Cultural Assimilation during the Age of Mass Migration*

Ran Abramitzky
Stanford University and NBER

Leah Boustan
UCLA and NBER

Katherine Eriksson
UC Davis and NBER

[INCOMPLETE DRAFT]

June 2015

We study cultural assimilation during the Age of Mass Migration to the US through an analysis of immigrants' naming practices. Using newly-assembled data on immigrants and their US-born children from complete-count Census data, we find that immigrant parents gave sons and daughters less foreign-sounding names after they spent more time in the US, erasing around 25 percent of the name gap with natives after 20 years. Similar patterns hold by birth order. Immigrants' name assimilation was equally rapid regardless of their socioeconomic status (proxied by literacy and home-ownership), but immigrants with foreign last names or who lived in an immigrant enclave shifted name choice more rapidly. We link brothers from their childhood households in 1920 to adulthood in 1940 and find that brothers who received more foreign names completed fewer years of schooling, earned less, and were more likely to be unemployed compared to their siblings. These results hold even for brothers born only a year apart or who received similarly foreign names at birth that diverged over time. We also show that immigrants assimilated culturally along other dimensions, including inter-marriage, application to US citizenship, and ability to speak English.

I. Introduction

We study the cultural assimilation of European immigrants in the United States during the Age of Mass Migration (1850-1913). Almost 30 million immigrants moved to the US during this era of open borders to European immigration. The foreign-born share of the US population remained around 15 percent throughout this period, a level only reached again in 2010. As early as 1890, Congress sought to restrict immigrant entry and finally succeeded in passing strict immigration quotas in the early 1920s.[1]

In a previous paper, we studied immigrant assimilation into US labor markets in this period; we found that the average immigrant held an equally-paid occupation to natives upon first arrival, and that immigrants and natives experienced a similar pace of occupational upgrading over time (Abramitzky, Boustan and Eriksson, 2014). Yet, much like today, public support for immigration restriction was based primarily on the perception that immigrants failed to assimilate into US society, instead maintaining their distinct cultural norms, continuing to speak foreign languages and live in enclave communities.[2] Senator Henry Cabot Lodge, a leading advocate for border closure, argued in 1891 that new immigrant groups "are from races most alien to the body of the American people," and that immigration is "bringing to the country people whom it is very difficult to assimilate and who do not promise well for the standard of civilization in the United States—a matter as serious as the effect on the labor market."

---

[1] Congress passed a literacy test for entry to the US in 1917 and a set of country-specific quotas that favored northern and western European countries in 1921 (modified in 1924). Goldin (1994) reviews the political economy of this legislation.

[2] Opponents of open immigration in the early twentieth century relied on cultural arguments to defend border restriction (Higham, 1955; Jones, 1992, chapter 9; Fetzer, 2000; King, 2000). Recent survey evidence suggests that, again today, cultural concerns outweigh economic factors in determining individual attitudes towards immigration policy (Citrin, Green, Muste and Wong, 1997; Hainmueller and Hiscox, 2007).

Despite the claims of contemporary observers, we know little about the extent and speed of cultural assimilation during this period. This paper aims to fill this gap. We examine a series of measures of cultural assimilation, including rates of inter-marriage, application for US citizenship and facility with English, with a focus on the names that immigrant parents select for their children as they spend more time in the US. We also follow the children of immigrants from their childhood households into adulthood in order to study the consequences of having received a more/less foreign name on subsequent labor and marriage market outcomes.

Standard measures of cultural assimilation reveal a strong tendency toward assimilation both across and within generations. Only 28 percent of first-generation immigrants who married after arrival partnered with a US native or a first- or second-generation immigrant from another sending country; this rate of inter-marriage increased to 53 percent by the second generation. By 1930, over 66 percent of first-generation immigrants had applied for US citizenship and almost all first-generation immigrants reported some ability to speak English. However, the extent of cultural assimilation varied by country-of-origin, with notably lower assimilation rates for the Finns, Italians, and Portuguese, and to some extent the Russians (some of whom were Jewish).

The names that parents select for their children offer another useful (and relatively unexplored) measure of cultural assimilation. Naming is a pure choice, unconstrained by financial limitations or by discrimination on the part of natives. Furthermore, parents often had the opportunity to select names for many children, allowing us to trace out an "assimilation profile" with time spent in the US. Marriage, on the other hand, was usually an once-in-a-lifetime event, determined not only by who you wanted to marry but also by who wanted to marry you. Naming patterns are correlated with the other measures of cultural assimilation and have been used in a number of other contexts as a measure of social distance (Zelinsky, 1970;

2

Lieberson and Bell, 1992; on African-Americans, see Lieberson and Mikelson, 1995 and Fryer and Levitt, 2004; on Hispanics, see Sue and Telles, 2007; and on immigrants to Europe, see Algan, et al., 2013 and Gerhards and Hans, 2009). To measure the empirical foreignness of each first name, we use the newly-digitized complete-count historical US censuses to assign the relative probability that a particular name was held by a foreigner (versus a native) in various years.

In a time series of naming patterns, we find that immigrant parents selected increasingly foreign-sounding names for children born between 1900 and 1915, as immigrant sending countries shifted toward Southern and Eastern Europe. These swings in naming practices were sizeable enough to be culturally meaningful (for example, they were more than half as large as the notable adoption of distinctively black names among the African-American community in the late 1960s and early 1970s).[3]

We then analyze the process of name assimilation within households, using the millions of cases available in the 1920 complete-count US Census. Immigrant parents chose less foreign names for their children as they spent more time in the US, erasing around 25 percent of the gap in name choice with natives after 20 years in the US. We interpret this pattern as being driven some combination of decisions to stay in the US (rather than return to the home country), learning about US culture, and changes in attitudes about integrating into US society. We show that name shifts were not driven by the convention of naming first-born sons after their parents or grandfathers by controlling directly for children who share a name with parents and by documenting a decline in name foreignness at later steps down the birth order (e.g., second to

---

[3] The name of the average child of immigrants increased from 46 to 53 on the Foreignness Index in fifteen years (Figure 1). Fryer and Levitt (2004, Figure 3) implies that the name of the average black child increased by 12 points in a similar period.

3

third son). Furthermore, such birth order effects are not present for children who were themselves born abroad or for children of native-born parents.

Eschewing foreign names was a form of assimilation open to households regardless of literacy or wealth (as proxied by homeownership). Furthermore, it appears that parents were just as likely to shift their name choices for sons and daughters. We document substantial heterogeneity in the speed of name-based assimilation across country-of-origin groups, with groups that start out more distant from the native norm (as proxied by the foreignness of their last names or residence in an immigrant enclave) assimilating more quickly.

Finally, we study the consequences of receiving a foreign name on later labor and marriage market outcomes. Names encode a signal of ethnic identity that could augment self-identification with one's ethnic group or could attract discrimination from teachers and employers. We link over one million immigrant brothers across historical Censuses from their childhood families in 1920 into adulthood in 1940, and find that brothers with more foreign names completed fewer years of schooling, received lower annual earnings, and were more likely to be unemployed. This pattern holds when comparing brothers who were born within one or two years of each other (in order to hold constant family resources), or when identifying the effect of name foreignness from changes in naming trends between a child's birth and his labor market entry.

II. Related literature

Our paper contributes to a number of literatures, including scholarship on immigrant assimilation, ethnic and racial discrimination in the labor market, and the inter-generational transmission of cultural traits.

In economics, studies of immigrant assimilation have mainly focused on labor market outcomes – in particular, whether immigrants' occupations and earnings converge to those of natives with time spent in the destination.[4] In a recent study, we show that initial gaps between immigrants and natives varied substantially by sending country, with some immigrants starting out ahead of natives and others starting out behind. Yet, on average, immigrants held similar occupations to natives upon arrival and, regardless of starting point, immigrants moved up the occupational ladder at the same rate as natives, preserving the size of initial gaps over time (Abramitzky, Boustan and Eriksson, 2014).[5]

There is a long tradition of studying cultural assimilation in sociology. One important indicator of cultural assimilation is the rate of inter-marriage between immigrants and US natives or members of other immigrant groups (Gordon, 1964; Lieberson and Waters, 1988). Pagnini and Morgan (1990) document high levels of endogamy, especially for first-generation immigrants from southern and eastern Europe in 1910 (see also Angrist, 2002). Alba and Golden (1986) show that in-group marriage for the descendants of European immigrants declined across birth cohorts, primarily because the children of mixed parentage were themselves very likely to intermarry.

With time spent in the US, European immigrants also began to resemble natives in various social behaviors, including age at first marriage, completed family size, political participation and criminality, but often this transition took more than one generation (Watkins,

---

[4] See Chiswick, 1978; Borjas, 1985 and Lubotsky, 2007 for discussions of labor market assimilation in the contemporary period and the associated methodological issues.
[5] Earlier work on labor market assimilation in the Age of Mass Migration found that immigrants held substantially lower-paid occupations than natives upon first arrival, but that they converged with natives after spending some time in the US (Hatton, 1997; Hatton and Williamson, 1998, chapter 7; Minns, 2000). Differences are primarily due to the use of cross-sectional vs. panel datasets.

1994; Guinnane, Moehling and O'Grada, 2006; Foley and Guinnane, 1999; Shertzer, 2013; Moehling and Piehl, 2009, 2014). Watkins and London (1994) is the only study we know of to consider naming practices during the Age of Mass Migration; they document limited convergence between second-generation Italian and Jewish immigrants and natives of native parentage.[6] A related contemporary literature finds that immigrants' attitudes converge with those of natives along various dimensions, including political preferences and gender norms, but that this convergence takes time (Fernandez and Fogli, 2009; Alesina et al., 2011; Luttmer and Singhal, 2011; Blau, et al., 2013).

Our paper also contributes to a growing literature using names to document discrimination against ethnic and racial groups in the labor market. Goldstein and Stecklov (2015) show that, during the Age of Mass Migration, men with foreign names received lower occupation-based earnings, even after controlling for the average socio-economic status of fathers. Other work has assessed the economic return received by immigrants who change either their first or last name (Arai and Thoursie, 2009; Biavaschi, Giulietti and Siddique, 2013; Carneiro, Lee and Reis, 2015); the effects of having a distinctively African-American name on employment, education and health outcomes (Bertrand and Mullainathan, 2004; Fryer and Levitt, 2004; Figlio, 2005); or the consequences of having an identifiably ethnic surname among children in inter-ethnic marriages (Rubinstein and Brenner, 2014). Immigrants' earnings appear to increase substantially following a name change, suggesting that workers can avoid ethnic discrimination in the labor market by weakening the signal of their ethnicity. This finding is consistent with experimental evidence from Bertrand and Mullainathan (2004), who demonstrate

_____

[6] In related historical work, Hacker (1999) uses name choice as a measure of secularization (primarily among the native born), documenting a decline in biblical names over the nineteenth century and a positive association between biblical naming and family size.

that resumes attached to particularly white names receive more interview requests than do identical resumes assigned a distinctively black name.[7]

III. Data and definitions

A. Measuring the foreignness of given names

Naming practices provide a useful measure of cultural assimilation. As immigrants spend more time in the US, they learn more about US culture, including which names are currently popular among the native born. Furthermore, immigrants who hope to assimilate into US society might select more American or "native-sounding" names, compared to immigrants who plan to remain within an immigrant enclave or eventually return to their home country. Name choices are free from financial constraint or from the discriminatory barriers imposed by natives in other dimensions of assimilation (e.g., marriage or neighborhood location). Names of all children who remain in their parental household can be observed using historical Census data, which disclose the names of all respondents after 72 years. For parents with multiple children, we can trace out an assimilation profile based on the names bestowed to children born soon after arrival or after more time spent in the US.

