 1920 and who could be linked to the 1930 or 1940 census.

Figure 3. Densities of log GNI of first son by cohort
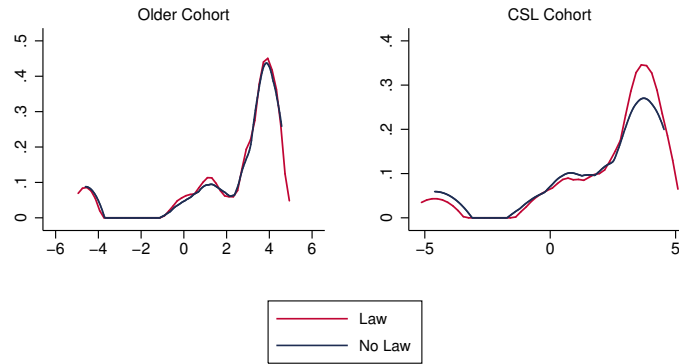
Notes: The figure illustrates, for the linked border dataset, the kernel density of the logarithm of the GNI of the first son. The panel on the left plots this density for the cohort too old to have been in school (by compulsory law) at the time German was banned; the right panel plots the density for the treated cohort.

Figure 4. Share of volunteers by cohort and law status
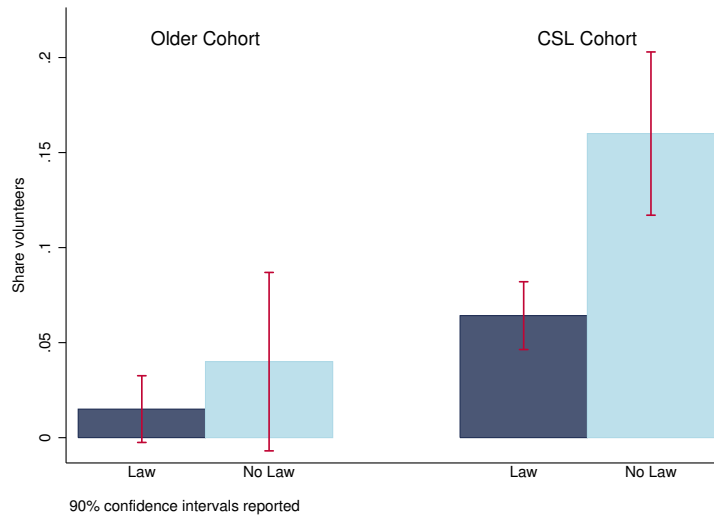
Notes: The bars on the left show the share of US Army volunteers by language law status for the cohort too old to have been in school (by compulsory law) at the time German was banned; the bars on the right plot the respective share for the treated cohort.

Figure 5. Estimated effects of a language ban by birth cohort bin
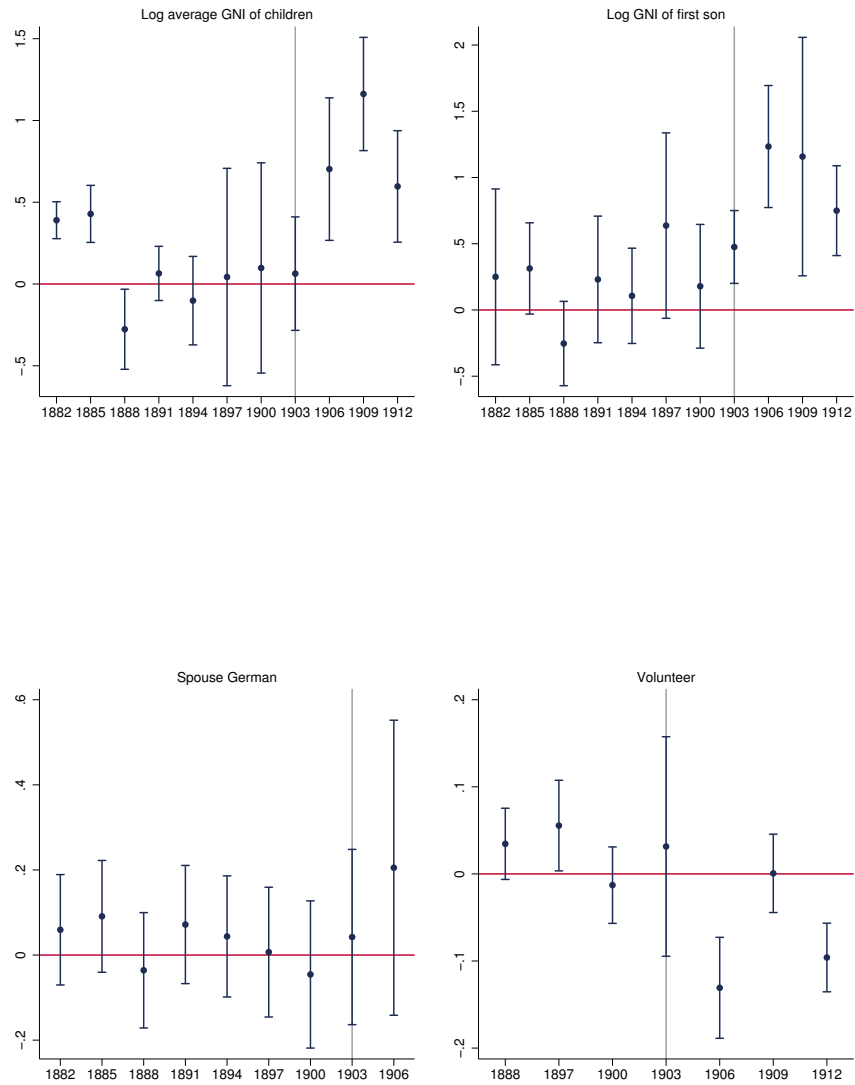
Notes: The figure shows coefficient estimates and 95% confidence intervals from a regression of each outcome on state and birth cohort fixed effects and a set of interactions of 3-year birth cohort bins with an indicator for a language ban. The grey line in year 1903 indicates the first cohort to be affected by the language ban.

Figure 6. Randomization inference
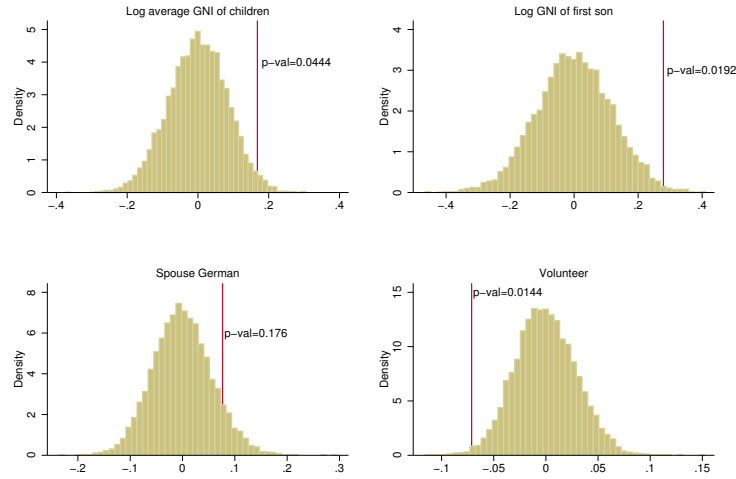
Notes: The figure plots, for each of the main regression outcomes, the distribution of coefficients resulting from 10,000 random assignments of cohorts and states to a language law. P-values are computed as the share of coefficients whose value is more extreme than the value estimated using actual assignment to a treated cohort status.

