

## Do immigrants assimilate more slowly today than in the past?\*

Ran Abramitzky  
Stanford University and NBER

ranabr@stanford.edu

Leah Boustan  
Princeton University and  
NBER

lboustan@princeton.edu

Katherine Eriksson  
UC Davis and NBER

kaeriksson@ucdavis.edu

February 2019

Using millions of historical Census records and modern birth certificates, we document that immigrants assimilated into US society at similar rates in the past and present. We measure cultural assimilation as immigrants giving their children less foreign names after spending more time in the US, and show that immigrants erase about one-half of the naming gap with natives after twenty years both historically and today. Immigrants from poorer countries choose more foreign names upon first arrival in both periods but are among the fastest to shift toward native-sounding names. We find substantial cultural assimilation for immigrants of all education levels.

\* An earlier version of this paper that focused on the historical data was titled “Cultural Assimilation During the Age of Mass Migration.” We acknowledge generous research support from the Russell Sage Foundation. We benefited from helpful comments at NBER DAE Summer Institute, European Economic Association Meetings, Canadian Network of Economic Historians (CNEH) meeting, Annual Meeting of the Chilean Economy Society (SECHI), the Munich “Long Shadow of History” conference, the Irvine conference on the Economics of Religion and Culture, the Cambridge conference on Networks, Institutions and Economic History, the AFD-World Bank Migration and Development Conference, the Economic History Association, Bergen-Stavanger Workshop in Applied Microeconomics, and Cide and El Colegio de México’s “The Long Reach of Institutions” conference. We also thank participants of seminars at Arizona State, Berkeley, Columbia, Kiel Institute, Michigan, Montreal, Ohio State, Pittsburgh/Carnegie Mellon, Princeton, Queens, Santa Clara (Civil Society Institute), Stanford, UCLA, UCSD, Warwick, Wharton, Wisconsin and Yale. We benefited from conversations with Cihan Artunc, Sascha Becker, Hoyt Bleakley, Davide Cantoni, Raj Chetty, David Clingingsmith, Dora Costa, Dave Donaldson, Joe Ferrie, Amy Finkelstein, Price Fishback, Matt Gentzkow, Avner Greif, Eric Hilt, Matt Jackson, Naomi Lamoreaux, Victor Lavy, Joel Mokyr, Kaivan Munshi, Martha Olney, Luigi Pascali, Santiago Perez, Hillel Rapoport, Christina Romer, David Romer, Jared Rubin, Fabian Waldinger, Ludger Woessmann, Gavin Wright, and Noam Yuchtman. Jaime Arellano-Bover, Helen Kissel, James Reeves, David Yang, and Tom Zohar provided excellent research assistance.

## I. Introduction

Immigrant cultural assimilation is a sharply divisive political issue in the United States. Many politicians and voters hold the nostalgic view that European immigrants assimilated quickly in the past, while new immigrant groups do not attempt to assimilate today.<sup>1</sup> Donald Trump gained traction as the 2016 Republican presidential nominee by accusing immigrants of refusing to assimilate into US society.<sup>2</sup> More broadly, public views on immigration policy closely correlate with beliefs about the possibility of cultural assimilation, rather than with concerns about the labor market (Citrin, et al., 1997; Hainmueller and Hiscox, 2010; Hainmueller and Hopkins, 2014).

The passionate debate about the possibility of cultural assimilation is not new. Writing in 1891 about immigrants from Southern and Eastern Europe, Senator Henry Cabot Lodge asserted 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.” Progressive reformers sought to “Americanize” immigrants through policies like compulsory schooling and English-only laws (Lleras-Muney and Shertzer, 2015; Bandiera, et al., 2015). Eventually, swayed by a growing coalition led by Senator Lodge, Congress passed strict immigration quotas in the early 1920s,

---

<sup>1</sup> There are many expressions of this nostalgic view in public life. One typical example was articulated by talk show host Rush Limbaugh in 2014: “Italian-Americans came, and they became Americans. They held on to their traditions... but they were Americans first, not Italians first... They essentially were assimilating into a distinct American culture that they craved to be part of... What’s happening to immigration now is there is no desire to assimilate... they are coming here and demanding that America accommodate their culture.” President Trump also favorably compared past Norwegian arrivals to immigrants today (Aizenmann, 2018; Van Dam, 2018).

<sup>2</sup> Of Muslim assimilation, candidate Trump said: “Assimilation has been very hard. It’s almost, I won’t say nonexistent, but it gets to be pretty close. And I’m talking about second and third generation — for some reason there’s no real assimilation” (quoted in DelReal, 2016). Cultural assimilation was also a pivotal issue in the 2017 candidacy of far-right French politician Marine Le Pen, and the vote for British exit from the European Union.

putting an end to the first Age of Mass Migration (Higham, 1955; Goldin, 1994; King, 2009 and Abramitzky and Boustan, 2017).

Our paper is the first rigorous comparison of the cultural assimilation of immigrants in the past and the present. We focus on two waves of mass migration to the United States, the first from Europe (1850-1913) and the second from Asia and Latin America (1965-present). Measuring and comparing cultural assimilation is a challenge because information on many cultural practices – things like food, dress, and accent – are not systematically collected.<sup>3</sup> The premise of this paper is that we can trace cultural assimilation by examining shifts in the names that immigrants give their offspring as they spend more time in the US. We draw on a rich literature in sociology and economics suggesting that names are signals of cultural identity, thereby offering a revealing window into the nature of cultural assimilation.<sup>4</sup> One benefit of name selection as a measure of cultural assimilation is that it is a pure choice, unlike other standard measures, like intermarriage, that reflect both a choice (who you want to marry) and a constraint (who wants to marry you).<sup>5</sup>

---

<sup>3</sup> One measure of cultural assimilation is fluency in English. Yet the Census data on English ability is quite coarse. Before 1980, the Census question on speaking English simply asked whether a respondent could speak English (yes or no). In more recent years, the Census divided those speaking English as a second language into three options (speaks very well, well or not well). Ward (2019) finds that English fluency during the Age of Mass Migration was associated with only a small upgrade in occupational income.