To develop a systematic measure of name foreignness, we use the complete-count 1940 US census to calculate R, the relative probability that a given name was held by a foreigner (versus a native). This measure has a natural interpretation; a relative probability of two means that a name is twice as likely to be used in the immigrant population as in the native population,

---

[7] In contrast, Fryer and Levitt (2004) show that, after controlling for family background, having a blacker name is not associated with poorer adult outcomes. Race is highly observable; therefore, even if black workers with racially ambiguous names garner more interviews, they might be rejected by employers later in the hiring process.

and a relative probability of 0.5 means the name is twice as likely to be found among natives as among immigrants.[8]

However, because the relative probability is sensitive to outliers (especially to names that are unpopular among natives, which produces small values of the denominator), we instead rely primarily on a related index used by Fryer and Levitt (2004) in the context of distinctly black names. In particular, the Foreignness Index is defined as:

$$Foreignness\ Index_{name} = \frac{\frac{\#\ foreigners_{name}}{total\ \#\ foreigners}}{\frac{\#\ foreigners_{name}}{total\ \#\ foreigners} + \frac{\#\ natives_{name}}{total\ \#\ natives}}$$

and ranges from zero to one, with a value of zero reflecting the fact that no men in the US with a given first name were foreign born (i.e., a distinctively native name) and a value of one assigned to a child whose first name is distinctively foreign. Note that the F-index is a simple function of R, equivalent to $R/(1-R)$. We discuss robustness to a number of alternative specifications of the F-index in Section IV.

The foreignness of a name can change over time with shifts in the naming practices of either natives or immigrants. Therefore, to capture the foreignness of a name in the year it was given, we calculate a child's Foreignness Index based only on individuals born up to twenty years before the child in question. Table 1 presents the most foreign, neutral and native names for the birth cohorts of 1900-20. Neutral names like David, George and Jeffrey were equally common among the children of foreign-born and native parents. The most foreign names in this period included Spanish and Portuguese names like Pedro and Francisco, Scandinavian names

---

[8] The formula of R is given by: $\quad R = \frac{\frac{\#\ foreigners_{name}}{total\ \#\ foreigners}}{\frac{\#\ natives_{name}}{total\ \#\ natives}}$

like Eric, and German names like Kurt. Very native names were often surnames used as first names like Troy and Nolan.

Figure 1 graphs the Foreignness Index separately for the sons of immigrant and native-born fathers in the birth cohorts of 1850 through 1920, as observed in the 1920 complete-count Census.[9] The Foreignness Index of the children of immigrants declines slightly from 0.51 to 0.46 for the birth cohorts of 1850 to 1900, before increasing to 0.53 by the birth cohort of 1915. The increase in F-index from 1900 to 1915 coincides with a shift in immigrant sending countries towards Southern and Eastern Europe. The gap in F-index between the children of foreign- and native-born fathers is maximized in 1910 at around 0.2. Native households used increasingly native-sounding names, reducing their average F-index from 0.39 in 1870 to 0.32 by 1910, perhaps as a means of differentiating their offspring from the children of foreigners. The shift in native naming occurred despite a growing number of second-generation immigrant parents in the native-born category.

We tested but did not find any evidence for breaks in naming trends following key political events in sending countries. For example, German-Americans faced increasing discrimination during World War I, and so we may expect German parents to respond by giving their children less identifiably German names (Moser, 2012). However, we find no trend break in naming practices in German households during World War I at the national level (see Appendix Table 1). Fouka (2015) shows that German immigrants in states that introduced anti-German

---

[9] We graph the F-index calculated from the 1920 (rather than the 1940) complete-count Census here so that we can extend the series back to the birth cohort of 1850. The F-index for this cohort is calculated from individuals born between 1830-49, who were already 71-90 years old in 1920. For younger birth cohorts, the series calculated from the 1920 and 1940 Censuses are nearly identical; any differences would be due to mortality, name changes, or in- or out-migration between 1920 and 1940.

language policies during the war responded by choosing visibly German names, perhaps as a show of community support.

Figure 2 offers the first evidence of assimilation in naming patterns with time spent in the US, graphing the distribution of name foreignness for the children living at home with native-born parents or with foreign-born parents who had been in the US less than/more than 10 years by the 1920 Census. The distribution of names bestowed by native parents is shifted to the left, dropping off substantially after an index value of 0.6. Within the set of foreign-born parents, the names given by recent arrivals can be easily distinguished from the names given to more long-standing immigrants. Recent immigrants are far more apt to give names with an index value above 0.6.

B. Creating a linked Census sample: 1920-40

[TO ADD, ALONG WITH COMPARISON OF MATCHED SAMPLE TO POPULATION]

IV. Household name choice in the 1920 Census

This section explores the naming choices of immigrant parents as they spend more time in the US. We interpret changes in naming with exposure to US society as evidence of cultural assimilation. However, parental time in the US is highly correlated with a child's rank in the birth order, which could in itself influence naming patterns (e.g., in some cultures, first-born sons are often named after their fathers or grandfathers).[10] We directly control for whether a son has

---

[10] For examples, see details of the Irish and Italian naming traditions here: http://homepages.rootsweb.ancestry.com/~cregan/patterns.htm and

10

the same name as his father and test for oldest son/daughter effects. Furthermore, we compare the children of first-generation immigrant parents to the children of second and third-or-higher generation parents, who may have been subject to the same naming traditions but should not be undergoing cultural assimilation over time, and to children born abroad. The section ends with a consideration of heterogeneity in the speed of name-based assimilation by country-of-origin, household wealth/literacy, and residence in an immigrant enclave.

In our main analysis, we estimate an assimilation profile relating parental name choice to time spent in the US. Using data from the complete-count 1920 Census, we run the following two regression specifications:

$$F\ Index_{ij} = \gamma_{ij} + \beta_2 X + \alpha_j + \varepsilon_{ij} \tag{1}$$

$$F\ Index_{ij} = \beta_1 Birth\ Order_{ij} + \beta_2 X_{ij} + \alpha_j + \varepsilon_{ij} \tag{2}$$

where the F-index (x 100) is the name Foreignness Index of child i in household j at birth. Equation (1) is estimated for children in households with a foreign-born household head. In this case, the main right-hand side variables are a vector γ_{ij} of indicators for parental years in the US by the date of child i's birth in household j. Equation (2) replaces this vector with a linear index of birth order rank among sons (or daughters) observed in the household, which allows us to compare households with foreign- and native-born heads. However, birth order may be a weaker proxy for exposure to US culture because it does not incorporate variation in timing of first birth and spacing between subsequent births. The sample includes non-black children aged 0-18 who were born in the US and are currently living with at least one parent in a non-southern state. We

http://www.italiangenealogy.com/articles/italian-genealogy/9-italian-naming-traditions-and-their-ramifications.

further restrict our attention to children whose mother was less than 43 years old in the 1920 Census; older mothers are more likely to have children who already left home, obscuring a child's place in the full birth order.

All regressions include household fixed effects and a series of dummy variables for the year of child i's birth.[11] With household fixed effects, the effect of parental time in the US (or: birth order) is identified by differences between siblings born after their parents spent more/less time in the US. In some specifications, we also control for other features of a child's name, including an indicator for whether he has the same name as his father, a measure of name frequency/commonness and indicators for whether the name is a saint or biblical name.

Figure 3 illustrates our research design with the example of the XX family observed in the 1920 Census manuscript. Parents XX and XX were born in country XX in XX and came to the US in XX. After arrival, the XX family had XX sons (names here) and XX daughters (names here). XX years in the US is associated with an XX change in the F-index. The XX family seems to have assimilated into US culture based on the name-based assimilation measure.

Figure 4 reproduces the pattern observed in the XX family for the full sample of households with two foreign-born parents in the 1920 Census. We present estimates of the effect of parental years in the US on child's F-index from equation (1), relative to the omitted category (0-3 years in the US). Immigrant parents gave both their sons and daughters less foreign names as they spent more time in US, consistent with a process of cultural assimilation. Sons and daughters born after their parents had spent over 20 years in the US scored 4-5 points lower on the Foreignness Index relative to their siblings born upon their parents' first arrival. The mean

---

[11] Child i's year of birth and place in the birth order are separately identified because households start their childbearing in different years. In some versions of Equation 1, we also control for child's rank in the birth order.

gap in the F-Index for the children of immigrants and natives in the 1920 Census was around 20 points, implying that immigrants closed 25 percent of this "cultural" gap with natives after spending some time in the US.

It is notable that parents shift their naming behavior at an equal pace for sons and for daughters. If name shifts reflect parental learning about US culture and/or interest in joining more integrated communities, we might not expect a differential response by child's gender. However, given the historical gender gap in labor force participation, we might expect larger name shifts for sons as parents become more aware of potential labor market discrimination by ethnicity. Alternatively, we might expect larger name shifts for daughters, given that parents are more open to novel names for girls, while boys tend to receive a more traditional set of names (Rossi, 1965; Sue and Telles, 2007).

Thus far, we have shown that immigrant parents shift to less foreign names as they spend more time in the US. Such a pattern may arise absent cultural assimilation if parents adhere to ethnic naming traditions, such as the practice of naming first-born sons after their father or grandfathers, who, by definition, are more likely to have names of European origin. We consider this possibility by comparing the children of foreign- and native-born parents, or children who were themselves born abroad.

The first panel of Figure 5 compares the naming patterns of US-born children in different household types, as well as the names of children who were themselves born abroad. We report the implied difference in name foreignness between the first and fourth child in the birth order controlling for other name attributes, based on estimates of equation (2). For the children of immigrant parents, moving from the first-born to fourth-born son lowers name foreignness by 1.5-2.0 points. The estimated change is smaller than in the parental time in US specification,

presumably because birth order does not capture all of the relevant variation in birth timing and spacing. There is no such effect of birth order on name foreignness for the sons of immigrants who were themselves born abroad (and a much smaller effect for daughters), suggesting that the observed pattern is not due to naming traditions imported from the home country. Similarly, the relationship between birth order and name foreignness is an order of magnitude smaller for the children of third-or-higher generation Americans (0.3 points on the F-index, or less, and statistically different from the estimates for immigrant parents).

We do find a sizeable effect of birth order on name foreignness for the children of second-generation immigrants. Yet, we would not expect second-generation immigrants who themselves were born in the US to become more aware of US culture at the time of their second or third birth. However, second-generation immigrants, who are themselves the children of immigrants, may continue to adhere to naming traditions by giving classic "ethnic" names to their first born children, or may be particularly influenced by their own parents in selecting the names of their oldest kids. We test this hypothesis by testing for unique differences in name foreignness between first-born and other sons.