Figure 7. Ethnic and religious composition and effects of the language ban
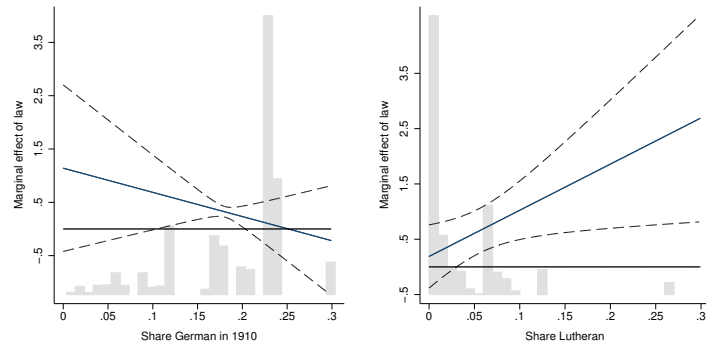
Notes: The figure plots the triple interaction coefficient from a regression specified in equation (4) against the share of first- and second-generation Germans (left panel) and the share of Lutheran church members (right panel) in a county in 1910. The dependent variable is the logarithm of the GNI of the first son. Dashed lines represent 90% confidence intervals. The underlying histograms show how the data is distributed across counties with different shares of Germans (left panel) and Lutheran church members (right panel) in 1910. In all cases, the data are restricted to native-born individuals with two German parents. Data on county shares of German ethnic stock is from ICPSR. Data on county shares of Lutheran church members are from the 1906 Census of Religious Bodies .

Table 1. Balancedness of border counties

| | All | | | Excluding Hamilton county | | |
	No Law	Law	Diff.	No Law	Law	Diff.
Population density	102.140	137.895	-35.756	102.140	99.660	2.480
	(147.415)	(228.339)	(51.003)	(147.415)	(114.772)	(35.926)
Share urban	0.206	0.319	-0.113	0.206	0.297	-0.091
	(0.261)	(0.309)	(0.076)*	(0.261)	(0.057)	(0.075)
Share foreign-born	0.033	0.051	-0.019	0.033	0.048	-0.015
	(0.037)	(0.057)	(0.013)*	(0.037)	(0.055)	(0.012)
Share German-born	0.017	0.028	-0.011	0.017	0.026	-0.010
	(0.020)	(0.028)	(0.006)**	(0.020)	(0.027)	(0.006)*
Share Lutheran	0.010	0.015	-0.005	0.010	0.016	-0.005
	(0.024)	(0.023)	(0.006)	(0.024)	(0.023)	(0.006)
Observations	29	27		29	26	

Notes: Data are from the 1910 county data in ICPSR and from the 1906 Census of Religious Bodies.

Table 2. Summary statistics: Border dataset

| | Found in 1930 | | | Found in 1940 | | |
	Mean	S.D.	N	Mean	S.D.	N
Married	0.742	0.437	8949	0.841	0.366	10807
Spouse of German ancestry	0.364	0.481	6455	–	–	–
Number of children	1.395	1.733	8996	1.43	1.721	10807
Log average GNI of children	2.685	1.964	4984	2.460	2.201	6396
Log GNI of first son	2.270	2.652	3674	2.400	2.524	4407
Lives in same state as 1920	0.905	0.293	8950	0.833	0.373	10807
Lives in same county as 1920	0.844	0.363	8950	0.666	0.478	10807
Years of education	–	–	–	8.505	2.788	10709
Yearly salary earnings	–	–	–	6.676	2.369	7624

Notes: The table shows summary statistics for males born 1880–1916 to German parents, who in 1920 lived in a county on either side of the border of Indiana and Ohio with Michigan or Kentucky and who were linked to the census of 1930 (left panel) or 1940 (right panel). See Section 3.2 for details on construction of the GNI variables.

Table 3. Most and least German-sounding names in the 1930 census

Highest-scoring				Lowest-scoring			
Name	Total	Germans	GNI	Name	Total	Germans	GNI
Hans	1272	324	96.15	Clyde	7350	0	0
August	5772	1260	95.33	Russell	6045	0	0
Gustave	1270	231	94.33	Melvin	5682	0	0
Karl	1538	268	93.91	Patrick	5367	0	0
Otto	5685	959	93.68	Leroy	5183	0	0
Christian	1214	179	92.67	Warren	5071	0	0
Herman	11423	1398	91.06	Marvin	4585	0	0
Emil	4256	515	90.96	Jim	4226	0	0
Adolph	3225	385	90.83	Glenn	3893	0	0
Conrad	1341	150	90.20	Leslie	3795	0	0

Notes: The table shows the values of the German name index for the 10 highest-scoring (left panel) and 10 lowest-scoring (right panel) names of males in the 1930 5% IPUMS sample. Highest-scoring names are chosen among names that appear at least 1,000 times in the 1930 sample and are ordered by their GNI value; lowest-scoring names are ordered by popularity. See Section 3.2 for details on construction of the GNI.

Table 4. Summary statistics: WWII Enlistments

	All			German parents		
	Mean	S.D.	N	Mean	S.D.	N
Age	30.58	5.014	77213	33.53	5.895	897
Married	0.211	0.408	77081	0.171	0.376	896
With dependents	0.111	0.315	77081	0.142	0.349	896
Volunteer	0.127	0.333	77213	0.077	0.267	897
High school graduate	0.429	0.495	77213	0.328	0.470	897
College graduate	0.063	0.244	77213	0.041	0.199	897

Notes: The table reports summary statistics for males who enlisted in the US Army between 1940 and 1942 and were linked to the 1930 census. The data comprises of cohorts born 1880–1916 in Indiana, Ohio, Michigan, and Kentucky. The right panel restricts the sample to individuals with German parents. Volunteers are identified as having a serial number in the 11 through 19 million series.