<sup>4</sup> Lieberman (2000) is the classic reference in the sociology of naming and includes some discussion of the name choices of immigrant parents in the US. Zelinsky (1970) and Lieberman and Bell (1992) study differences in name choices by region and by parental education, respectively. On names and social outcomes across various contexts, see Lieberman and Mikelson (1995), Fryer and Levitt (2004), Bertrand and Mullainathan (2004) and Figlio (2005) on blacks; Sue and Telles (2007) on Hispanics; Algan, et al. (2013), Smith and Macrauld (2009), and Gerhards and Hans (2009) on immigrants to Europe; and Rubinstein and Brenner (2014) on ethnic groups in Israel.

<sup>5</sup> Inter-marriage has been used extensively in sociology as a marker of cultural assimilation (see, for example, Gordon, 1964; Alba and Golden, 1986; Lieberman and Waters, 1988; Pagnini and Morgan, 1990 and Wildsmith, et al., 2003). Angrist (2002) and Meng and Gregory (2005) employ these measures in economics.

Furthermore, name choices will capture assimilation that can take place within co-ethnic couples, as well as between couples of different backgrounds.

To use naming decisions to study cultural assimilation, our first step is to construct “Foreignness Indices” for each name in the past and present indicating the relative probability that a given name is held by immigrants versus the native born. We then compare the name choices that immigrants selected for their children after they spent more time in the US, tracing out an assimilation profile in both the past and the present. We find that the speed of name-based assimilation was roughly similar in both periods. Although immigrants did not completely converge with natives in name choices in one generation either historically or today, immigrants reduce the naming gap with natives by about one-half after spending twenty years in the US.

Furthermore, we find that the rate of cultural assimilation varies substantially by country of origin and in an unexpected way, with the groups most often accused of a lack of assimilation actually assimilating most rapidly. In the early twentieth century, Southern and Eastern Europeans started out with the most foreign-sounding names, yet shifted toward American-sounding names for their children most rapidly with time spent in the US. The same is true for Mexican and Vietnamese immigrants today. We find substantial cultural assimilation for both literate and illiterate immigrants in the past, and for immigrants at all education levels today, including for high school dropouts.

Despite perceived economic returns to having a native-sounding name,<sup>6</sup> immigrants may be inclined to give their children ethnic names in order to retain their original cultural identity. The

---

<sup>6</sup> Immigrants who adopt native-sounding first names appear to experience wage growth after the name change (Arai and Thoursie, 2009; Biavaschi, Giulietti and Siddique, 2016; Carneiro, Lee and Reis, 2015). Relatedly, in an audit study, Oreopoulos (2011) finds that resumes with more foreign-sounding last names receive fewer call backs in Canada. Goldstein and Stecklov (2016) and our working paper both document differences in outcomes between men with more/less

fact that parents did not fully adopt native naming patterns within the first generation highlights the value that immigrants place on maintaining their cultural identity, a pattern consistent with Akerlof and Kranton's (2000) and Benabou and Tirole's (2011) theories of the economics of identity and Bisin and Verdier's (2000) model of cultural transmission within families. Shifting toward native-sounding names with time spent in the US may reflect learning about US culture, identifying more with US culture over time, or making a final decision to stay in the US rather than to return home.<sup>7</sup>

Beyond allowing the first rigorous comparison between past and present, our paper is also one of the first to document cultural assimilation for first generation immigrants. Many papers in economics and other social sciences show that the *children of immigrants* readily join US society. Second generation immigrants have similar fertility behavior, labor force participation and political preferences to the children of the native born, with only minor differences associated with the culture of their parents' home country.<sup>8</sup> However, it has been a challenge to find large longitudinal datasets that allow researchers to follow shifts in first generation immigrants' behavior over time.

We emphasize that our paper has no normative implications. That is, we do not wish to imply that immigrants *should* assimilate culturally into the US society. In fact, arguably part of what makes a society flourish is its openness to cultural diversity. Our paper simply documents

---

foreign-sounding first names in the cross-section, but we see no differences between pairs of brothers with more/less foreign names who were raised in the same household.

<sup>7</sup> Gould (1980) and Bandiera, Rasul and Viarengo (2013) document high rates of return migration to Europe during the Age of Mass Migration, reaching at least 25 percent of the immigrant flow. Dustmann (2016) finds similar rates of return migration from the US today.

<sup>8</sup> On assimilation across generations in the Age of Mass Migration, see: Watkins, 1994; Watkins and London, 1994; Guinnane, Moehling and O'Grada, 2006; Foley and Guinnane, 1999; for contemporary studies, see: Fernandez and Fogli, 2009; Alesina and Giuliano, 2011; Luttmer and Singhal, 2011; Blau, et al., 2013.

what immigrants do in practice. We also acknowledge that our estimates only capture assimilation into the general “average” (and often white) society and they do not capture the multi-dimensional nature of culture in the US.

## II. Measuring the foreignness of given names

Our analysis proceeds in two steps. First, we use the names of foreign- and native-born adults to calculate a Foreignness Index for each first name in the population. Second, we assign this index to the names selected for the (native-born) children of immigrant mothers.

Our measure of name foreignness is based on the relative probability ( $R$ ) that a given name is held by a foreign-born versus native-born individual. The formula for  $R$  is given by:

$$R_{name} = \frac{\frac{\# foreigners_{name}}{total \# foreigners}}{\frac{\# natives_{name}}{total \# natives}}.$$

This measure has a natural interpretation: for example,  $R = 2$  implies that a name is twice as likely to be used in the immigrant population as in the native population.