Table 2 estimates separate dummy variables for each step in the birth order separately for households with two, three, or four or more sons. For households with two foreign-born parents, each step along the birth order is associated with a 0.6-0.8 point decline in the F-index. However, for households with second- (and third-plus) generation parents, the linear birth order effect is driven by the difference between first-born sons and all higher-order brothers. First-born sons score up to 0.7 points higher on the F-index than their younger brothers, but we see no systematic differences between second and third or fourth sons.

The underlying coefficients, reported in Appendix Tables 2 and 3, suggest that controlling for having the same name as a parent does not change the main findings: that immigrants bestowed less foreign names on their US-born sons and daughters later in the birth order, consistent with a process of cultural assimilation. As expected, sons who shared their father's name scored higher on the F-index (4-9 points), presumably because the set of names used in the father's generation were, on average, more European in origin. Controls for name attributes – including name frequency and other name attributes (saint names, biblical names) – reduce what appears to be a spurious correlation between birth order and name foreignness for other household types.

The relationship between naming and birth order is robust to many alternative measures of name foreignness. Figure 5 (Panel B) reproduces our main results with the relative probability measure, which has a more natural interpretation than the F-index but is more sensitive to outliers. The effect of birth order on the relative probability measure of name foreignness monotonically declines with parental generation (first- to second- to third-generation immigrants), with no detectable effect for the children of third-plus generation parents or for children born abroad. For households with two foreign-born parents, the first son is given a name with a relative probability that is 0.33 points higher than the fourth son. The average gap in the relative probability between the children of immigrants and natives was 1.5 points in 1920 (compare mean values of 2.3 and 0.8, respectively). As before, this estimate implies that, as immigrants spent more time in the US (as proxied by birth order), they erased 25 percent of the "cultural" gap with natives (= 0.33/1.5).

Appendix Table 4 reports the effect of birth order on various measures of the F-index, focusing on households with a foreign-born head. Results are similar when adjusting the F-index

for phonetically-equivalent names using the NYSIIS algorithm, rather than using raw names (for example, treating Roberto and Robert as the same name); when fixing the F-Index in 1900, rather than assigning birth-cohort specific indices; when calculating the F-index using the 1920 (rather than the 1940) complete-count Census; or when using state-specific F-indices to allow for differential name trends by region.[12] Including southern residents in the analysis also leaves results unchanged. The coefficient of interest is 50 percent larger when using a country-specific F-index, which compares immigrants from a particular sending country to all other US residents. This pattern implies that immigrants were shifting away from names popular in their own culture towards names used both by natives and by other immigrant groups. Our main F-index compares the probability of a name being assigned to a foreign- or native-born individual, putting children of immigrants who were themselves born in the US in the 'native' category. Results are somewhat weaker when we instead assign second-generation immigrants to the 'foreign' category, as would be expected if immigrant parents are choosing names used by immigrant parents from other ethnic communities, rather than only those names used by native parents.

Finally, we explore variation in the speed of name-based assimilation among sub-categories of foreign-born parents. Name choice is an inexpensive means for immigrants to express their interest in integrating into US culture. Table 3 shows that the relationship between birth order and name foreignness was similar for literate and illiterate household heads and for immigrants who owned or rented their housing unit, suggesting that name-based assimilation was available to and used similarly by immigrants of all socioeconomic statuses.

---

[12] Fixing the F-index at a point in time ensures that the results are not driven by a mechanical relationship between birth order and trends in the F-index over time. Any differences in the F-index calculated in 1920 and 1940 would be due to mortality, name changes, or net-migration by birth cohort.

Appendix Table 5 subdivides households with foreign-born heads and/or spouses into categories based on age at arrival in the US, inter- vs. intra-marriage, and gender of foreign-born parent (mother vs. father). Name-based assimilation in immigrant households in which at least one parent arrived in the US as a child was not appreciable faster or slower than in households in which both parents arrived as adults. Households with immigrant parents from two different countries of origin shift name choices nearly twice as fast as immigrant parents in endogamous marriages, perhaps because they are less constrained by adherence to a single ethnic tradition. Among households with only one foreign-born parent, having a foreign-born father is associated with somewhat faster name-based assimilation than having only a foreign-born mother.

The speed of name-based assimilation differs by country of origin. Figure 6 shows that immigrants from most sending countries gave later-order sons less foreign names (with the exceptions of immigrants from Russia and Wales). Yet, the magnitude of this shift varies from 4 points on the F-index between first and fourth son for the Portuguese to less than 0.5 points for the Finns.

Immigrants from sending countries that were culturally distant from the US or that faced high levels of discrimination may have had the largest benefit from name-based assimilation, but they also may have experienced the highest costs of assimilation, in terms of foregoing aspects of their cultural identity. Table 4 proxies for cultural distance in two ways: first, by calculating an analogous F-index for surnames, positing that households with more foreign-sounding last names hail from more dissimilar cultures, and then by comparing immigrants who lived in enclaves to those who lived in more integrated areas.[13] In both cases, we then interact birth order with our

_____

[13] We note that, for some households, living in an enclave was a choice that may have signaled a strong interest in retaining their cultural identity. In this case, we would expect residents of

measure of cultural distance to examine whether immigrants that started out further from the native norm took more steps to assimilate or found themselves further behind. We find that households with more foreign last names and those that live in counties with more residents from their own home country shift toward native-sounding names more rapidly. In this sense, cultural assimilation generated convergence between immigrant families from different backgrounds.

V. Labor market consequences of having a foreign name: 1940 Census

The previous section documents that immigrant parents selected increasingly native-sounding names for their children as they spent time in the US, a form of cultural assimilation. We now turn to the labor market consequences of receiving a foreign-sounding name. We link men from their childhood household in 1920 to the 1940 Census and study whether brothers who received less foreign names attended more years of schooling, earned higher wages and were less likely to be unemployed, after controlling for age and place in the birth order. Differences in outcomes could be driven by preferential treatment by teachers or peers at school or by employers who use names as a signal of ethnic identity, or by the effect of one's name on self-identification with an ethnic community. Recent immigrants are more likely to give their children foreign names, but also may have fewer financial or cultural resources. To control for differences in access to resources in early childhood, even among siblings, our most restrictive specifications compare brothers who were born within one or two years of each other, or brothers who received equally foreign names at birth that diverged in foreignness over time due to differences in naming trends.

---

enclaves to be slower to adopt native-sounding names. Furthermore, residents of enclaves may have had less exposure to influences from the wider culture.

Our main regression equation estimates the relationship between an adult outcome *y* and the F-index value of a man's name at birth:

$$y_{ij} = \beta_1 F\ Index_{ij} + \beta_2 Age_{ij} + \beta_3 Birth\ Order_{ij} + \beta_4 X_{ij} + \alpha_j + \varepsilon_{ij} \tag{3}$$

We include household fixed effects in order to compare brothers given names with different attributes, and control for an individual's age and birth order rank. In some specifications, our vector of controls *X* also includes the F-index of an individual's name in adulthood, which is separately identified by changes in the popularity of names within different groups over time. Our sample includes sons of foreign-born parents in the 1920 Census that can be successfully matched to the 1940 Census. Our overall matched sample contains over one million observations and XX brother pairs.

Table 5 presents estimates of the relationship between name foreignness and our three main outcomes of interest: years of schooling, annual earnings, and unemployment. The first column uses variation across households in the linked sample, while the second column adds household fixed effects. We find that a wholesale shift from zero to one on the F-Index (from a completely foreign to a completely native name) is associated with nearly one less year of schooling, or two-thirds of a year between brothers; $936 lower annual earnings in 2010 dollars, or a $420 (2 percent) decline in earnings between brothers; and a nearly 3 point increase in the probability of unemployment. Smaller effects of name foreignness within households suggest that a portion of the population-wide estimate is driven by the correlation between names and family background. A more realistic 20 percent shift in the F-index (the typical gap between the children of immigrants and natives) would be associated with one less month of schooling, a 0.4 percent decline in annual earnings and a 0.6 point increase in the probability of unemployment.

The previous section shows that brothers born later in the birth order were given, on average, less foreign-sounding names. Yet, even after controlling for birth order, differential spacing between births could create an association between name foreignness and aspects of the home environment, including parental income and degree of cultural assimilation. In column 3, we restrict our analysis to brothers who were born within two years of each other and thus likely grew up in a very similar setting. The results are quantitatively similar in this subsample, although the coefficient on annual earnings loses significance.

The ethnic signal of names that parents select for their children at birth can be attenuated (or augmented) as the name becomes more/less popular among certain groups. For example, Eric and Kurt, two of the most foreign names in the data in 1940, are commonly given by native parents today. More relevant to an employer's perception of a worker's ethnic identity might be the Foreignness Index of his name at the time of labor market entry. Column 4 includes two F-indices on the right-hand side – one calculated at birth and the other at labor market entry. By controlling for the F-index at birth, we can identify the effect of name foreignness based on trends in name popularity over time, which are hard to predict and therefore likely exogenous to family background. As expected, we find that F-index at labor market entry is more quantitatively important than the F-index at birth, although the point estimate (for annual earnings) is not statistically significant, reflecting the high demands that this specification requires of the data (identification comes from households with two or more matched brothers whose names follow different trends over a twenty year period).

One mechanism through which brothers with less foreign names may experience better labor market outcomes is via better educational attainment. In Column 5, we test whether the association between name foreignness and schooling can fully account for declines in annual

earnings and employment. We find that the effect of name foreignness on unemployment is unaffected by this additional control but that the relationship between names and earnings is entirely explained by lower levels of schooling among brothers with more foreign names.

Appendix Table 6 considers a series of additional labor market outcomes, including the subcomponents of annual earnings (hourly wages, weeks worked during the year, and hours worked during the week) and various forms of employment. Consistent with the effect of name foreignness on unemployment at the time of the Census, brothers with more foreign names work less time during the year in both hours and weeks but, conditional on being employed, they do not receive a lower wage. The effect of name foreignness on employment is equal and opposite to the effect on unemployment, implying no effect of names on the probability of being out of the labor force in this prime-age sample (the omitted category). Yet, brothers with foreign names were no more likely to hold a public works job through the New Deal, despite facing higher unemployment, which could reflect some discrimination in access to public relief employment.

VI. Additional measures of cultural assimilation: Inter-marriage, citizenship and facility with English

Thus far, we have used the name choices that immigrants make for their children as a measure of cultural assimilation. To provide a fuller picture of the cultural assimilation process, this section considers a wider set of measures of cultural assimilation, including rates of inter-marriage, application for US citizenship and facility with English.

Inter-marriage has been used extensively in sociology as a marker of cultural assimilation (Gordon, 1964; Lieberson and Waters, 1988; see also Meng and Gregory, 2005 in economics). Rates of inter-marriage reflect direct preferences over who to marry as well as the degree of cultural segregation between groups. That is, immigrants may be more likely to marry each other

simply because they interact more regularly in segregated neighborhoods or schools. Inter-marriage is a relatively stringent measure of cultural assimilation because it requires not only that immigrants want to integrate into their new society, but also that natives (or members of other immigrant groups) are willing to interact with them (Kalmijn, 1998).