Table 5. Baseline results: Border data set

	[1]	[2]	[3]	[4]	[5]
Panel A		Dep. Variable: Log average GNI of children			
Law × CSL age	0.168**	0.347***	0.432***	0.557***	0.393***
	(0.0635)	(0.0253)	(0.0478)	(0.144)	(0.0398)
Observations	8348	8348	8348	8348	5789
R-squared	0.0220	0.0330	0.0414	0.0416	0.0498
Panel B		Dep. Variable: Log GNI of first son			
Law × CSL age	0.279*	0.562***	0.637***	0.294**	0.576***
	(0.132)	(0.0680)	(0.102)	(0.103)	(0.0765)
Observations	5832	5832	5832	5832	4048
R-squared	0.0289	0.0483	0.0633	0.0644	0.0731
Panel C		Dep. Variable: Spouse German			
Law × CSL age	-0.0664	0.0765***	0.0662***	0.0900***	0.0660***
	(0.0476)	(0.00170)	(0.00457)	(0.0115)	(0.00446)
Observations	6455	6455	6455	6455	4366
R-squared	0.0391	0.0483	0.0729	0.0730	0.0954
State Controls × Cohort FE	N	Y	Y	Y	Y
County FE	N	N	Y	Y	Y
State trends	N	N	N	Y	N
Excluding Cincinnati	N	N	N	N	Y

Notes: The sample consists of males, born 1880–1916 in the US to German parents, living in a border county in 1920 and who were linked to the 1930 census (Panel C) or the 1930 and 1940 census (Panels A and B). All regressions include residence state in 1920 and birth cohort fixed effects, and controls for the following name string properties: first and last name length and first and last name commonness. Regressions in Panels A and B include a census year indicator. State controls interacted with birth cohort dummies include the share of Germans in the state in 1910 and a border segment indicator in Panels A and B and log population and log population of first and second generation Germans in the state in 1910 in Panel C. Standard errors are clustered at the state×cohort level. Significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 6. Baseline results: WWII enlistments

Dep. variable:	Volunteer					
	[1]	[2]	[3]	[4]	[5]	[6]
Law × CSL age	-0.0712***	-0.0683***	-0.0897***	-0.0442**	0.0841	-0.0243
	(0.00416)	(0.00651)	(0.00972)	(0.0130)	(0.0812)	(0.0142)
Law × CSL age × German parents						-0.0631***
						(0.00888)
Observations	897	896	896	896	1254	77081
R-squared	0.0742	0.114	0.139	0.140	0.0356	0.0821
Controls	N	Y	Y	Y	Y	Y
Share German in state in 1910 × Cohort FE	N	N	Y	Y	N	N
State trends	N	N	N	Y	N	N

Notes: The sample consists of males born 1880–1916 in Indiana, Ohio, Michigan, or Kentucky, who enlisted in the US Army between 1940 and 1942 and who were linked to the 1930 census. In columns [1]–[4] it is restricted to individuals with German parents and in column [5] to individuals with Italian parents. All regressions include state-of-birth and birth cohort fixed effects and control for the following name string properties: first and last name length and first and last name commonness. Columns [3]-[4] include interactions of the indicator for German parents with state-of-birth and birth cohort dummies. Columns [2]–[6] control for marital status, the number of dependent family members and enlistment year fixed effects. Standard errors are clustered at the state of birth×cohort level. Significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 7. Non-German mothers

	[1]	[2]	[3]	[4]
Dep. Variable:	Log average GNI of children	Log GNI of first son	Spouse German	Volunteer
Law × CSL age	-0.189**	-0.0904	0.0176*	-0.0952***
	(0.0554)	(0.0894)	(0.00765)	(0.0213)
Observations	6548	4536	4491	1041
R-squared	0.0548	0.0776	0.0724	0.105

Notes: The sample consists of linked males, born 1880–1916 in the US to a German father and a non-German mother, living in a border county in 1920 (columns [1]–[3]) or born in Indiana, Ohio, Michigan or Kentucky (column [4]). All regressions include (1920 residence or birth) state and birth cohort fixed effects, and controls for the following name string properties: first and last name length and first and last name commonness. Regressions in columns [1]–[3] include county fixed effects. State controls interacted with birth cohort dummies include the share of Germans in the state in columns [1], [2] and [4], a border segment indicator in columns [1]–[2], and log population and log population of first and second generation Germans in the state in 1910 in column [3]. Standard errors are clustered at the state×cohort level. Significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 8. Educational and labor market outcomes

	Both parents German	Only father German	Both parents German	Only father German
	[1]	[2]	[3]	[4]
Dep. Variable	Years of schooling		Log yearly wage income	
Law × CSL age	0.247*	0.279***	-0.121	-0.0565
	(0.111)	(0.0307)	(0.0770)	(0.108)
Observations	11218	9477	7905	6676
R-squared	0.0952	0.106	0.0234	0.0325

Notes: The sample consists of males, born 1880–1916 in the US to a German father, living in a border county in 1920 and who were linked to the 1940 census. All regressions include residence state in 1920 and birth cohort fixed effects, county fixed effects and controls for the following name string properties: first and last name length and first and last name commonness. When the dependent variable is log yearly wage income, the dataset is restricted to salaried workers. Standard errors are clustered at the state×cohort level. Significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

A   Appendix (For Online Publication)

A.1   A simple model of ethnic identity backlash

This section provides a basic conceptual framework for understanding how language restrictions at school affect the formation of ethnic identity among immigrant children. I construct a simple model of intergenerational transmission of ethnic identity; the model borrows from Bisin and Verdier (2001) and Bisin et al. (2011), and it features both vertical and horizontal socialization. A child's sense of ethnicity is the end product of parental investment and socialization in the school. Using this model, I derive predictions for the effect of a language ban on the younger generation's sense of ethnic identity and for how this effect varies according to the initial strength of parents' ethnic identification.

Modeling the intergenerational transmission of ethnic identity

Consider a population that consists of a majority and a minority group. The two groups are differentiated by some external attribute, which is exogenous to the individual. In the context of our specific case study, this attribute is ethnic — in particular, German — origin. Hence we use subscript G to denote the minority group and subscript N to denote the majority. Within the minority group of German ethnics, there are two types of individuals: "mainstream" ($i = m$) and "oppositional" ($i = o$). Mainstream types are assimilated into the majority Anglo-Saxon culture and follow its norms, whereas oppositional types actively try to maintain their German culture and resist assimilation by the mainstream. Although members of the German minority can be either mainstream or oppositional, the majority group is assumed to consist only of mainstream types.

Families are composed of a parent and a child. Children (marked by superscript c) initially inherit the type or trait, $i = m, o$, of their parents (marked by superscript p); however, they can switch to a different trait after exposure to interactions with peers, role models, or other cultural partners in society. I assume that such "horizontal" socialization occurs in school, which every child attends. In school, the child interacts with teachers and peers and is paired to a role model, who with probability q_i is of a different type than the child's inherited type. If this is the case, then the child switches type with probability $1 - \lambda_i^c$. Here λ_i^c denotes the intensity of the child's identification with his initial type and is the result of parental investment. Since oppositional parents are more likely to feel strongly about the identity of their children, I assume that parameter values are such that $\lambda_o^c > \lambda_m^c$.

After the family, the school is assumed to be the main (and, for simplicity, the only) socialization pool entered by the child. The school's ethnic character — in other words,

the importance it places on German education — thus determines how likely it is that the child will become oppositional later in life. Recall that q_i denotes the probability of the child meeting a role model of type different than her parents. For oppositional children, the probability q_o of meeting a mainstream role model is lower in a school that emphasizes the transmission of German ethnicity (as in, e.g., a school that teaches the German language).