The relative probability that a name is given to a U.S.-born child of foreign born parents is sensitive to outliers, especially for names that are unpopular among natives, which results in small values of the denominator. Thus, our main results are based on a normalized index used by Fryer and Levitt (2004) in the context of distinctly black names. In particular, the Foreignness Index is defined as:

$$Foreignness Index_{name} = 100 \cdot \frac{\frac{\# foreigners_{name}}{total \# foreigners}}{\frac{\# foreigners_{name}}{total \# foreigners} + \frac{\# natives_{name}}{total \# natives}}$$

and ranges from zero to 100, 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 100 assigned to a child whose first name is distinctively foreign. Note that the F-index is a simple function of  $R$ , equivalent to  $R/(1+R)$ .<sup>9</sup>

Calculating the F-index is very data intensive because it requires having accurate counts of foreign- and native-born residents with each given name. In the historical data, this is straightforward because the 1920 census contains information on the first name and country of birth of every resident of the US. However, because our modern data comes from birth records, we use the names of birth parents, some of whom were born abroad, to compute the modern F-index. Our preferred Foreignness Index in the historical data uses parents and non-parents, but we show below that results are nearly identical when constructing a historical F-index from the names of parents only.<sup>10</sup>

The foreignness of a given name changes over time as names come into and out of fashion. We thus construct the F-index from individuals born soon before the children in the sample. To ensure that we have enough data to accurately reflect each name, we construct the F-index from ten birth cohorts in each period (1895-1905 for the past; 1985-1995 for the present). Recall that our F-index in the modern data is based on the name of birth parents.<sup>11</sup> For the historical data, we

---

<sup>9</sup> Appendix Table 5 presents the pairwise correlation between alternative measures of cultural assimilation (intermarriage, English fluency, and application to US citizenship) and this names-based measure for the historical data. The correlations are highly significant and in the expected direction, providing validation for our measure of cultural assimilation. However, we acknowledge that we cannot compare any of the observable indicators of cultural assimilation with unobserved, home-based practices.

<sup>10</sup> We prefer using the F-index constructed from all names, rather than only the names of parents, when possible to maximize the number of observations used to measure the index value for each name. Yet the correlation between indices constructed with all names or with parents only is 0.954.

<sup>11</sup> Parents born between 1985 and 1995 are observed having children primarily after 2005.

also present results using a birth cohort specific F-index to better capture naming trends calculated from cohorts born in  $t-1$  to  $t-20$  for a child born in  $t$ .

One concern in the modern data is that birth certificates record the birth places (and thus foreign-born status) of mothers, but not of fathers. We impute birth places for men in the modern data by assuming that all foreign- (native-)born mothers were paired with foreign- (native-)born fathers.<sup>12</sup> As a result, our modern F-index is more accurate for girl names than for boy names. Alternatively, we calculate an F-index based on the *names of the children themselves*. In this case, we classify children of foreign-born mothers as “foreign” and children of native-born mothers as “native,” thereby estimating whether immigrants shift away from the names commonly given by other immigrants to the names commonly given by native-born mothers. This alternate F-index treats boy and girl names symmetrically in the modern data.

The average name gap between the children of foreign-born and native-born mothers was 19 points in the past and 18 points today. Appendix Table 1 lists the most foreign, neutral and native names for boys and girls for our main F-index in the past and present. The most foreign girl names in the past included Yiddish names like Yetta and Italian names like Carmela and Antonia. Names like Elizabeth and Catherine were equally common among the native- and foreign-born. Today, the most foreign names are Spanish names like Salma and names like Abigail and Olivia are equally common by mother’s nativity status.

---

<sup>12</sup> We acknowledge that imputing father’s birthplace will introduce some measurement error into the F-index values of boys’ names. Yet more than 83 percent of foreign-born women with a child present in the household living in California in 2000 had a foreign-born spouse (our calculation from IPUMS). Recall also that this assumption is only required for calculating the F-index, not for the analysis of immigrants’ name choices.

### III. Estimating the relationship between time in the US and name choice

We use the Foreignness Index to explore the naming choices of immigrant parents as they spend more time in the US. Ideally, we would estimate assimilation profiles using the following specification:

$$FIndex_{ijt} = \alpha_j + \beta_1 YearsUS_{ijt} + \beta_2 BirthOrder_{ijt} + \gamma_{ijt} + \varepsilon_{ijt} \quad (1)$$

The dependent variable is the Foreignness Index of child's  $i$ 's name born to mother  $j$  at time  $t$ . The main right-hand side variable is mother's years spent in the US by the child's year of birth  $t$ . By including mother fixed effects ( $\alpha_j$ ), the effect of mother's time in the US is identified by differences between siblings born in different years. We control for the child's rank in the birth order to account for traditions in some cultures to name oldest sons and daughters after relatives who likely have more foreign-sounding names. We also add a set of indicators for child's birth year ( $\gamma_{ij}$ ) in five-year bins to absorb secular trends in naming. We later show that results are robust to using three year instead of five year bins.

We can estimate equation 1 directly in the historical data by using standard Census variables to measure a child's birth year and his/her mother's years spent in the US at the time of birth.<sup>13</sup> However, we cannot use equation 1 for our past/present comparison because we do not observe mother's arrival year in the CA birth records. Instead, we use mother's age in birth year  $t$  as a proxy for mother's time spent in the US. This proxy relies on a demographic identity, whereby mother's years spent in the US by birth  $t$  is a function of mother's age in year  $t$  and mother's age at arrival in the US (*mother's years in the US in birth year  $t$  = mother's age in birth year  $t$  – age*

---

<sup>13</sup> In particular, we observe the ages of all children living in the household in the 1920 Census, along with mother's current age and arrival year in the US.

at arrival in the US). For example, a mother who arrived in the US at age 20 and had a child at age 30 would have been in the US for 10 years by the time of that birth. Controlling for a mother fixed effect absorbs a mother's age at arrival in the US, resulting in a one-for-one relationship between mother's years in the US and mother's age at birth year  $t$  (net of measurement error).

For our past/present comparison, we thus estimate the relationship between name foreignness and mother's age at the time of birth:

$$FIndex_{ijt} = \alpha_j + \beta_1 AgeMother_{ijt} + \beta_2 BirthOrder_{ijt} + \gamma_{ijt} + \varepsilon_{ijt} \quad (2)$$

To validate the use to mother's age as a proxy for mother's years spent in the US, we start by estimating both equations 1 and 2 in the historical data to demonstrate that the two approaches produce similar results (see Figure 1 below). Both age and year of arrival are measured with error in the Census; there is a well-known problem of both "age heaping" and "year heaping" on the zeroes and fives with more immigrants than expected reporting arriving in the US in, say, 1900 and 1905 than 1899 or 1901. A regression of mother's years in the US at birth on mother's age at birth and a set of mother fixed effects in the 1920 Census data results in a coefficient of 0.81 (st. err = 0.004) demonstrating that mother's years in the US and age at birth are closely but not perfectly related. Any difference between the estimates of equations (1) and (2) in the historical data is due to measurement error in either mother's arrival year or mother's age.<sup>14</sup>