We construct the proportion of endogamous marriages for first- and second-generation immigrants by country of origin in 1930, excluding immigrants who were likely married abroad (based on reported immigration year and age at first marriage). In particular, we calculated the share of married immigrants whose spouse was either a first- or second-generation immigrant from the same country of origin; here, we report patterns from the 1930 5 percent IPUMS sample.[14]

Figure 7 presents endogamy rates for immigrants from 16 sending countries. Panel A reports means from the raw data and Panel B controls for the size of each immigrant group and the sex ratio within the group by state (a proxy for the marriage marriage). These controls are designed to address mechanical differences in the likelihood of meeting a potential spouse from one's own group. First-generation immigrants exhibited a strong tendency toward endogamy, which weakened by the second generation. The mean probability of endogamous marriage falls substantially from 61 percent for the first generation to 32 percent for the second generation (endogamy rates were slightly higher for women; see Appendix Figure 1). The in-group marriage of immigrants who arrived as children (which we term the 1.5 generation) fits between these two

---

[14] We focus on 1930 because it was the first Census to add the "age at first marriage" variable. Previous work on inter-marriage in this period (e.g., Pagnini and Morgan, 1990; Angrist, 2002) analyze earlier Censuses and so they cannot separate marriages that occurred in the US from those that occurred prior to migration. If the endogamy rate for the first generation is biased upward by marriages that occurred abroad, the rate of marital assimilation by the second generation will also be overstated. Including marriages that occurred abroad will differentially bias the endogamy rate for country-of-origin groups that tended to migrate in families, rather than as single individuals.

values. We also find sizeable variation in the endogamy rate across countries of origin. For example, 89 percent of first-generation immigrants from Italy were married to another first- or second-generation Italian, compared with only 28 percent of first generation immigrants from Scotland. The Finns, Portuguese and Russians also exhibit a strong tendency toward endogamy, while the French and the English are unlikely to marry fellow countrymen. There is considerable persistence in the endogamy rate across generations; the within-country correlation between the first and second-generation endogamy rates is 0.90.

Figure 8 consider two additional aspects of cultural assimilation: the decision to apply for US citizenship (Panel A) and facility with English (Panel B). Each panel plots the average likelihood of the given activity for first-generation immigrants by gender and country of origin. Over 75 percent of immigrants from most sending countries received citizenship or had started the application process by 1930 (at which point, the average immigrant had been in the US for 24.5 years). As with the rates of inter-marriage, applications for US citizenship were lower among the Portuguese, Italians, Finns, three groups with high return migration rates, and to some extent Russians (Gould, 1980). Across all the sending countries, women exhibited lower rate of citizenship application than men. This pattern is consistent with higher female endogamy, because one route to citizenship is via marriage to a US citizen. Immigrants from most sending countries reported near-universal facility with English, although the ability to speak English was somewhat lower for the Portuguese, Italians, and Finns, and for women relative to men. Note that the historical Census measure simply indicates the ability to speak *some* English; unlike today, this question did not ask about language spoken at home or distinguish between levels of English-speaking ability.

Table 6 presents the pairwise correlation between the various measures of cultural assimilation at the country-of-origin and individual level. Endogamy and the average foreignness of names given to sons born in the US are coded to be negatively correlated with cultural assimilation, while applications for US citizenship and the ability to speak English are positively correlated with cultural assimilation. These measures are all related with each other in the expected way (e.g., immigrants in an endogamous marriage are less likely to be able to speak English or to have applied for US citizenship, and select more foreign names for their children). The correlations are all highly significant at the individual level, providing strong validation for using name choice as a measure of cultural assimilation.

VII. Conclusion

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122, no. 3 (2014).

Alba, Richard D. and Reid M. Golden. "Patterns of Ethnic Marriage in the United States." *Social Forces* 65, no. 1 (1986).

Alesina, Alberto and Paola Giuliano. 2011. "Preferences for Redistribution." In *Handbook of Social Economics*, edited by Alberto Bisin and Jess Benhabib, p. 93-131. Amsterdam: Elsevier.

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

Arai, Mahmood, and Peter Skogman Thoursie. "Renouncing Personal Names: An Empirical Examination of Surname Change and Earnings." *Journal of Labor Economics* 27, no. 1 (2009): 127-147.

Bertrand, Marianne, and Sendhil Mullainathan. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94, no. 4 (2004): 991-1013.

Biavaschi, Costanza, Corrado Giulietti, and Zahra Siddique. "The Economic Payoff of Name Americanization." IZA Discussion Paper No. 7725, 2013.

Blau, Francine D., Lawrence M. Kahn, Albert Yung-Hsu Liu, and Kerry L. Papps. 2013. "The Transmission of Women's Fertility, Human Capital, and Work Orientation Across Immigrant Generations." *Journal of Population Economics* 26 (2): 405-35.

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

Carneiro, Pedro, Sokbae Lee and Hugo Reis. 2015. "Please Call me John: Name Choice and the Assimilation of Immigrants in the United States, 1900-30." Institute for Fiscal Studies Working Paper.

Chiswick, Barry R. "The Effect of Americanization on the Earnings of Foreign-born Men." *Journal of Political Economy* (1978): 897-921.

Citrin, Jack, Donald P. Green, Christopher Muste, and Cara Wong. "Public Opinion Toward Immigration Reform: The Role of Economic Motivations." *Journal of Politics* 59, no. 03 (1997): 858-881.

Fernández, Raquel, and Alessandra Fogli. 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American Economic Journal: Macroeconomics* 1(1): 146-77.

Fetzer, Joel S. Public *Attitudes toward Immigation in the United States, France, and Germany*. New York: Cambridge University Press, 2000.

Foley, Mark C., and Timothy W. Guinnane. "Did Irish Marriage Patterns Survive the Emigrant Voyage? Irish-American Nuptiality, 1880-1920." *Irish Economic and Social History* 26 (1999): 15-34.

Fouka, Vasiliki. 2015. "Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I." Manuscript.

Figlio, David N. "Names, Expectations and the Black-white Test Score Gap." National Bureau of Economic Research Working Paper No. w11195, 2005.

Fryer, Roland G., and Steven Levitt. "The Causes and Consequences of Distinctively Black Names." *Quarterly Journal of Economics* 119, no. 3 (2004): 767-805.

Gerhards, Jürgen, and Silke Hans. "From Hasan to Herbert: Name-Giving Patterns of Immigrant Parents between Acculturation and Ethnic Maintenance." *American Journal of Sociology* 114, no. 4 (2009): 1102-1128.

Goldin, Claudia. "The Political Economy of Immigration Restriction in the United States, 1890 to 1921." In *The Regulated Economy: A Historical Approach to Political Economy*, pp. 223-258. University of Chicago Press, 1994.

Goldstein, Joshua R. and Guy Stecklov. 2015. "From Patrick to John F.: Ethnic Names and Occupational Success in the Last Era of Mass Migration." *American Sociological Review*.

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

Gould, John D. 1980. "European Inter-Continental Emigration. The Road Home: Return Migration from the USA." *Journal of European Economic History* 9 (1): 41-112.

Guinnane, Timothy W., Carolyn M. Moehling, and Cormac Ó Gráda. "The Fertility of the Irish in the United States in 1910." *Explorations in Economic History* 43, no. 3 (2006): 465-485.

Hacker, J. David. "Child Naming, Religion, and the Decline of Marital Fertility in Nineteenth-century America." *History of the Family* 4, no. 3 (1999): 339-365.

Hainmueller, Jens, and Michael J. Hiscox. "Attitudes toward Highly Skilled and Low-skilled Immigration: Evidence from a Survey Experiment." *American Political Science Review* 104, no. 01 (2010): 61-84.

Hatton, Timothy J. "The Immigrant Assimilation Puzzle in Late Nineteenth-Century America." *Journal of Economic History* 57, no. 01 (1997): 34-62.

Hatton, Timothy J. and Jeffrey G. Williamson. *The Age of Mass Migration: Causes and Economic Impact*. New York: Oxford University Press, 1998.

Higham, John. *Strangers in the Land: Patterns of American Nativism, 1860-1925*. Rutgers University Press, 1955.

Jones, Maldwyn Allen. *American Immigration: Second Edition*. Chicago, University of Chicago Press, 1992.

Kalmijn, Matthijs. "Intermarriage and Homogamy: Causes, Patterns, Trends." *Annual Review of Sociology* (1998): 395-421.

King, Desmond. *Making Americans: Immigration, Race, and the Origins of the Diverse Democracy*. Harvard University Press, 2009.

Lieberson, Stanley, and Mary C. Waters. *From Many Strands: Ethnic and Racial Groups in Contemporary America*. Russell Sage Foundation, 1988.

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

Lieberson, Stanley, and Kelly S. Mikelson. "Distinctive African American Names: An Experimental, Historical, and Linguistic Analysis of Innovation." *American Sociological Review* (1995): 928-946.

Lubotsky, Darren. "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings." *Journal of Political Economy* 115, no. 5 (2007): 820-867.

Luttmer, Erzo F. P., and Monica Singhal. 2011. "Culture, Context, and the Taste for Redistribution." *American Economic Journal: Economic Policy*, 3 (1): 157-79.

Meng, Xin, and Robert G. Gregory. "Intermarriage and the Economic Assimilation of Immigrants." *Journal of Labor Economics* 23, no. 1 (2005): 135-174.

Minns, Chris. "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, no. 4 (2000): 326-350.

Moehling, Carolyn, and Anne Morrison Piehl. "Immigration, Crime, and Incarceration in Early Twentieth-Century America." *Demography* 46, no. 4 (2009): 739-763.

Moehling, Carolyn M., and Anne Morrison Piehl. "Immigrant Assimilation into US Prisons, 1900–1930." *Journal of Population Economics* 27, no. 1 (2014): 173-200.

Moser, Petra. 2012. "Taste-Based Discrimination Evidence from a Shift in Ethnic Preferences after WWI." *Explorations in Economic History* 49 (2): 167-88.

Pagnini, Deanna L., and S. Philip Morgan. "Intermarriage and Social Distance among US Immigrants at the Turn of the Century." *American Journal of Sociology* (1990): 405-432.

Rossi, Alice S. "Naming Children in Middle-Class Families." *American Sociological Review* (1965): 499-513.

Rubinstein, Yona, and Dror Brenner. "Pride and Prejudice: Using Ethnic-Sounding Names and Inter-Ethnic Marriages to Identify Labour Market Discrimination." *Review of Economic Studies* 81, no. 1 (2014): 389-425.

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

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

Sue, Christina A., and Edward E. Telles. "Assimilation and Gender in Naming." *American Journal of Sociology* 112, no. 5 (2007): 1383-1415.

Watkins, Susan Cotts, ed. *After Ellis Island: Newcomers and Natives in the 1910 Census*. New York: Russell Sage Foundation, 1994.

Watkins, Susan Cotts, and Andrew S. London. "Personal Names and Cultural Change: A Study of the Naming Patterns of Italians and Jews in the United States in 1910." *Social Science History* (1994): 169-209.

Zelinsky, Wilbur. "Cultural Variation in Personal Name Patterns in the Eastern United States." *Annals of the Association of American Geographers* 60, no. 4 (1970): 743-769.