Given the socialization mechanism just described, the probability with which a parent of trait i will end up with a child who shares that same trait is given by

$$P_{ii} = 1 - q_i(1 - \lambda_i^c) \tag{5}$$

Similarly, the probability that a parent of trait i will end up with a child who instead exhibits trait $j \neq i$ is

$$P_{ij} = q_i(1 - \lambda_i^c) \tag{6}$$

Later in life, each child takes an action $a \in \{G, N\}$. Action G ("German") is accepted as appropriate by the German minority, whereas action N ("Native") is compatible with the mainstream values and is taken by an assimilated individual. For instance, one could have $G = \{\text{marrying a German ethnic}\}$ versus $N = \{\text{marrying a native}\}$ or $G = \{\text{giving one's child a German name}\}$ versus $N = \{\text{giving one's child an Anglo-Saxon name}\}$.

I define preferences such that a mainstream minority type always prefers action N to action G, while an oppositional type always prefers G to N. In particular, utility is given by $U_o(G) = a\lambda_i^c + b\overline{\lambda_o} + d$ and $U_m(N) = F$, with $a, b, d, F > 0$. I normalize to 0 the payoff resulting from taking an action not corresponding to one's type.

For minority oppositional types, utility is increasing in the child's intensity of identity. It is also increasing in a social interaction component, denoted by $\overline{\lambda_o}$, which captures the strength of German identity among the child's peers. The intuition for this latter component is straightforward. The utility that the oppositional child derives from a German action is greater if his immediate environment supports that action. A mainstream child derives utility $U_m^c = F > 0$ from action N and 0 otherwise.[26]

---

[26]Mainstream minority children may also enjoy a psychological benefit when interacting with mainstream peers whose intensity of identity is similar to their own. The converse of this mechanism is the "acting white" phenomenon (Austen-Smith and Fryer, 2005), or the psychological cost sustained by minority individuals who do not conform to the norms of their group. Here, I assume that the psychological benefit to mainstream children is negligible when compared with the direct gain F of assimilation. In other words, the benefit of undertaking the action prescribed by one's type is more

Parents are characterized by what Bisin and Verdier (2001) call *imperfect empathy*, a form of paternalistic altruism whereby parents care about their children's action but evaluate it using their own preference parameters. So in my setup oppositional parents whose intensity of identity is λ_o^p will derive utility $U_o^p = a\lambda_o^p + b\overline{\lambda_o} + d$ if their child is oppositional (takes action G) or zero utility if their child is mainstream (takes action N). Conversely, mainstream parents gain utility $U_m^p = F$ if their child is mainstream or zero utility if their child is oppositional. Parents can influence their child's choice of action by undertaking a costly socialization investment in her identity. A stronger ethnic identity will reduce the likelihood of the child abandoning the parental trait for a random role model in school and will also increase the child's incentives to take parentally desired actions later in life.

Assuming investment costs are quadratic, the problem of an oppositional parent can be written as follows:

$$\max_{\lambda_o^c} \quad (1 - q_o(1 - \lambda_o^c))(a\lambda_o^p + b\overline{\lambda_o} + d) - \frac{\lambda_o^{c2}}{2}$$

This expression gives the optimal intensity of identity chosen by the parent as

$$\lambda_o^{c*} = q_o(a\lambda_o^p + b\overline{\lambda_o} + d) \tag{7}$$

Since $\lambda_o^{c*} = \overline{\lambda_o}$ in equilibrium, we can write

$$\lambda_o^{c*} = \frac{q_o(a\lambda_o^p + d)}{1 - bq_o}$$

Note that λ_o^{c*} is increasing in q_o (i.e., in the likelihood of the child meeting a mainstream role model). This relation makes intuitive sense. Parents will invest more in their child's identity when it is threatened — that is, when the child interacts more frequently with role models belonging to a different type.

The model exhibits a unique and stable steady state, in which $\lambda^{ss} = \frac{dq_o}{1 - q_o(a+b)}$. To ensure that the problem has an interior solution, I further assume that $1 \geq q_o(a+b+d)$.

An analogous maximization problem for the mainstream parent yields

$$\lambda_m^{c*} = q_m F$$

Just as with their oppositional counterparts, the investment of mainstream parents is

_____

psychological in nature for oppositional individuals but more tangible in nature for mainstream individuals.

increasing in the payoff of ending up with a mainstream child and in the probability that the child will meet an oppositional role model at school. For both types of parents, the vertical and horizontal transmission of culture are substitutes.

Implications of a language ban in elementary school

A school's ethnic character is determined by the probability q_i that a child is exposed at school to an ethnic trait different to her own. In this context, a German language ban at school corresponds to an increase in q_o and, equivalently, a reduction in q_m (i.e., a reduced likelihood that the child meets an oppositional role model). If we denote the share of oppositional types among the minority by p and the share of the minority in society by s, then we can think of q_o, or the probability of an oppositional child encountering a mainstream type, as $q_o = (1 - \kappa s) + \kappa s(1 - p)$. Similarly, the probability of a mainstream child meeting an oppositional role model can be written as $q_m = \kappa s p$. Here κ represents how likely it is to meet a minority role model at school, so a language law can be thought of as a reduction in κ.

The law sets in motion two opposing forces. Its immediate effect, acting through horizontal socialization, is a reduction in the share of oppositional children; this effect is mediated by a decrease in the relative importance of oppositional role models at school. The second, indirect effect acts through vertical transmission. Given that socialization in the school and in the family are substitutes, oppositional parents will react to the weakening of the school's ethnic socialization function by increasing their own investment in the child's identity. This increase in parental investment can be high enough to counteract the language law's direct effect. Hence we can state the following proposition

Proposition 1 (Backlash). *Starting from the steady state, children of oppositional parents are more likely to become oppositional in response to a language-banning policy if the steady state intensity of oppositional identity is high enough.*

Proof. For a backlash to occur, we need

$$\frac{dP_{oo}}{dq} = -(1 - \lambda_o^c) + q_o \frac{a\lambda_o^p + d}{(1 - bq_o)^2} > 0 \tag{8}$$

Evaluating the above derivative at the steady state, we can rewrite the condition for a backlash as

$$\lambda^{ss} > \frac{1 - bq_o}{2 - bq_o}$$

$\square$

45

To gain some intuition on the content of Proposition 1, it is useful to revisit equation (7), which describes the optimal investment for oppositional parents. That expression makes clear that the law increases parental investment through two channels. The first one is the direct reaction of oppositional parents to the now diminished role of the school, which formerly served as a substitute for their own investment. Hence parents now invest more, which leads to higher λ_o^c (i.e., to a stronger sense of identity) for those children who remain oppositional. In turn, that behavior increases $\overline{\lambda_o}$, the average ethnic identity among oppositional children. Equation (7) shows that this dynamic feeds directly into the parental decision inducing parents to make additional investment in cultural identity. If this amplification effect is strong enough, then the share of oppositional children will actually increase as a result of the school language law. Propositions 3 and 4 posit that such a "backlash" result is more likely if the minority community is small and/or if oppositional parents strongly identify with their type to begin with.