Using mother's age at the time of birth as a proxy for mother's years in the US relies on the assumption that mother's age at birth does not have a *direct effect* on name choice beyond its

---

<sup>14</sup> We note that mothers arrive in the US at somewhat older ages today (age 24 on average) than they did in the past (age 20 on average). A direct comparison between our past and present results relies on the assumption that years in the US have a similar effect on cultural assimilation for women of somewhat different ages. This assumption seems innocuous in this context given that we find a very linear relationship between years spent in the US and name foreignness.

effect through years in the US. We provide two pieces of evidence to support this assumption. First, we consider the name choice of native-born mothers. Our assumption implies that there should be a decline in name foreignness with age for foreign-born mothers due to cultural assimilation but not for native-born mothers (which is what we find; see Figures 1 and 3 below). Second, we consider the names given to children who themselves were born abroad before their mothers moved to the US. We can only conduct this analysis in the historical data. Before migrating to the US, mothers may still select US-sounding names for their children prospectively if a move to the US is already planned, but we expect such assimilation to be more muted than for children born after migration. We indeed find a smaller reduction in name foreignness with mother's age for children born abroad than for children born in the US (see Appendix Figure 1 below).

#### **IV. The name choice of immigrant mothers in the early twentieth century**

This section presents results for the early twentieth century. Consistent with a process of cultural assimilation, we observe that immigrant mothers gave their children less foreign names as they spent more time in US. Panel A of Figure 1 graphs the estimated relationship between indicators for mother's years in the US and name foreignness for all immigrant households in the 1920 Census (equation 1); confidence intervals are suppressed because, with millions of observations, all estimates are highly statistically significant. Children born after their mother had spent 20 years in the US scored 13 points lower on the Foreignness Index relative to their siblings born upon their parents' first arrival.<sup>15</sup> We show separate results by gender below.

---

<sup>15</sup> In our main analysis, we focus on non-black children who were born outside the South because few immigrants lived in the South in 1920. We also restrict the sample to the children of mothers

In Panel B of Figure 1, we replace mother’s years in the US at the time of the birth with our proxy using mother’s age at the time of the birth (equation 2). For foreign-born mothers, we find a similar effect of time spent in the US (as proxied by mother’s age) on name foreignness – a reduction of 9.5 points on the Foreignness Index for twenty years of age. In contrast, native-born mothers actually shift slightly *toward* more foreign-sounding names as they age (by 1.4 points), perhaps because, in this historical period, mothers became more aware of a wider and more cosmopolitan set of names as they aged. Contrary patterns of mother’s age on name choice for native-born mothers reinforces our interpretation of naming shifts for immigrant mothers as cultural assimilation. The mean gap in the F-Index for the children of immigrants and natives in the 1920 Census was around 18 points, implying that immigrants closed half of this “cultural gap” with natives after spending some time in the US.<sup>16</sup>

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 abandoning elements of their cultural identity. Figure 2 documents that the speed of name-based assimilation differed by country of origin, reporting the linear effect of mother’s age in birth year  $t$ . Immigrants from English-speaking countries and from Scandinavia gave their children less foreign-sounding names on average (F-index of 39, compared to 35 for the native-born) and exhibit *no decline* in name foreignness with time spent in the US; Scotland is one exception. In contrast, immigrants from

---

who were less than 40 years old in 1920 to minimize inaccurate measures of birth order due to the departures of older sibling from the childhood home.

<sup>16</sup> If we instead use a birth-cohort specific F-index based on the 20 years of cohorts prior to a child’s birth, we find a reduction of 7.2 points on the F-index (Appendix Figure 2). We used this measure in a previous working paper that focused on the historical results (Abramitzky, Boustan, Eriksson, 2016).

German-speaking countries, as well as Southern and Eastern Europe (e.g., Russia and Portugal) gave their children highly-foreign names on average (F-index = 53) but also exhibited the fastest rates of name assimilation as they spent more years in the US (declines of around 0.5 points for each year in the US). This result is not simply due to “regression to the mean”; without concerted effort on the part of immigrants to adapt to US society, the names given by immigrant parents would not mechanically drift toward the native norm.<sup>17</sup>

Both literate and illiterate immigrants shifted away from foreign-sounding names with time spent in the US. Table 1 compares the pace of cultural assimilation by mother’s country of origin and literacy status for the six countries that exhibit assimilation in Figure 2 and have at least 1,000 mother-child observations in each category (literate/not); recall that the literacy rate for some countries was more than 95 percent and so this comparison is only feasible in some cases. Illiterate mothers from Germany and Finland assimilated more rapidly than their illiterate counterparts, whereas for mothers from Austria and Portugal this pattern was reversed. For Russian- and Italian-born mothers, literate and illiterate mothers assimilated at a similar pace.

## **V. The name choice of immigrant mothers in California today**

This section presents comparable results on the relationship between name foreignness and mother’s years in the US (as proxied by mother’s age at birth) in the modern data. Our modern

---

<sup>17</sup> There is little correlation between cultural and economic assimilation at the country level. In particular, the correlation between changes in occupation-based earnings from Abramitzky, Boustan and Eriksson (2014, Figure 3) and changes in name-based assimilation (Figure 2 here) is 0.04 (or 0.17 without Finland).

sample is the universe of births occurring in the state of California between 1989 and 2015.<sup>18</sup> California housed one third of the foreign-born population living in the US in 1990.

Our research design requires comparing multiple births that occur to the same mother. We link mothers with at least two lifetime births across birth records using mothers' first name, maiden name, places of birth, and exact birth date. We can only link births that occur after 1988 in California. Appendix Figure 3 documents the match rate by year. The match rate is lower at the beginning of the sample because many reported births occurred before the dataset began. Overall, we match 80 percent of native-born mothers and 67 percent of foreign-born mothers because foreign-born mothers have a larger share of births occurring outside of California. Panel A of Appendix Table 2 compares the characteristics of matched and unmatched mothers. Not surprisingly, matched mothers are more likely to be born in California. They are also less likely to be high school dropouts. Restricting the sample to births after 1998 to mothers born in California (Panel B) eliminates the educational differences between matched and unmatched observations. We thus replicate our modern results for the subsets of mothers who had all of their reported births in California or who had all reported births after 1998 and find very similar patterns in each case (see Appendix Figures 4 and 5; compare to Figure 3 described in next paragraph).