Figure 1: Average name foreignness, Children of native- and foreign-born fathers, Birth cohorts of 1850-1920
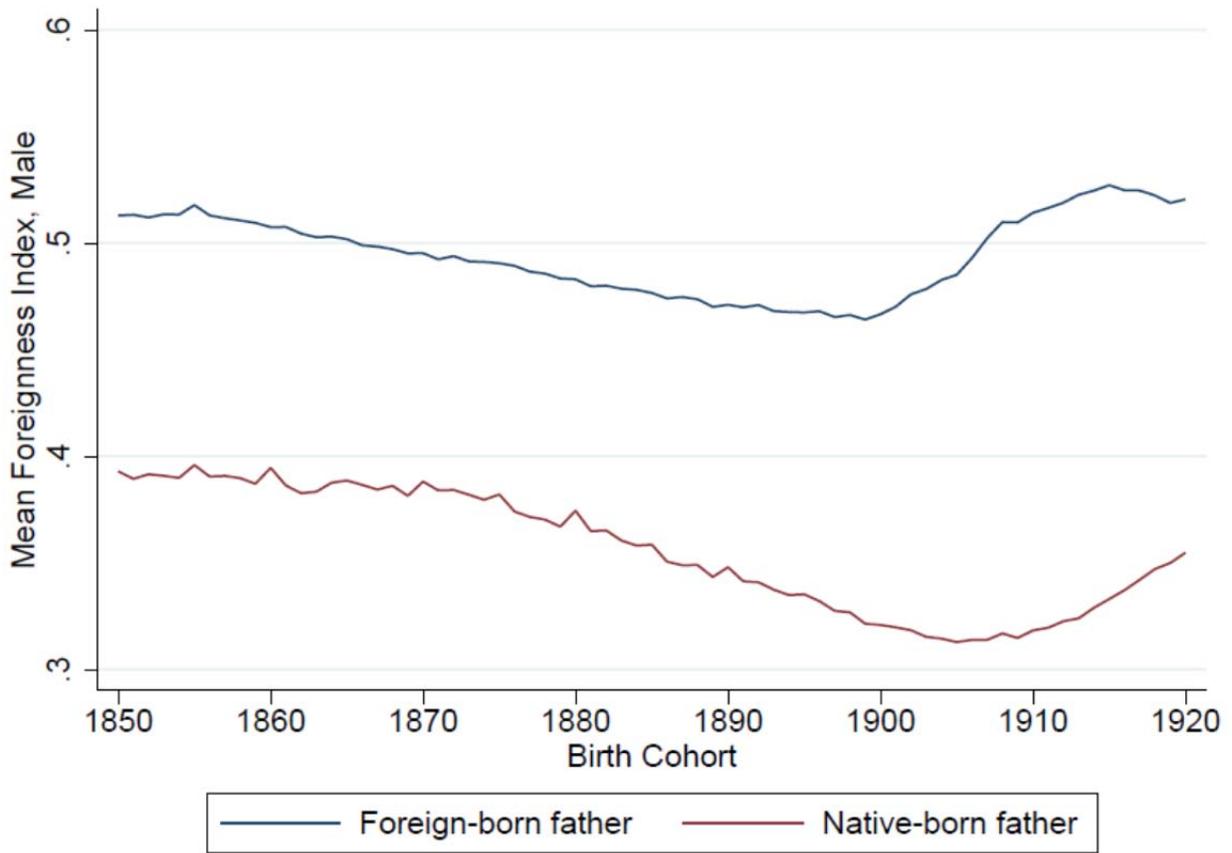
Notes: Sample is restricted to men born in the US. Father's nativity is determined using the father's place of birth variable. The Foreignness Index is calculated for each name and birth cohort using the complete-count (100 percent) sample of the 1920 census. The F-Index value for men in birth cohort *t* is based on men born in *t-1* through *t-20*.

**Figure 2: Kernel density estimates of name Foreignness Index in 1920,
Children of native-born or foreign-born in the US more/less than 10 years**
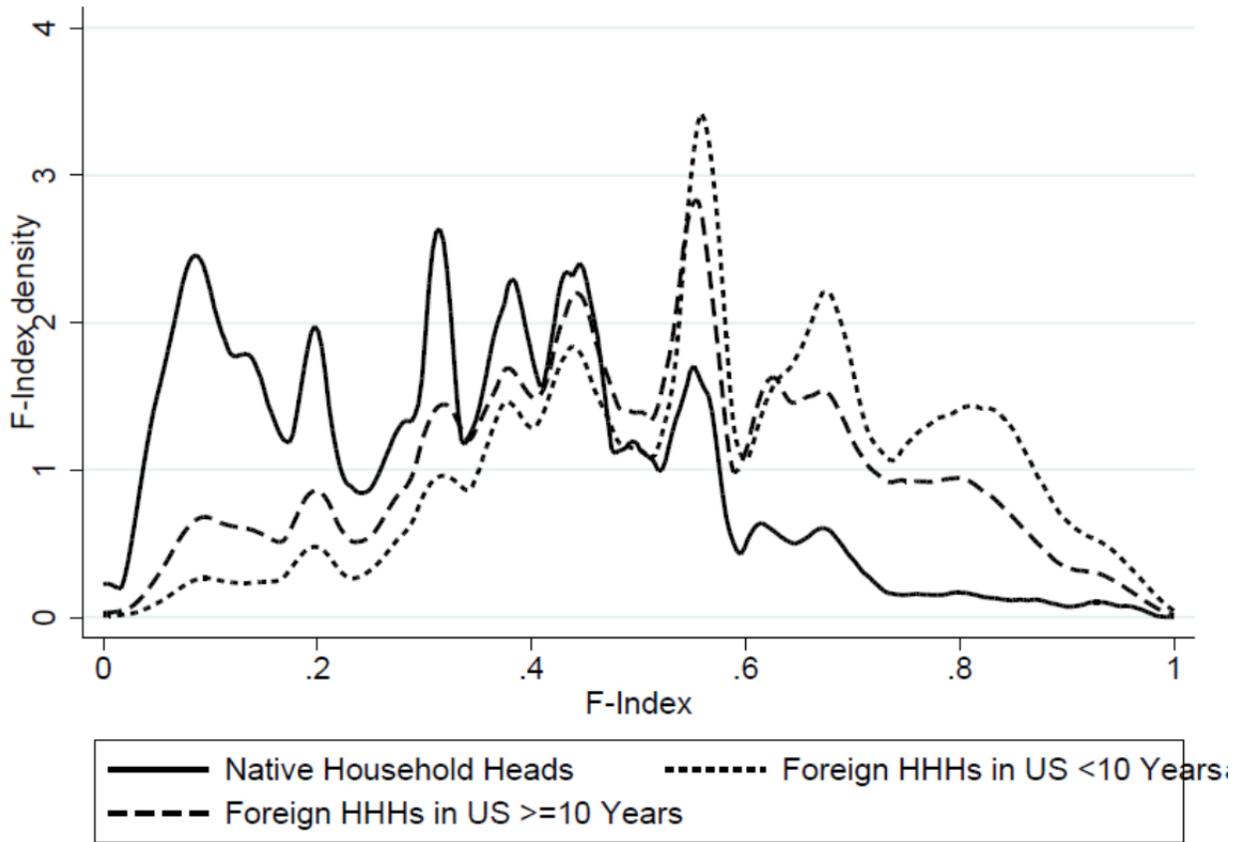
Notes:

Figure 3: 1920 census manuscript of the XX family

Figure 4: Immigrants gave sons and daughters less foreign names after spending more time in US, 1920 (Dependent variable = F-index *x 100*)
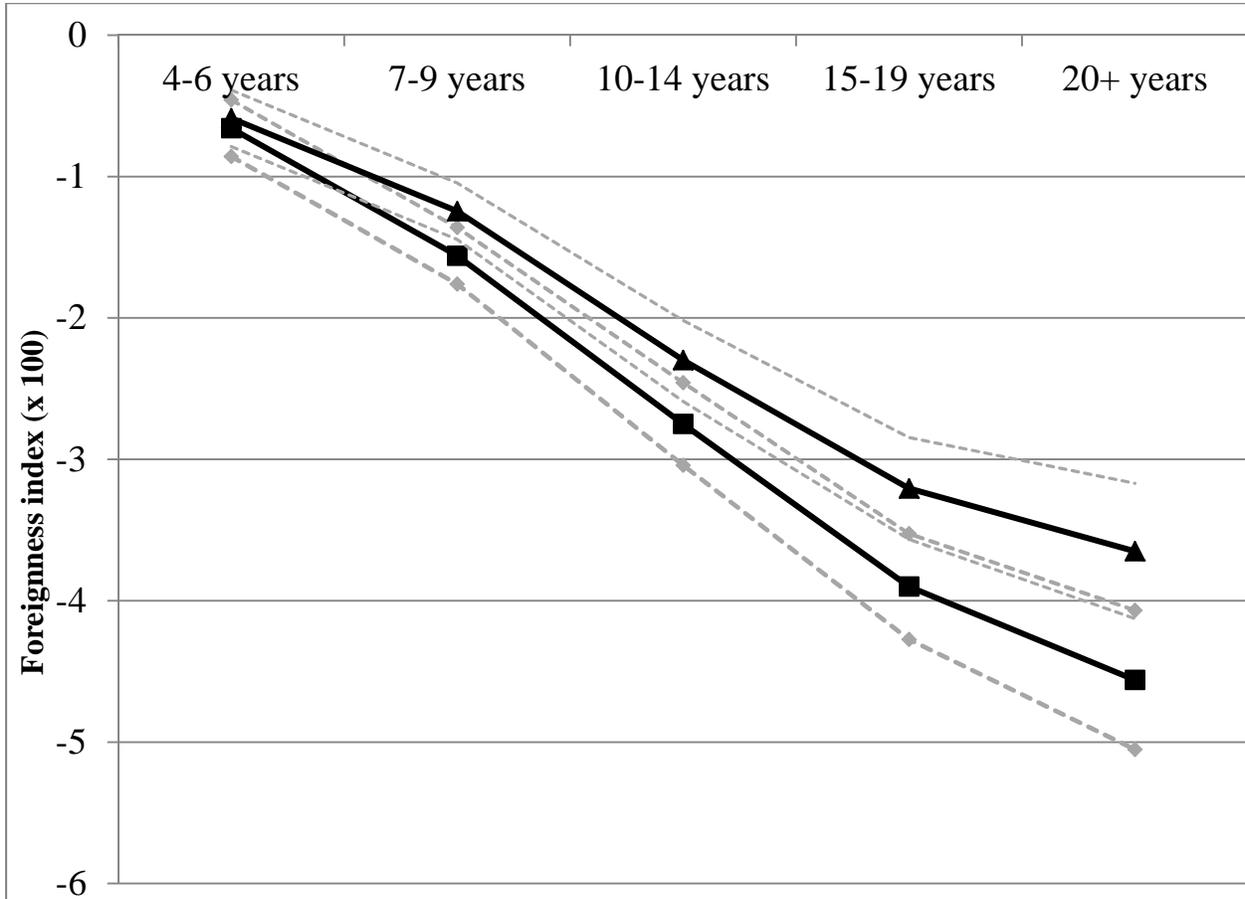
Notes: Squares = sons and triangles = daughters. Reported coefficients from estimates of Equation 1, a regression of F-index on indicators for years household head spent in the US by time of child's birth. Regressions also include household fixed effects and controls for child's birth year and rank in birth order. Data from 1920 complete-count Census. Sample includes non-black children aged 0-18 living with their parents in a non-southern state. Households must have foreign-born head and spouse (mother) must be less than 43 years old ($N = 2{,}230{,}017$).