Although the *sign* of the average reaction is indeterminate, the next proposition predicts how each type of minority parent reacts.

Proposition 2. *The difference between oppositional and mainstream parents — with regard to their respective shares of oppositional children — increases in response to a language-banning policy.*

Proof. Using the transition probabilities and the fact that $q_m = 1 - q_o$, we can write this difference as

$$D \equiv P_{oo} - P_{mo} = \lambda_o^c + q_o(\lambda_m^c - \lambda_o^c)$$

Taking the derivative with respect to q_o yields

$$\frac{\partial D}{\partial q_o} = (1 + \frac{1}{1 - bq_o})\lambda_o^c - 2\lambda_o^m > 0$$

where the inequality follows from $\lambda_o^c > \lambda_m^c$ and $1 - bq_o \in (0, 1)$. $\qquad\square$

Proposition 2 shows that a language ban always leads to an increase in heterogeneity within the minority group. Both types of minority parents adjust their identity investment in response to the law, with mainstream parents now investing less and oppositional parents more. However, the combination of a stronger identity and the effect of social interactions for oppositional parents ensures that their reaction will always be more pronounced than that of their mainstream counterparts. As a result, the spread in the shares of oppositional children born to the two types of parents will increase.

The following two statements identify conditions under which the language law's backlash effect will be more pronounced.

Proposition 3. *A backlash from oppositional types is more likely when the share s of the minority community is small.*

Proposition 4. *A backlash from oppositional types is more likely when the initial identity of parents is strong (i.e., when λ_o^p is large).*

Proof. Recall that the condition under which a backlash occurs is $\lambda_o^c > \frac{1-bq_o}{2-bq_o}$. Then proposition 3 follows from the fact that q_o is increasing in s, the left-hand side of the previous expression is increasing in q_o and the right-hand side is decreasing in q_o. For proposition 4, recall that λ_o^c if increasing in λ_o^p. A backlash is more likely if the share of the minority is small or if oppositional types have a stronger identity. $\qquad\square$

The mechanism of cultural distinction elicits higher investment levels from oppositional parents who belong to a small minority. When the share of the minority group is small, the child is unlikely to meet an oppositional role model at school; in this case, the parents must replace horizontal socialization with their personal effort. The result is both high initial λ_o^c and high average identity among oppositional types after the introduction of a language policy. When an oppositional child's utility from social interactions is high, the initial increase in parental investment spurred by the language ban is amplified considerably. This increases the likelihood of parental compensation outweighing the law's first-order assimilation effect and producing a backlash.

The intuition behind Proposition 4 is similar to that behind Proposition 3. In both cases, initial parental investment is high enough to ensure a high average identity — and thus a high utility benefit from social interactions — for oppositional children. These conditions make for a strong amplification mechanism and a more pronounced reaction of oppositional parents to assimilation policies.

A.2 Data Construction

Census record linking

Here I list in more detail the steps I take to link records from the 1920 census to the 1930 and 1940 census years.

1. I begin by using the NYSIIS phonetic equivalent of first and last name, the exact birthplace and the birthyear to assign 1920 records to all possible matches in the target (1930 or 1940) complete-count census dataset. I allow for a band of 2 years around the birthyear recorded in 1920. In cases of multiple matches where only

the birth year differs, I keep the match(es) with the minimum distance between birthyears.

2. In this subset of matched records, I compute the string distance between first and last names in 1920 and 1930 or 1940. I use the Jaro-Winkler algorithm, as it has been implemented in STATA. The algorithm yields a distance measure taking values between 0 and 1, with 1 implying that the strings considered are a perfect match. I sum up the JW measures for first and last names. In cases of multiple matches, I keep those with the minimum value in this composite measure.

3. I next choose a threshold of JW distances below which I drop all matched observations. A higher JW measure implies a better match, but restricting the matched dataset only to matches that have the maximum JW score would imply discarding a number of potentially correct matches, in which e.g. the names are slightly misspelled. To determine a cutoff that balances precision and power, I ignore JW=2, and I find the structural break in the match rate as a function of the JW. Figure A.1 illustrates these breaks for each linking procedure. The slope of the match rate is relatively flat for JW values below the breakpoint and becomes steep thereafter. Intuitively, any increase in the JW above the breakpoint results to a much larger loss in terms of sample size than increases below that value.

4. Because a few of the multiple matches have identical JW distances, I take one final step to filter them. For names with a middle initial, I keep only links for which the middle initial matches between 1920 and a later census year. I then discard all remaining multiple matches.

5. In the few cases where the same record from 1930 or 1940 has been matched to multiple records in 1920, I keep the match with the minimum birthyear difference and discard all remaining multiple matches.

The match rate is approximately 30% for the 1930 and 36.5% for the 1940 census.[27] Table B.3 lists characteristics and name string properties that affect the probability of a successful match. I control for these string properties in all regressions.

It is noteworthy that the characteristics affecting the probability that an observation is linked do not vary systematically across cohorts or between the two sides of a border.

[27]This is roughly comparable to the match rates of previous work linking individuals between censuses. Long and Ferrie (2013) and Collins and Wanamaker (2013) report match rates of about 20%. Abramitzky, Boustan and Eriksson (2016) have a match rate of 35%. Parman (2011) and Feigenbaum (2015) achieve match rates of 50% and 56% respectively, but they use either manual or machine learning matching techniques.

Figure A.1. Jaro-Winkler cutoffs
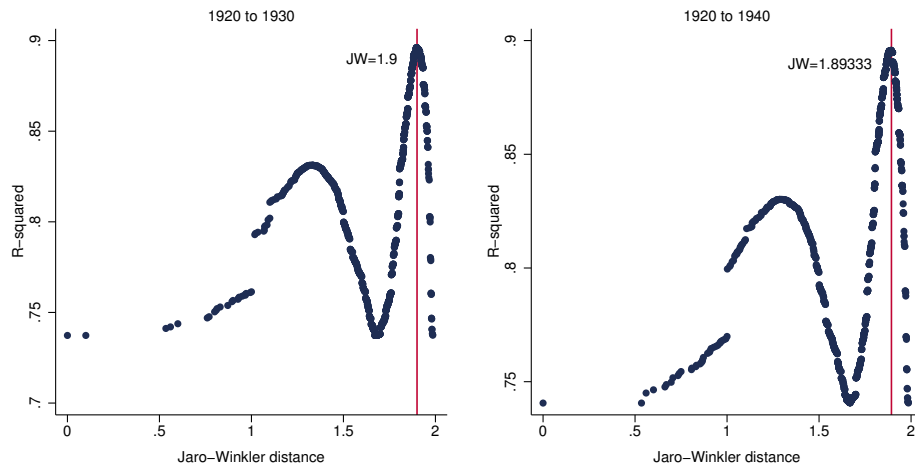
Notes: The figure plots the values of the R-squared from piecewise linear regressions with two segments. A regression is estimated for each value of the Jaro-Winkler distance as breakpoint and the chosen cutoff is the breakpoint value that maximizes the R-squared. The left panel plots this procedure for matches between 1920 and 1930, and the right panel for matches between 1920 and 1940.