Figure 3 presents our main analysis of the contemporary data, reporting estimates of changes in the Foreignness Index of child's name with mother's age at birth from equation 2. We find that immigrant mothers give less foreign-sounding names to children when they are older (i.e., children who were born after their mother spent more time in US). After 20 years, mothers eliminate 9.8 points on the name Foreignness Index, very similar in magnitude to the historical

---

<sup>18</sup> Our dataset excludes 2011, which has incomplete information. There is a large literature in economics and public health using the California birth certificate data to look at birth outcomes. Currie and Moretti (2007) also link mothers over time by name, and date and state of birth.

data. As in the past, immigrants eliminate half of the naming gap with natives after 20 years in the US. Again, akin to the historical data, native-born mothers shift slightly toward foreign-sounding names with age (1.5 points after 20 years).<sup>19</sup>

Table 2 analyzes variation in the speed of cultural assimilation by group. Panel A subdivides foreign-born mothers by country of origin. Mexican immigrants give their children the most foreign-sounding names (F-index = 50), whereas other immigrants choose substantially less foreign names (average F-index = 26).<sup>20</sup> However, Mexican immigrants shift away from foreign-sounding names most rapidly, reducing the F-index by 0.51 points for each year spent in the US. Vietnamese immigrants also demonstrate a fast pace of assimilation (0.34 points per year), whereas immigrants from China and the Philippines shift toward native-sounding names at a slower rate.<sup>21</sup>

Panel B of Table 2 subdivides mothers within country of origin by highest degree obtained in the data. We focus on mothers from Mexico and Vietnam, the two countries that are separately identified in the data and exhibit substantial name-based assimilation. For these mothers, we find rapid assimilation in all educational categories. If anything, less educated mothers appear assimilate most rapidly, but we cannot statistically distinguish the coefficients across education groups.

---

<sup>19</sup> Results are similar in both past and present when using 3-year bands for children's year of birth to control for changes in name trends, rather than 5-year bands (see Appendix Figure 6).

<sup>20</sup> Anecdotally, Asian immigrants select native-sounding names on their birth certificate for the purpose of social interaction in the US while maintaining an Asian name for household use.

<sup>21</sup> The California birth certificates also separately identify mothers from Cuba, Japan and Canada. There are not enough observations in these sub-samples to analyze separately. The remainder of mothers in the data are combined into a "rest of the world" category. It is hard to interpret the coefficient for this group because mothers in this group come from many different sending countries who may have children at different ages on average.

## VI. Alternative measures of the F-index

Sociologists have documented that parents are more open to new or creative names for girls, while boys tend to receive a more traditional set of names, and so parents may be more willing to shift towards native-sounding names for daughters (Rossi, 1965; Sue and Telles, 2007). On the other hand, parents may be more concerned about the possible negative consequences of having an ethnic-sounding name in the labor market for sons than for daughters, especially in the past. Appendix Figure 7 shows that immigrants shift towards more native-sounding names somewhat more rapidly for daughters than for sons both in the past and the present (a disparity of five F-index points by gender after 20 years in the US in both periods).

In the modern data, differential results by gender could be driven by measurement error in the F-index of boy names (see Section II). Appendix Figure 8 reproduces results by gender using an alternative F-index constructed from the names of the children of foreign-born versus native-born mothers; this index treats boys and girls names symmetrically. Using this metric, differential name choice by gender attenuates in the historical data and disappears in the modern data. It appears that parents shift their naming behavior with time spent in the US at a roughly equal pace for sons and for daughters.

Appendix Table 3 presents results from alternative F-indices that provide insight into the strategies that immigrant families used to assimilate through name choice. Row 2 replaces the overall F-index with a country-specific index. This measure asks whether immigrants shift away from common ethnic names with time spent in the US. Estimates of name-based assimilation are similar in this case, suggesting that the relevant margin is foreign- versus native-sounding names. Immigrants avoided the names commonly used by other immigrant groups (that is, Italian immigrants did not shift from names like Antonio to names like Hyman, but instead choose neutral

names like John and Frank). Row 3 standardizes all names to their phonetic equivalents, rather than using raw names using the NYSIIS algorithm (for example, treating Roberto and Robert as the same name). Results are similar in the historical data but, in the modern data, the extent of name-based assimilation is cut in half. This pattern suggests that in California today, Anglicizing Spanish names is one strategy that immigrant parents use to assimilate. Indeed, inspecting common names at different percentiles of the F-index distribution reveals some names like Mary and Anna (low F-index) that have Spanish equivalents like Maria and Ana (high F-index).

Throughout the paper, we attempt to standardize the historical and modern data. However, there are still a few remaining differences that we investigate here. First, the historical data covers the entire United States, while the modern data is only for California. We replicated our historical analysis in Appendix Table 4 (row 1) for California only and find similar patterns. Alternatively, we try reweighting the California data so that the immigrant mothers in our sample reflect the country-of-birth distribution of immigrant mothers nationwide. In doing so, the reduction in name foreignness with time spent in the US attenuates somewhat (from -9.8 points to -8.4 points after 20 years in the US; see Appendix Figure 9), primarily because of the high share of Mexican-born mothers in California with high rates of name-based assimilation. Second, the modern data records legal name as reported on a birth certificate, while the historical data were collected by Census enumerators who may instead write down a different name that is used in social settings. We restrict the historical data in row 2 to children six years of age or younger who have not yet been to school, and thus would not have had the time or the pressure to change their birth name. Again, patterns are similar. Third, the F-index for the modern birth certificate data is based on the names of parents only, while the historical Census data uses parents and non-parents. We instead replicate

the results using an F-index based on parents only in the 1920 Census (Appendix Figure 10) and results are entirely comparable.

## **V. Conclusion**

We study the cultural assimilation of immigrants during two Ages of Mass Migration drawing on historical census data for the early twentieth century and California birth certificate records for today. Both then and now, immigrants chose less foreign names for their children as they spent more time in the US, reducing the difference in name choice with natives by about one-half both in the past and today after 20 years in the US. There are substantial differences in the pace of name-based assimilation by country of origin, with the immigrant groups most often accused of a lack of assimilation actually assimilating most rapidly. In the past, Portuguese and Russians were among the fastest to assimilate, and today Mexican immigrants have the fastest rate of name assimilation.