Figure 5: Immigrants gave less foreign names to children later in birth order

Panel A: Foreign-Born Index
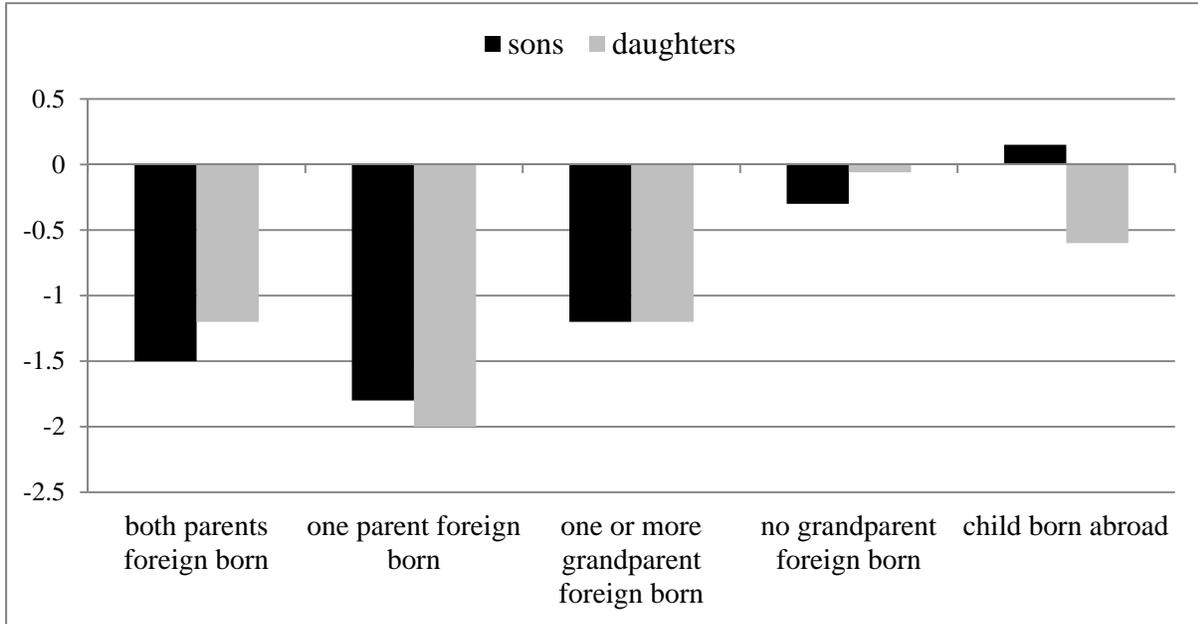
Panel B: Relative Probability Index
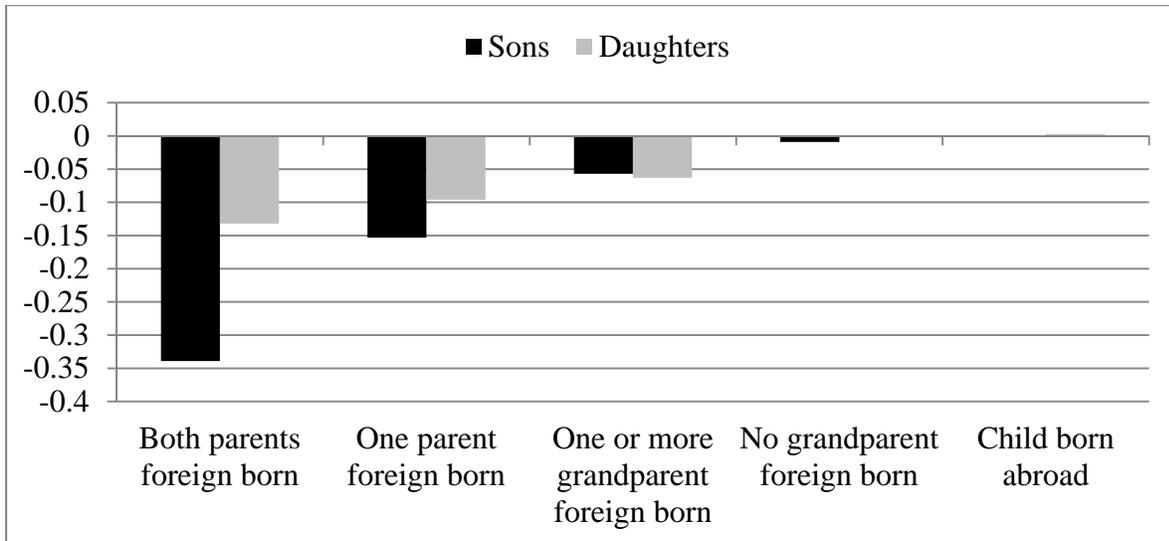
Notes: Data from the complete-count 1920 census. Bars report the implied effect on F-index x 100 (panel A) or relative probability (Panel B) of moving from first son to fourth son in birth order. Underlying coefficients for Equation 2, controlling for household fixed effects, child's birth year, whether child has same name as father, and other attributes of name (frequency, whether saint or biblical name), reported in Appendix Tables 2 and 3. Sample includes non-black children aged 0-18 living with their parents in a non-southern state. Spouse of household head (mother) must be less than 43 years old.

Figure 6: Effect of birth order on foreignness of name, by sending country
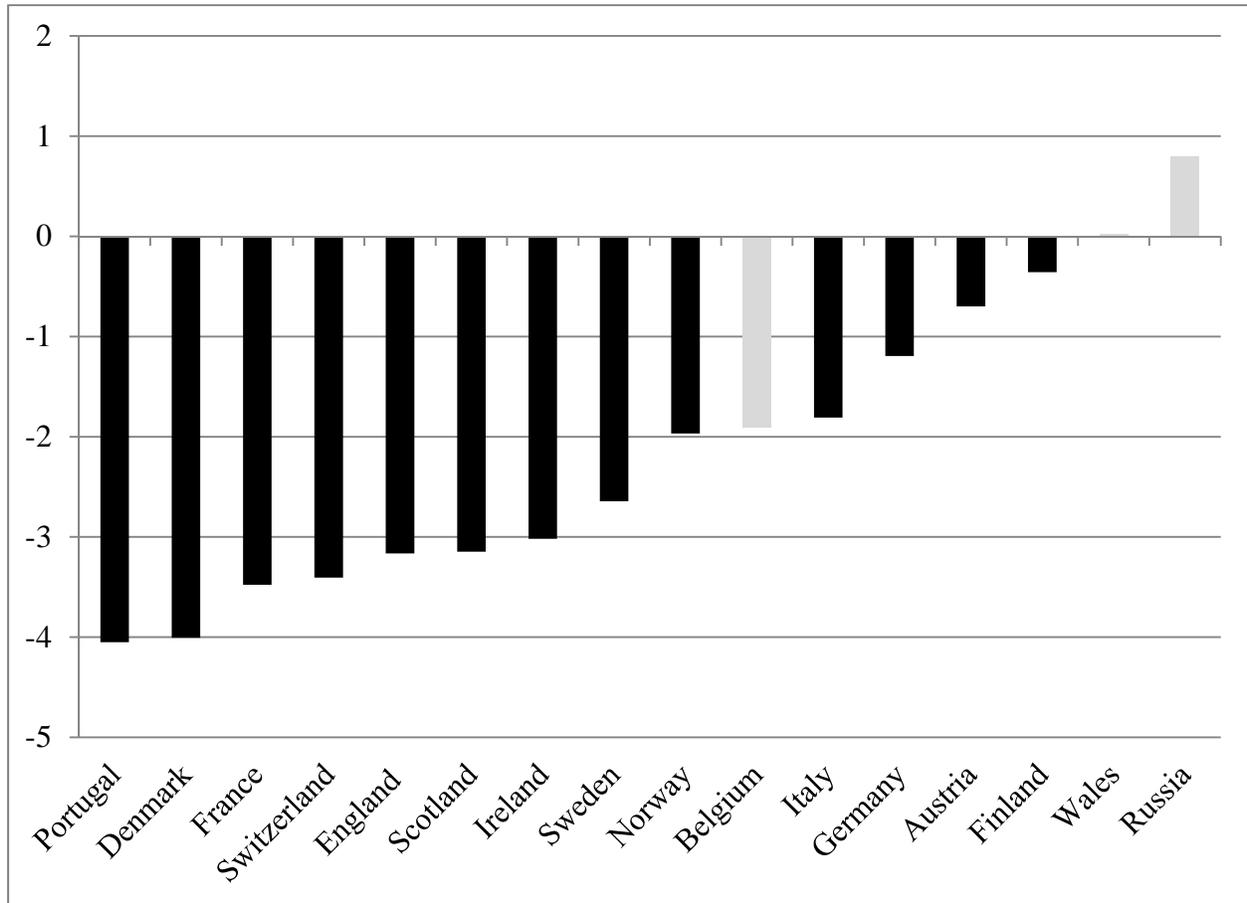
Notes: Reported coefficients from estimates of Equation 2, a regression of name foreignness (F-index x 100) on linear birth order rank, controlling for household fixed effects and child's birth year. The bars report the implied effect on the F-Index of moving from the first son to the fourth son in the birth order; dark bars are statistically significant at the 5 percent level. Data is from the complete-count 1920 census. Sample includes non-black children aged 0-18 living with their parents in a non-southern state. Spouse of household head (mother) must be less than 43 years old.