Figure A.2. Estimated effect of the language ban on the probability of a match: Border data set
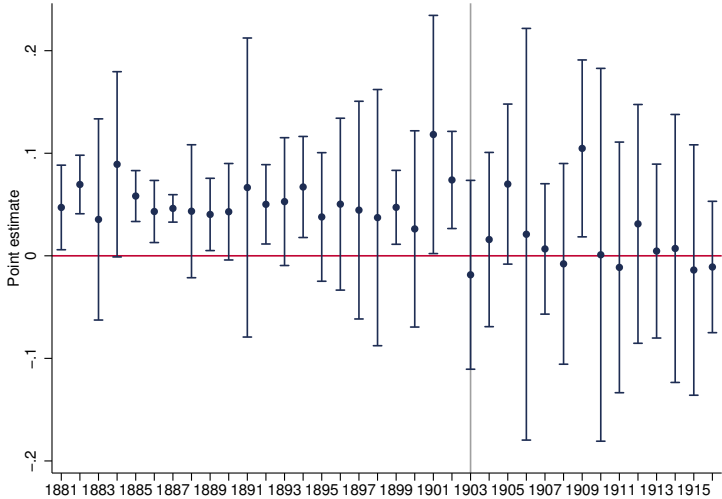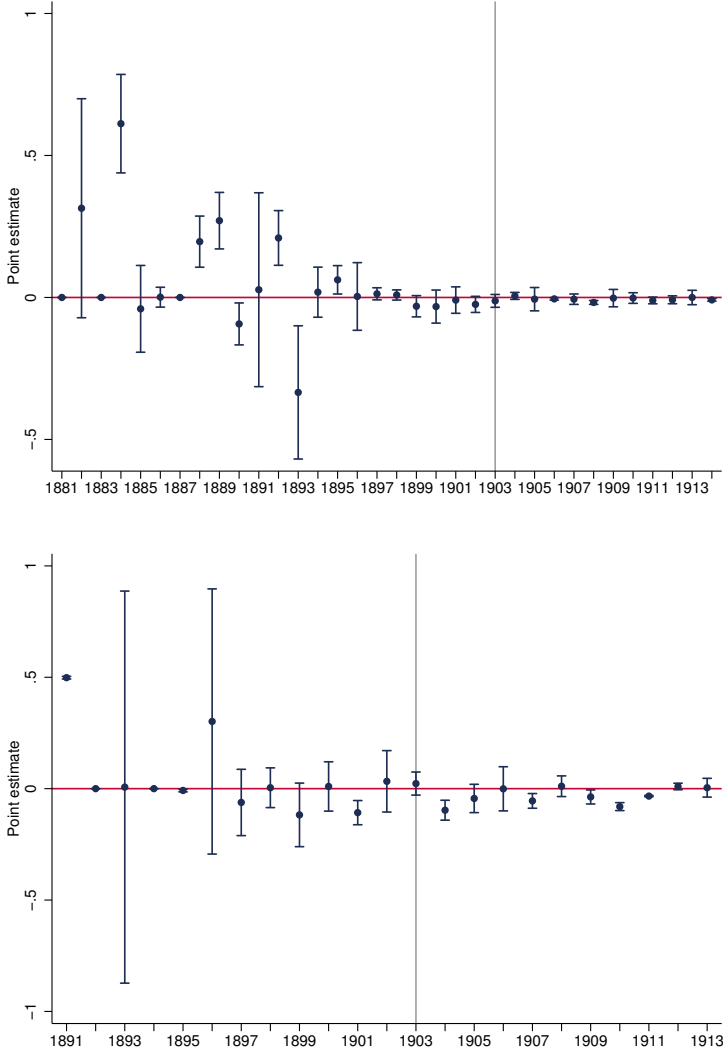

Notes: The figure shows coefficient estimates and 95% confidence intervals from a regression of an indicator for a record that was matched on state and birth cohort fixed effects and a set of interactions of birth cohort fixed effects with an indicator for a language ban. Standard errors are clustered at the state×cohort level. The data consist of all males who were born 1880–1916 in the US to German parents and lived in a border county in 1920. The grey line in year 1903 indicates the first cohort to be affected by the language ban.

Figure A.2 plots interaction coefficients of birth cohort indicators with a dummy for states with a language ban. The dependent variable is an indicator for observations that were matched across censuses. There is somewhat of a downward trend in the match rate throughout the sample period, but no clear indication of a break for treated cohorts.[28] This indicates that there is no systematic difference in the probability of a successful match that could bias the difference-in-differences analysis. Figure A.3 does the same for the entire sample of enlisted men and for volunteers.[29] Volatility in the match rate is high, especially for older cohorts, but there are no systematic differences in the linking probabilities across states and cohorts.

---

[28]The estimated interaction coefficient in a difference-in-differences regression with an indicator for matched records as the dependent variable is small and insignificant (-0.0338 [p-value: 0.134]).

[29]The match rate is approximately 17%. Parman (2015) uses identical criteria to match a sample of WWII enlistees to the 1930 census and reports a match rate of about 8%, after an additional manual inspection of the matched sample.

Figure A.3. Estimated effect of the language ban on the probability of a match: WWII enlistments

Notes: The figure shows coefficient estimates and 95% confidence intervals from a regression of an indicator for a record that was matched on state and birth cohort fixed effects and a set of interactions of birth cohort fixed effects with an indicator for a language ban. Standard errors are clustered at the state of birth×cohort level. The data consist of all enlisted men (upper panel) and volunteers (lower panel) who were born 1880–1916 in in Indiana, Ohio, Michigan, or Kentucky. The grey line in year 1903 indicates the first cohort to be affected by the language ban.

German Name Index

The German Name Index follows Fryer and Levitt (2004) and is constructed as follows

$$\text{GNI}_{\text{name},s} = \frac{\Pr(\text{name}|\text{German}_s)}{\Pr(\text{name}|\text{German}_s) + \Pr(\text{name}|\text{non-German}_s)} \times 100$$

To compute this index, I use information from the 1930 5% IPUMS sample (Ruggles et al., 2010) and define Pr(German) as the share of foreign-born individuals in the census that were born in Germany. I compute this index separately for men and women and drop from the analysis all names that appear fewer than 10 times in the data, so that resulting index values are not driven by rare names.