Our findings suggest that immigrants' identification with US culture grows stronger with time spent in the country. The gradual adoption of American-sounding names appears to have been part of a process of assimilation through which newcomers learned US culture, made a commitment to build roots in their adoptive country, and came to identify as Americans. The naming patterns also highlight the tradeoff that immigrant families face between maintaining their cultural identity and assimilating into society at large. Giving an ethnic-sounding name can enhance self-identification with an ethnic group but, at the same time, this signal of ethnic identity might generate discrimination from teachers and employers. This tradeoff is still salient for immigrants in the US and ethnic minorities around the world today (Ramakrishnan, 2015; Ali, 2015).

Overall though, lessons from the Age of Mass Migration suggest that fears that immigrants cannot or will not fit into American society are misplaced. It would be a mistake to determine immigration policy based on the belief that immigrants will remain foreigners, preserving their old ways of life and keeping themselves at arm's length from the dominant culture.<sup>22</sup> The evidence suggests that over time immigrant populations come to resemble natives, and that new generations form distinct identities as Americans.

---

<sup>22</sup> Fouka (2015) shows that forced assimilation policies can even generate backlash.

## 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).
- Abramitzky, Ran and Leah Platt Boustan. "Immigration in American Economic History." *Journal of Economic Literature*, 55, no. 4 (2017): 1311-45.
- Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. "Cultural Assimilation during the Age of Mass Migration." NBER Working Paper 22381, July 2016.
- Aizenman, Nurith. "Trump Wishes We Had More Immigrants From Norway. Turns Out We Once Did." *National Public Radio*, January 12, 2018.
- Akerlof, George A. and Rachel E. Kranton. "Economics and Identity." *Quarterly Journal of Economics* (2000): 715-53.
- 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.
- Ali, Wajahat. "For Muslim Americans: Baby Aiden or Baby Mohammed?" *New York Times* October 3, 2015.
- 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.
- Bandiera, Oriana, Imran Rasul, and Martina Viarengo. "The Making of Modern America: Migratory Flows in the Age of Mass Migration." *Journal of Development Economics* 102 (2013): 23-47.
- Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. "Nation-Building Through Compulsory Schooling During the Age of Mass Migration." Unpublished, University College London (2015).
- Benabou, Roland and Jean Tirole. "Identity, Morals and Taboos: Beliefs as Assets." *Quarterly Journal of Economics* 126 (2011): 805-855.

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." *Journal of Labor Economics* 2016.

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

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 (3) (1997): 858-881.

Currie, Janet and Enrico Moretti. "Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight." *Journal of Labor Economics* 25:2 (2007).

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

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.

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.

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

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

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

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

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

Limbaugh, Rush. "Hyphenated Americans and Immigration," *The Rush Limbaugh Show*, March 31, 2014.

Lleras-Muney, Adriana, and Allison Shertzer. "Did the Americanization Movement Succeed? An Evaluation of the Effect of English-Only and Compulsory Schooling Laws on Immigrants." *American Economic Journal: Economic Policy* 7, no. 3 (2015): 258-290.

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.

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.

Oreopoulos, Philip. "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes." *American Economic Journal: Economic Policy* 3, no. 4 (2011): 148-171.

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.

Ramakrishnan, Karthick. "'Strange' vs. 'Simple Old American' Names." *Los Angeles Times*. May 28, 2015.

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.

Smith, Malcolm T., and Donald M. MacRaild. "Paddy and Bidy No More: An Evolutionary Analysis of the Decline in Irish Catholic Forenames among Descendants of 19th Century Irish Migrants to Britain." *Annals of Human Biology* 36, no. 5 (2009): 595-608.

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

Van Dam, Andrew. "Norway Was Once the Kind of Country Trump Might've Spit On. Now Its People Don't Even Want To Come Here." *Washington Post*, January 12, 2018.

Ward, Zachary, "The Low Return to English Fluency during the Age of Mass Migration," Working paper (2019).

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

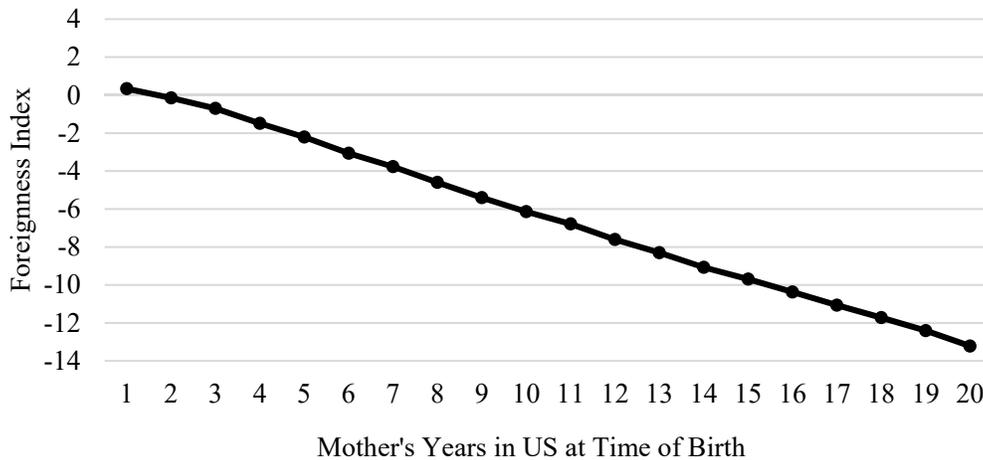
Watkins, Susan Cotts, and Andrew S. London. "Personal Names and Cultural Change: A Study of the Naming Patterns of Italians and Jews in the United States in 1910." *Social Science History* (1994): 169-209.

Wildsmith, Elizabeth, Myron P. Gutmann, and Brian Gratton. "Assimilation and intermarriage for US immigrant groups, 1880–1990." *The History of the Family* 8.4 (2003): 563-584.