Figure 7: Share of first and second generation immigrant men in endogamous marriage, by country of origin, 1930
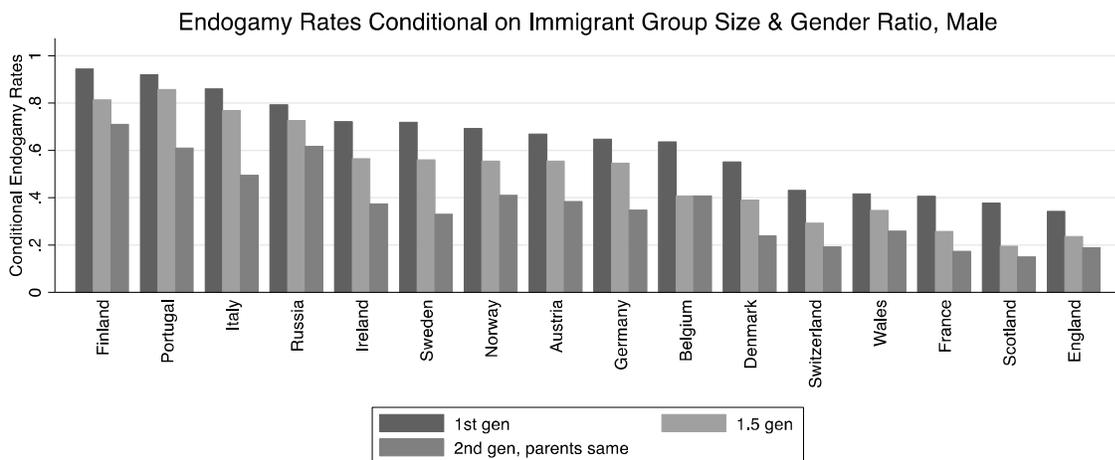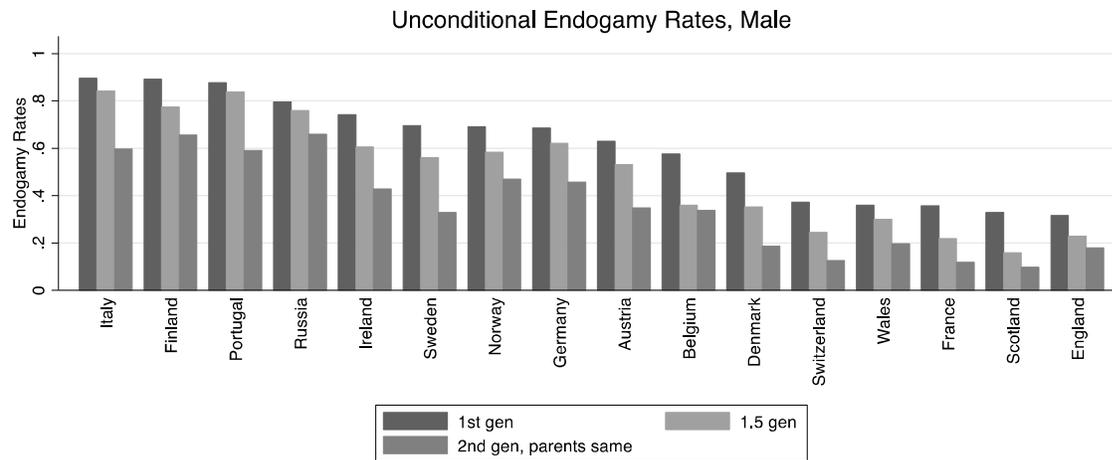
Note: Figure based on men in IPUMS 5% sample of 1930 census who are currently married and whose age at first marriage occurred after arrival in the US. Men whose spouse (or spouse's parents) were born in the same country of origin as he (or his parents) are considered to be in an 'endogamous' marriage. Panel B reports country of origin fixed effects from a regression whose dependent variable is a dummy equal to one for endogamous marriage with controls for the group size and gender ratio of the corresponding immigrant group at the state level. Immigrant group size is defined as the number of immigrants (first or second generation) with a particular ancestry, relative to the total population. Gender ratio is defined as the ratio between total number of male to female immigrants (first or second generation) with a particular ancestry.

Figure 8: Share of immigrants who engaged in other forms of cultural assimilation by country of origin, 1930

A. Applied for US citizenship

Likelihood of Applying to Citizenship Conditional on Immigrant Group Size & Gender Ratio, 1st Gen
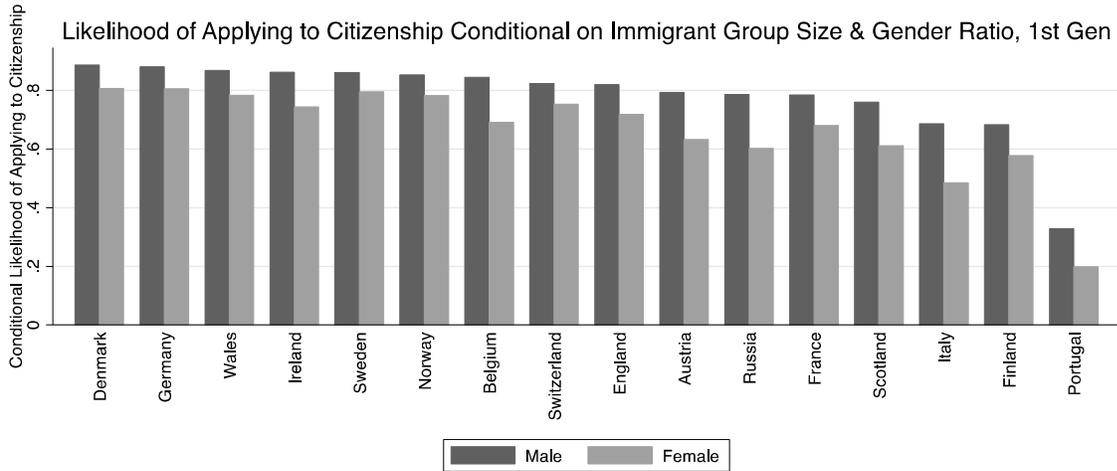
B. Reports ability to speak English

Likelihood of Speaking English Conditional on Immigrant Group Size & Gender Ratio, 1st Gen
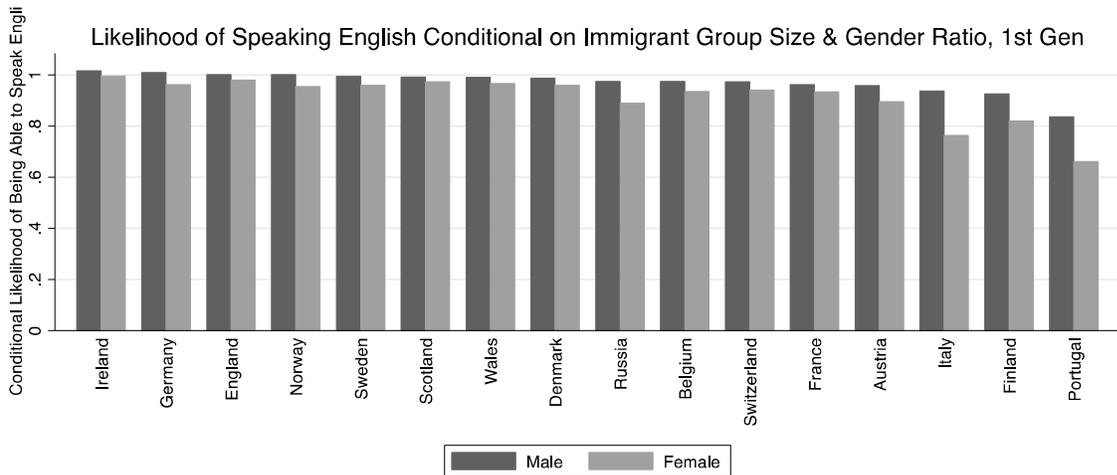
Note: Figure is based on IPUMS 5% sample of 1930 census. The sample is restricted to individuals who were born outside of the US in one of the listed countries. Panel A reports country of origin fixed effects from a regression whose dependent variable is a dummy equal to one for immigrants who report being naturalized US citizens or having applied for first papers. The regression also controls for the group size and gender ratio of the corresponding immigrant group at the state level. Panel B follows the same format for a regression whose dependent variable is a dummy equal to one for immigrants reporting the ability to speak English. See the notes to Figure 7 for definitions of the other controls.

Table 1: Examples of foreign, neutral, and native names (1900-1920 birth cohorts)

Most foreign (F-Index >0.95)	Most neutral (F-Index = 0.5)	Most native (F-Index <0.05)
Pedro	David	Troy
Francisco	George	Nolan
Jesus	Jeffrey	Denver
Eric	Daniel	Shelby
Juan	Albert	Herschel
Jose	Jack	Galen
Kurt	Bennie	

Notes: Names selected for having high/lowest/most neutral F-index values in 1940 complete-count Census. Currently, we report such names for all birth cohorts in the data (1870-1940). Need to update to only 1900-20 cohorts.

Table 2: Name foreignness and rank in birth order, Sons of foreign-born and native-born parents

	2 sons	3 sons	4 or more sons
A. Both parents foreign born			
=1 if second born	-0.662***	-0.659***	-0.366***
	(0.087)	(0.093)	(0.133)
=1 if third born		-1.220***	-1.222***
		(0.151)	(0.185)
=1 if fourth or more			-1.860***
			(0.264)
N	516,417	375,690	209,224
B. Parents born in US; 1+ grandparent born abroad			
=1 if second born	-0.692***	-0.756***	-0.550***
	(0.092)	(0.113)	(0.172)
=1 if third born		-0.826***	-0.861***
		(0.183)	(0.243)
=1 if fourth or more			-0.628*
			(0.347)
N	413,497	238,171	117,937
C. Parents and grandparents born in US			
=1 if second born	-0.372***	-0.181***	-0.170
	(0.059)	(0.070)	(0.105)
=1 if third born		-0.147	-0.086
		(0.113)	(0.149)
=1 if fourth or more			0.286
			(0.212)
N	1,080,826	648,824	75,150

Notes: All specifications contain controls for child's age, a dummy for same name as father, name frequency and household FE.

Table 3: Similar birth order effect for literate/illiterate household heads and for homeowners/renters

Dependent variable = F-index *x* 100; Sample = Both parents are foreign born

| | Household head is… | | | |
	Literate	Not literate	Homeowner	Renter
Birth order	-0.497***	-0.431***	-0.490***	-0.509***
	(0.042)	(0.085)	(0.059)	(0.049)
N	1,378,050	269,729	604,170	1,042,427

Notes: Data from complete-count 1920 census. Sample includes non-black children aged 0-18 living with their parents in a non-southern state. To observe complete birth order, sample restricted to households in which mother is less than 43 years old. Linear birth order variable defined over sons within a household. All regressions control for child's birth year, a dummy for same name as father, frequency of name, and household FE.

Table 4: Name-based assimilation by proxies of cultural distance

A. F-index of last name	
Birth order	0.189***
	(0.106)
Birth order *x* F-index of last name	-1.415***
	(0.112)
N	1,149,188
B. % county from home country	
Birth order	-0.536***
	(0.055)
Birth order x % from home country	-1.207***
	(0.372)
N	1,745,338

Notes: Data from the complete-count 1920 census. Sample includes non-black children with a foreign-born head of household who were between the ages of zero and 18 and lived with their parents in a non-southern state. To observe complete birth order, sample restricted to households in which mother is less than 43 years old. All regressions control for child's age, a dummy equal to one if child has same name as father, and name frequency.

Table 5: The effect of name foreignness on education, earnings and unemployment

Dep. variable	No HH FE	With household FE			
		Full sample	Brothers, 1 year apart	F-index at birth/age 20	Control for education
		PANEL A **Dep. Variable: Highest grade (Mean=10.07)**			
Highest grade	-0.915***	-0.677***	-0.799***		---
	(0.016)	(0.039)	(0.103)		
N	747,533	747,533	208,296		
		PANEL B **Dep. Variable: Annual earnings (Mean=$19,383)**			
Annual earnings	-936.16***	-449.63**	-592.51	3.311	611.75***
	(139.3)	(223.92)	(854.85)	(672.84)	(218.04)
				-404.87 (688.86)	
N	681,309	681,309	191,939	681,309	672,068
		PANEL C **Dep. Variable: Unemployed (Mean=0.100)**			
Unemployed	0.027***	0.029***	0.024**		0.024***
	(0.001)	(0.003)	(0.01)		(0.003)
N	1,039,035	1,039,035	286,700		1,021,886

Note: Sample includes men matched between 1920 and 1940 complete-count Censuses. Men must be 0-18 in 1920, living at home with parents and living with a mother who is 43 years old or younger. Sample restricted to men whose fathers were foreign-born (second-generation immigrants) who were not self-employed in 1940. All regressions control for child's age. Columns 2-5 add household fixed effects. Columns 2, 4 and 5 also control for a child's rank in the birth order.