A.3   Additional Robustness Checks

Migration

One challenge to identification is endogenous sorting across the border. Given that the census nearest to the passage of language legislation is 1920, I do not observe individuals in the data set until *after* the law was enacted. It is conceivable that parents with a strong desire to send their children to a German school could have moved across the border in response to (or in anticipation of) legislation. I would then be identifying the effect of legislation on the selected group of non-movers, and it could be biased in any direction — most likely in one of higher assimilation, since these individuals would be characterized by a weaker ethnic identity to begin with. Since I do not know the migration history of individuals in the years before 1919, I can assess the relevance of sorting only imperfectly: by examining the share of people who were born in a state other than the one in which they are observed in 1920. This share is plotted in the lower panel of Figure A.4. While the two sets of states are clearly on a different trend, the trend break in in-migration happens already in the 1890's and is steeper in Michigan and Kentucky. If the concern is that German families with school-aged children in Indiana and Ohio move out of those states and into their neighboring states in response to the legislation, this does not seem to be reflected in the graph.

Another way in which differential migration could bias my findings is if relatively more assimilated German-Americans (i.e. those with lower endogamy rates) were more mobile and thus more able to migrate out of states that banned German in the schools. Out-migration rates are not directly observable in my data. I assess the possibility that German out-migration rates were higher for younger cohorts in Indiana and Ohio using the 1920 1% IPUMS sample. Figure A.5 plots the share of males with German parents who were born in one of the four states in the data set, but who did not live in the same

Figure A.4. Migration

Notes: The figure compares the share of "movers" (i.e., people who were born in a different state than the one in which they lived in 1920) across birth cohorts and across states with and without a language ban. The data consist of males, born 1880–1916 in the US to German parents, who lived in a border county in 1920 and could be linked to the 1930 or 1940 census.

Figure A.5. Assessing out-migration in 1920
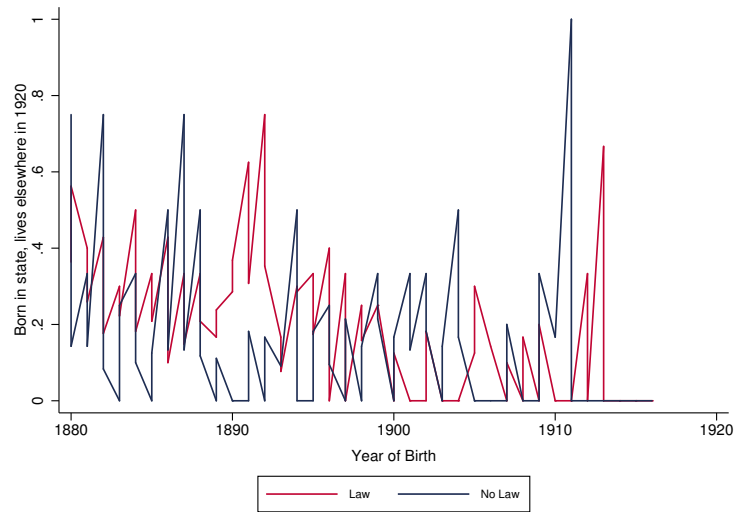
Notes: The figure plots the share of males with German parents born 1880–1916 in Indiana and Ohio (red line) or Michigan and Kentucky (blue line), who lived in a state different from their state of birth in 1920. Data is from the 1920 1% IPUMS sample.

state in 1920. With the exception of a jump in the out-migration rate for cohorts born around 1890 in Ohio and Indiana, this share, though volatile, is similar (and low) for states with and without a law and does not differ for cohorts affected by the language ban.

As an additional check, I estimate my baseline specification by dropping all "movers", i.e. individuals not born in the state in which they are observed residing in 1920. Estimated coefficients remain fairly stable or become larger, suggesting that migration is not the primary driver of the observed effects of the language ban.

A.4 Countrywide evidence

I use the 1930 5% and 1960 1% IPUMS samples (Ruggles et al., 2010) to examine how language laws affected the assimilation of German-Americans in the country as a whole.[30] I focus my attention on cohorts born to a German father between 1880 and 1916. Data on English-only laws are from Edwards (1923), Hood (1920) and Ruppenthal (1919), who provide references to all language-related legislation enacted in the United States until 1919.

I match individuals in the census with legislation enacted in their state of birth, so that exposed cohorts are those born in a state with an English-only law and in school according to the compulsory schooling law at the time that the language law was in effect. In the 1960 census, a person's foreign-born wife is coded as "foreign-born" and without any details about her particular birthplace. For this reason, in the 1960 sample I can distinguish only between whether an individual's mother is foreign or native; I am unable to determine whether she is German. I estimate

$$Y_{isc} = \alpha + \beta T_{cs} + \lambda_c + \theta_s + \varepsilon_{isc} \tag{9}$$

where T_{cs} is an indicator for individuals born in a state with a law and who were within the age range for compulsory schooling at the time that law was in place. The terms λ_c and θ_s signify cohort and state of birth/residence fixed effects.

Column [1] of Table A.2 reports the estimates of a regression like equation (9) in the pooled 1930 and 1960 IPUMS samples of second-generation German males born 1880–1916 to German parents. Column [2] controls for a linear state-specific trend and column [3] adds interactions of baseline state characteristics with birth cohort fixed effects. These are: the share of first- and second-generation Germans in the state

---

[30]I can only observe the ethnic background of the spouse of a native person in 1930 and 1960.

Table A.1. Dropping movers

	[1]	[2]	[3]	[4]	[5]
Panel A	Dep. Variable: Log average GNI of children				
Law × CSL age	0.167*	0.358***	0.450***	0.445**	0.357***
	(0.0734)	(0.0193)	(0.0308)	(0.167)	(0.0584)
Observations	7102	7102	7102	7102	4765
R-squared	0.0261	0.0386	0.0473	0.0474	0.0596
Panel B	Dep. Variable: Log GNI of first son				
Law × CSL age	0.333*	0.604***	0.684***	0.340	0.558***
	(0.150)	(0.0567)	(0.0614)	(0.189)	(0.0688)
Observations	4947	4947	4947	4947	3319
R-squared	0.0354	0.0560	0.0686	0.0701	0.0839
Panel C	Dep. Variable: Spouse German				
Law × CSL age	-0.0737	0.114***	0.103***	0.114***	0.104***
	(0.0591)	(0.00574)	(0.00869)	(0.0119)	(0.00901)
Observations	5485	5485	5485	5485	3606
R-squared	0.0398	0.0513	0.0764	0.0764	0.0996
State Controls × Cohort FE	N	Y	Y	Y	Y
County FE	N	N	Y	Y	Y
State trends	N	N	N	Y	N
Excluding Cincinnati	N	N	N	N	Y

Notes: The sample consists of males, born 1880–1916 in the US to German parents, living in a border county in 1920 and who were linked to the 1930 census (Panel C) or the 1930 and 1940 census (Panels A and B). All regressions include residence state in 1920 and birth cohort fixed effects, and controls for the following name string properties: first and last name length and first and last name commonness. Regressions in Panels A and B include a census year indicator. State controls interacted with birth cohort dummies include the share of Germans in the state and a border segment indicator in Panels A and B and log population and log population of first and second generation Germans in the state in 1910 in Panel C. Standard errors are clustered at the state×cohort level.
Significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table A.2. Endogamy in IPUMS