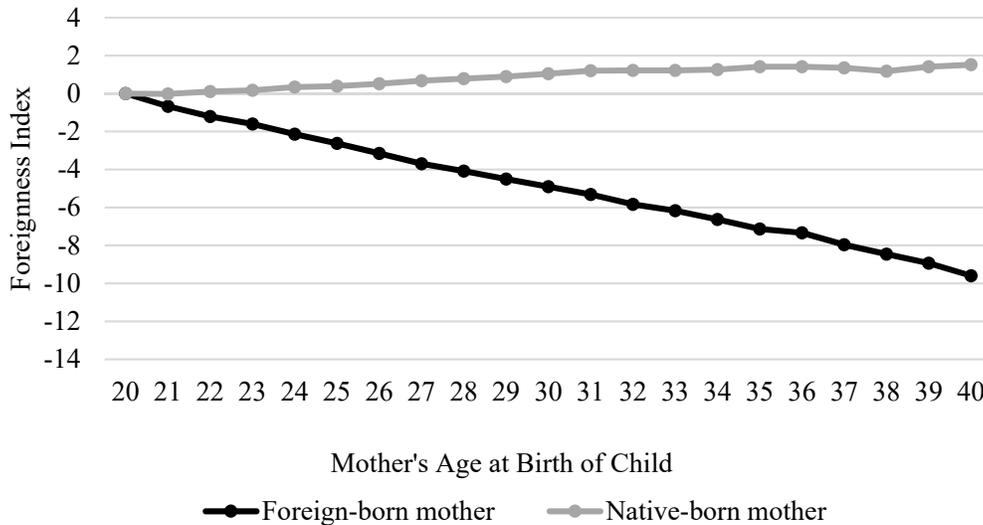
Zelinsky, Wilbur. "Cultural Variation in Personal Name Patterns in the Eastern United States." *Annals of the Association of American Geographers* 60, no. 4 (1970): 743-769.