Table 6: Pairwise correlations of various measures of cultural assimilation

	Variables	1	2	3
A.	**16 sending countries**			
1	Endogamy rate (-)			
2	Able to speak English (+)	-0.642***		
3	Applied for Citizenship (+)	-0.476*	0.944***	
4	Average F-index of sons (-)	0.489*	-0.767***	-0.666***
B.	**Individual (*N* = 23,043)**			
1	Endogamy rate (-)			
2	Able to speak English (+)	-0.065***		
3	Applied for Citizenship (+)	-0.075***	0.258***	
4	Average F-index of sons (-)	0.247***	-0.106***	-0.156***

Note: IPUMS 5% sample of 1930 census. (+) and (-) indicate positive and negative indicators of cultural assimilation. All samples restricted to white men who were born abroad and were 10 years or older. For *endogamy rate*, sample is further restricted to men who were currently married, and whose marriage took place after arrival in the US. For *average F-index of sons*, sample is restricted to men whose spouse is younger than 43 years old and whose oldest child is below 18 years old or less. The list of 16 sending countries underlying the correlations in Panel A can be found in Figure 2. The sample underlying the individual correlations in Panel B imposes all of the above restrictions (age, marital status, spouse's age and so on).

Appendix Figure 1

Share of first and second generation immigrant women in endogamous marriage, by country of origin, 1930

Endogamy Rates Conditional on Immigrant Group Size & Gender Ratio, Female
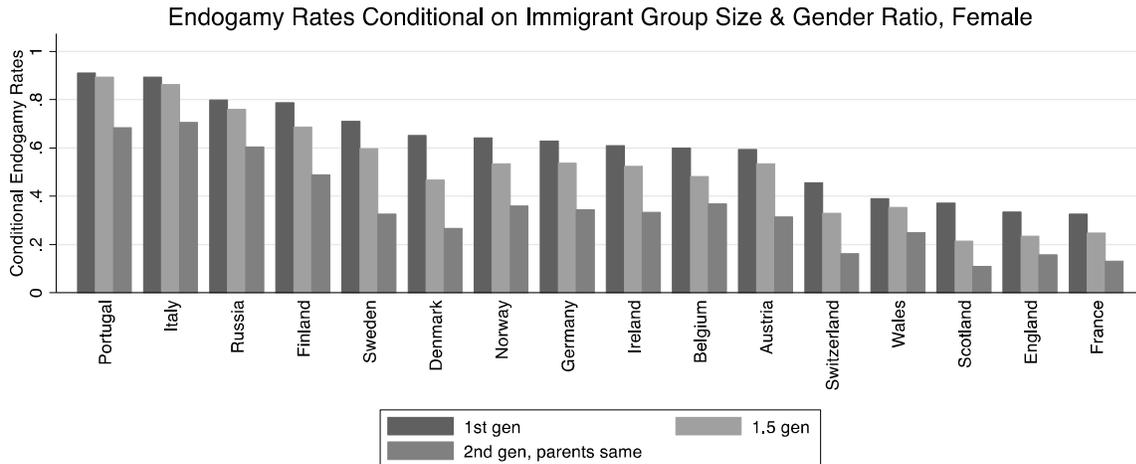
Note: Figure based on women in IPUMS 5% sample of 1930 census who are currently married and whose age at first marriage occurred after arrival in the US. Women whose spouse (or spouse's parents) were born in the same country of origin as she (or her parents) are considered to be in an 'endogamous' marriage. Panel B reports country of origin fixed effects from a regression whose dependent variable is a dummy equal to one for endogamous marriage with controls for the group size and gender ratio of the corresponding immigrant group at the state level. See notes to Figure 7 for definitions of control variables.

Appendix Table 1: No trend breaks in naming practices for German families during World War I

Trend Break Year	Coefficient	Std. Error	F-statistics
1910	0.006	[0.0083]	0.572
1911	0.003	[0.0081]	0.137
1912	0.002	[0.0079]	0.057
1913	0.006	[0.0077]	0.612
1914	0.004	[0.0077]	0.229
1915	0.004	[0.0076]	0.276
1916	0.003	[0.0076]	0.163
1917	0.004	[0.0077]	0.217
1918	0.000	[0.0077]	0.003
1919	0.002	[0.0078]	0.075
1920	0.006	[0.0079]	0.617

Notes:

Appendix Table 2: Did immigrants give less foreign name to sons later in the birth order?
Dependent variable = F-index *x* 100

	(1)	(2)	(3)	(4)
A. Both parents foreign born: Mean DV = 56.1				
Birth order	-0.464***	-0.339***	-0.495***	
	(0.038)	(0.039)	(0.038)	
=1 if dad's name		4.109***	4.887***	
		(0.067)	(0.065)	
Name freq., share			-243.8***	
			(1.159)	
Saint/biblical	N	N	N	Y
N	1,808,882	1,720,297	1,644,617	
B. One parent foreign born: Mean DV = 40.4				
Birth order	-1.047***	-0.663***	-0.574***	
	(0.059)	(0.060)	(0.061)	
=1 if dad's name		7.745***	6.950***	
		(0.094)	(0.094)	
Name freq., share			118.2***	
			(2.120)	
Saint/biblical	N	N	N	Y
N	793,644	759,569	732,862	
C. Child born abroad (*N* = 174,872) Mean DV = 61.1				
Birth order	-0.091	0.053	0.048	
	(0.277)	(0.353)	(0.207)	
=1 if dad's name		1.807***	3.055***	
		(0.460)	(0.306)	
Name freq., share			-333.2***	
			(6.014)	
Saint/biblical	N	N	N	Y
N	187,712	179,653	174,872	

Continuation of Appendix Table 2

	(1)	(2)	(3)	(4)
D. Parents born in US; 1+ grandparents born abroad: Mean DV = 36.7				
Birth order	-1.079***	-0.566***	-0.415***	
	(0.045)	(0.046)	(0.061)	
=1 if dad's name		8.541***	6.791***	
		(0.070)	(0.069)	
Name freq., share			235.5***	
			(1.627)	
Saint/biblical	N	N	N	Y
N	1,415,121	1,364,514	1,307,359	
E. Parents born in US; No grandparents born abroad: Mean DV = 33.5				
Birth order	-0.640***	-0.385***	-0.140***	
	(0.028)	(0.029)	(0.029)	
=1 if dad's name		6.039***	3.620***	
		(0.051)	(0.050)	
Name freq., share			402.1***	
			(1.063)	
Saint/biblical	N	N	N	Y
N	3,898,104	3,704,419	3,444,189	

Notes: Data is from the complete-count 1920 census. Sample includes non-black children aged 0-18 living with their parents in a non-southern state. To observe complete birth order, sample restricted to households in which mother is less than 43 years old. Linear birth order variable defined over sons within a household. All regressions control for child's age and household FE.

Appendix Table 3: Did immigrants give less foreign name to daughters later in the birth order? Dependent variable = F-index *x* 100

	Both parents FB	One parent FB	1+ grandpar FB	No grandpar FB	Foreign born child
Birth order	-0.372***	-0.525***	-0.440***	-0.022	-0.223
	(0.048)	(0.080)	(0.064)	(0.041)	(0.358)
N	1,641,950	706,847	1,245,108	3,372,954	169,798

Notes: Data is from the complete-count 1920 census. Sample includes non-black children aged 0-18 living with their parents in a non-southern state. To observe complete birth order, sample restricted to households in which mother is less than 43 years old. Linear birth order variable defined over sons daughters within a household. All regressions control for child's age, dummy for same name as mother, frequency of name, and household FE

Appendix Table 4: Robustness to measurement of F-Index
Dependent variable = F-index x 100; coefficient on linear birth order

	(1)
Baseline	-0.606***
N = 2,151,775	(0.044)
Adjust names with NYSIIS	-0.588***
	(0.041)
Fix F-index in 1900 [+]	-0.440***
	(0.058)
F-index calculated in 1920 [+]	-0.511***
	(0.052)
F-index, separate by country	-0.976***
N = 1,586,246	(0.075)
F-index, by state of birth	-0.640***
N = 2,055,583	(0.045)
F-index with 2nd generation as foreign [+]	-0.358***
	(0.047)
Birth order = All kids [+]	-0.274***
	(0.035)
Include South [+]	-0.483***
N = 1,709,630	(0.048)

Notes: Data from the complete-count 1920 census. Sample includes non-black children with a foreign-born head of household who were between the ages of zero and 18 and lived with their parents in a non-southern state. To observe complete birth order, sample restricted to households in which mother is less than 43 years old. All regressions control for child's age, a dummy equal to one if child has same name as father, and name frequency. Rows marked with [+] only include children with two foreign-born parents and need to be updated.

Appendix Table 5: Results for sub-groups of immigrant households
Dependent variable = F-index x 100; coefficient on linear birth order

	(1)
Baseline – 2 parents foreign	*-0.495****
N = 1,664,645	*(0.038)*
Both parents arrived as adults	-0.434***
N = 1,056,898	(0.061)
One or more parents arrived as child	-0.565***
N = 607,747	(0.080)
Parents from same sending country	-0.455***
N = 1,529,548	(0.050)
Parents from different sending countries	-0.828***
N = 135,097	(0.197)
Baseline – 1 parent foreign	*-0.574****
N = 732,911	*(0.061)*
Only mother foreign	-0.406***
N = 271,996	(0.131)
Only father foreign	-0.641***
N = 460,915	(0.102)

Notes: Data from the complete-count 1920 census. Sample includes non-black children with two foreign-born parents (rows 1-5) or one foreign-born parent (rows 6-8) who were between the ages of zero and 18 and lived with their parents in a non-southern state. To observe complete birth order, sample restricted to households in which mother is less than 43 years old. All regressions control for child's age, a dummy equal to one if child has same name as father, and name frequency.

Appendix Table 6: Name foreignness and other labor market outcomes

	Mean	Without HH FE	With HH FE
Hourly wage	11.04	1.17	4.74
(*N* = 619,651)		(1.65)	(6.16)
Weeks worked per year	44.84	-0.807***	-0.963***
(*N* = 710,030)		(0.069)	(0.193)
Hours worked per week	43.80	-1.369***	-1.531***
(*N* = 667,237)		(0.072)	(0.204)
Employed	0.839	-0.025***	-0.027***
(*N* = 1,021,886)		(0.002)	(0.004)
Public emergency wkr	0.044	0.006***	0.003
(*N* = 858,214)		(0.001)	(0.003)
Self employed	0.161	-0.011***	-0.009***
(*N* = 836,973)		(0.002)	(0.005)

Note: Sample includes men matched between 1920 and 1940 complete-count Censuses. Men must be 0-18 in 1920, living at home with parents and living with a mother who is 43 years old or younger. Sample restricted to men whose fathers were foreign-born (second-generation immigrants) who were not self-employed in 1940. All regressions control for child's age and rank in birth order.