	Both parents German			Only father German		
	[1]	[2]	[3]	[4]	[5]	[6]
Law × CSL age	0.0233*	0.0270**	0.0207*	0.0111	−0.0126	−0.00884
	(0.0117)	(0.0123)	(0.0108)	(0.0175)	(0.0205)	(0.0196)
Observations	33432	33432	33432	17385	17385	17385
R-squared	0.0920	0.0948	0.0983	0.0649	0.0695	0.0770
Residence state FE	Y	Y	Y	Y	Y	Y
State of birth trends	N	Y	Y	N	Y	Y
State of birth controls × Cohort FE	N	N	Y	N	N	Y

Notes: Reported values are derived from a linear probability model. Regressions are estimated in the pooled 1930 5% and 1960 1% IPUMS samples; standard errors (reported in parentheses) are clustered at the state-of-birth level. The dependent variable is an indicator for a spouse that is German-born or has either parent German. The sample consists of males born in the United States (excluding Hawaii, Alaska, and the District of Columbia) during the period 1880–1916 to a German father. German mothers are identified as born in Germany in 1930 or as being foreign-born and married to a German-born spouse in 1960. All regressions contain both state-of-birth and birth cohort fixed effects (FE) as well as a census year indicator. Columns [2], [5], and [8] control for a linear trend specific to the state of birth. The state-of-birth controls that are interacted with birth cohort dummies in columns [3], [6], and [9] include the share of the German stock, the sex ratio, and the 1910 sex ratio among first- and second-generation Germans in the state of birth.
Significance levels: *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$.

in 1910, the sex ratio (computed as the ratio of males to females) in the state and the same ratio among first- and second-generation Germans. The magnitude of the estimated coefficient varies little across specifications. Being in a potentially affected cohort is associated with roughly a 2 percentage point (p.p.) increase in the probability of having a German spouse; compared with the 37.5% average endogamy rate in the entire sample of males with two German parents, this effect is small, but not negligible.

Results are less clear for individuals born to mixed couples of German fathers and non-German mothers. Columns [4]–[6] report a non significant effect of language laws on endogamy for this group; that effect becomes negative with the inclusion of state trends and becomes zero when the specification includes interactions between cohorts and state controls.

Table A.2 verifies that the results for Indiana, Ohio and their neighboring states go through for the country as a whole. Language policies seem to have no assimilation effect — at least as measured by intermarriage rates. On the contrary, such policies

lead to increasing endogamy for the more German group and to increasing variance in endogamy rates within the larger German population.

B Additional Tables

Table B.1. Summary statistics: Non-German mothers, border sample

	Found in 1930			Found in 1940		
	Mean	S.D.	N	Mean	S.D.	N
Married	0.634	0.482	7309	0.814	0.389	9286
Spouse of German ancestry	0.268	0.443	4491	–	–	–
Number of children	1.133	1.619	7349	1.333	1.624	9286
Log average GNI of children	2.601	2.012	3380	2.373	2.308	5286
Log GNI of first son	2.136	2.762	2474	2.310	2.600	3614
Lives in same state as 1920	0.902	0.367	7315	0.833	0.372	9286
Lives in same county as 1920	0.840	0.367	7315	0.666	0.471	9286
Years of education	–	–	–	9.147	2.987	9187
Yearly salary earnings	–	–	–	6.676	2.342	6513

Notes: The table shows summary statistics for males born 1880–1916 to a German father and non-German mother, who in 1920 lived in a county of Indiana (IN) and Ohio (OH) that bordered Michigan (MI) or Kentucky (KY) and who were linked to the census of 1930 (left panel) or 1940 (right panel). See Section 3.2 for details on construction of the GNI variables.

Table B.2. Summary statistics: Non-German mothers, WWII Enlistments

	Mean	S.D.	N
Age	31.58	5.328	1042
Married	0.185	0.389	1041
With dependents	0.103	0.304	1041
Volunteer	0.105	0.306	1042
High school graduate	0.424	0.494	1042
College graduate	0.058	0.235	1042

Notes: The table reports summary statistics for all males who enlisted in the US Army between 1940 and 1942 and were linked to the 1930 census. The sample comprises of cohorts born 1880–1916 in Indiana, Ohio, Michigan, and Kentucky, to a German father and a non-German mother. Volunteers are identified as having a serial number in the 11 through 19 million series.

Table B.3. Characteristics affecting the probability of a match: Border data set

	Both parents German			Only father German		
	1920 data set	Found in 1930	Found in 1940	1920 data set	Found in 1930	Found in 1940
Demographics						
Mother native born	—	—	—	0.842	0.932	0.933
				(0.364)	(0.251)	(0.250)
Born in different state	0.149	0.139	0.139	0.128	0.121	0.120
	(0.356)	(0.346)	(0.346)	(0.334)	(0.327)	(0.325)
Name string properties						
First name length	5.613	5.633	5.623	5.716	5.746	5.746
	(1.419)	(1.195)	(1.188)	(1.433)	(1.207)	(1.195)
Last name length	6.672	6.784	6.631	6.789	6.749	6.653
	(1.937)	(1.871)	(1.795)	(1.985)	(1.870)	(1.830)
First name commonness	5.047	4.978	5.424	4.449	4.417	4.729
	(5.132)	(4.920)	(5.036)	(4.950)	(4.724)	(4.830)
Last name commonness	0.156	0.074	0.205	0.148	1.161	1.198
	(0.804)	(0.415)	(0.938)	(0.763)	(0.780)	(0.874)
Observations	30987	9262	11323	27999	10451	9578
Match rate		29.9%	36.5%		37.3%	34.2%

Notes: The table reports means and standard deviations (in parentheses) for several characteristics of the border data set. Columns [1] and [4] refer to all males born 1880–1916 (respectively) to German parents and to a German father and a non-German mother, who lived in a border county in 1920. Columns [2]–[3] and [5]–[6] refer to the part of these data sets that could be linked to the census of 1930 and 1940. "Name commonness" is computed as the share of people in the 1920 1% IPUMS sample with the same first or last name (multiplied by 1,000).

Table B.4. Characteristics affecting the probability of a match: WWII enlistments

	All enlistments	Records found in 1930 census
Name string properties		
First name length	5.697	5.788
	(1.291)	(1.208)
Last name length	6.363	6.565
	(1.730)	(1.670)
First name commonness	3.781	2.899
	(4.900)	(4.206)
Last name commonness	0.549	0.410
	(1.499)	(1.224)
Observations	460835	77213
Match rate		16.7%

Notes: The table reports means and standard deviations (in parentheses) for several of characteristics of the WWII enlistments data set. The first column refers to all males— born 1880–1916 in Indiana, Ohio, Michigan, or Kentucky— who enlisted in the US Army between 1940 and 1942. The second column refers to the part of this data set that could be linked to the 1930 census. "Name commonness" is computed as the share of people in the 1930 1% IPUMS sample with the same first or last name, (multiplied by 1,000).