**Figure 1: Immigrants selected less foreign names after spending time in US, 1920 census**

**Panel A: By mother's years in the US at the time of child's birth**

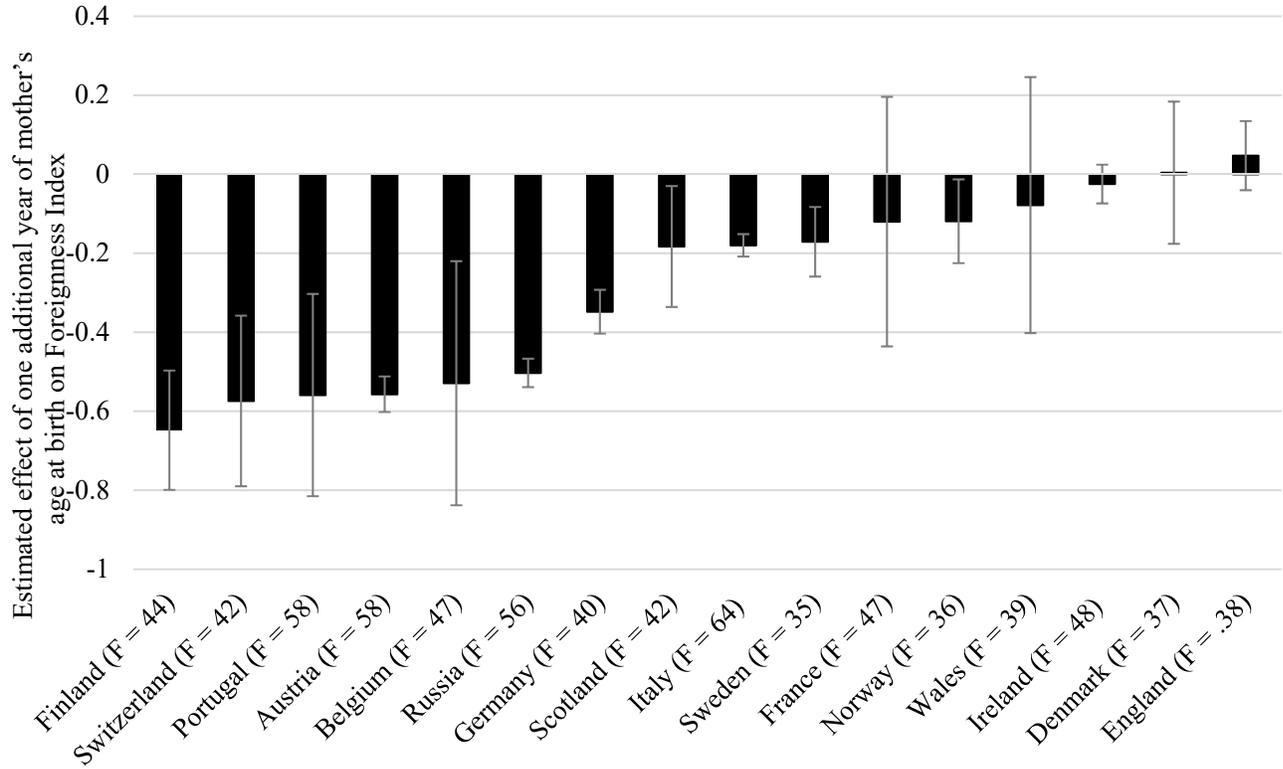


**Panel B: By mother's age at the time of child's birth**



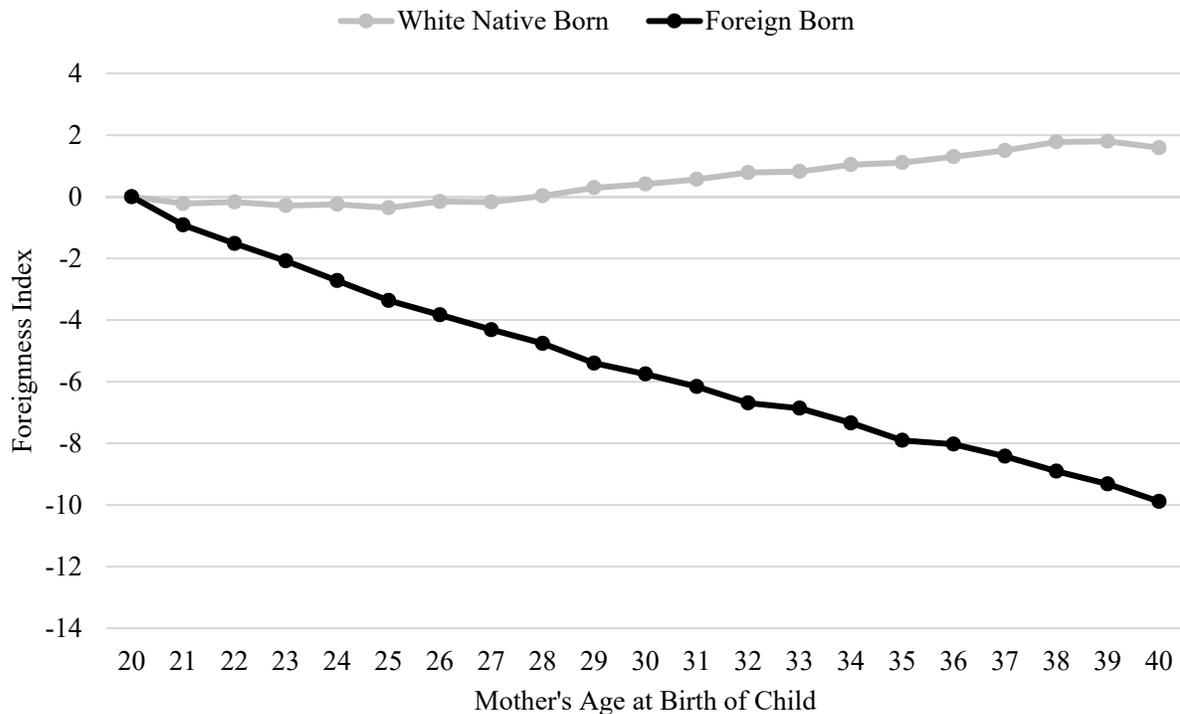
Notes: Panel A reports the coefficient estimates from equation 1, a regression of the Foreignness Index of a child's name on a set of indicators for mother's years in the US when the child was born. Panel B reports the coefficient estimates from equation 2, a regression of the Foreignness Index of a child's name on a set of indicators for mother's age when the child was born. Regressions also include child's rank in the birth order, indicators for child's age in five-year bands and a set of mother fixed effects. Data is from the complete-count 1920 census. Sample includes white children aged 0-15 who were born in a non-Southern state and were living with their parents in 1920. All children have mothers who were less than 43 years old in 1920 so that we can observe a (relatively) complete birth order, and who were between 20-40 years old at the time of the child's birth. Panel A includes only mothers who had been in the US between 0-20 years at the time of the child's birth. For Panel A,  $N(3,735,637)$ . For Panel B,  $N(\text{foreign born}) = 4,425,208$ ;  $N(\text{native born}) = 9,645,775$ . The sample size is smaller in Panel A because of missing information on mother's year of arrival in the US. We suppress confidence intervals because regressions are based on millions of observations and all coefficients are highly significant.

**Figure 2: Effect of time in US on name foreignness by sending country, 1920 census, by mother's age at birth**



Notes: Reported coefficients from equation 2 estimated separately by mother's country of birth. Regressions control for child's rank in the birth order, indicators for child's age in five-year bands and a set of mother fixed effects. Country labels also report average F-index for names given to children of mothers from each country of origin. Data is from the complete-count 1920 census. Sample details in notes to Figure 1.

**Figure 3: Immigrants select less foreign names after spending time in US, California birth certificates (1989-2015), by mother's age at birth**



Notes: This figure reports coefficient estimates of equation 2, a regression of the Foreignness Index of a child's name on a set of indicators for mother's age when the child was born. Regressions are estimated separately for children of foreign-born (N = 3,841,048) and white native-born mothers (N = 4,160,299). Regressions also include child's rank in the birth order, indicators for child's age in five-year bands and a set of mother fixed effects. The sample includes all children born to a mother aged 20-40 in California from 1989-2015. We suppress confidence intervals because regressions are based on millions of observations and all coefficients are highly significant.

**Table 1: Effect of time in US on name foreignness, 1920 census  
By mother's country of origin, literacy status, and mother's age at birth**

		Dependent Variable = Foreignness Index	
		Mother is...	
		Literate	Not literate
Austria	Mom age at birth	-0.628*** (0.04)	-0.397*** (0.092)
	<i>N</i>	296,480	63,820
Finland	Mom age at birth	-0.633*** (0.121)	-1.774*** (0.579)
	<i>N</i>	50,619	2,935
Germany	Mom age at birth	-0.341*** (0.043)	-0.882** (0.358)
	<i>N</i>	254,979	5,119
Italy	Mom age at birth	-0.216*** (0.03)	-0.207*** (0.051)
	<i>N</i>	508,047	219,065
Portugal	Mom age at birth	-0.825*** (0.258)	-0.252 (0.26)
	<i>N</i>	11,241	10,390
Russia	Mom age at birth	-0.576*** (0.029)	-0.666*** (0.096)
	<i>N</i>	529,041	73,898

Notes: This table reports coefficients from estimates of equation 2, a regression of the Foreignness Index of a child's name on a linear measure of mother's age at the time of birth. Regressions control for child's rank in the birth order, indicators for child's age in five-year bands and a set of mother fixed effects, and are conducted separately by mother's country of origin and literacy status. All data is from the 1920 complete-count census. Sample details in notes to Figure 1. The six countries included in the table exhibit assimilation in Figure 2 and have at least 1,000 mother-child observations in each category (literate/not).

**Table 2: Cultural assimilation by group, California data (1989-2015)**

Dependent Variable = Foreignness Index

	(1)	(2)	(3)	(4)	(5)
<b>Panel A: By country of origin</b>					
	Mexico	China	Philippines	Vietnam	Rest world
Age at birth	-0.513*** (0.025)	-0.100 (0.136)	0.025 (0.124)	-0.341*** (0.088)	-0.358*** (0.049)
<i>N</i>	2,435,812	164,149	123,895	146,403	907,969
<i>Mean F-index</i>	49.7	25.0	22.4	23.5	34.0
<b>Panel B: By education level and country of origin</b>					
	Less HS	HS grad	Some college	College plus	
Vietnam	-0.374 (0.318)	-0.386** (0.166)	-0.323* (0.175)	-0.249* (0.138)	
<i>N</i>	21,296	44,331	32,784	46,190	
Mexico	-0.547*** (0.035)	-0.461*** (0.045)	-0.489*** (0.067)	-0.516*** (0.105)	
<i>N</i>	1,397,249	625,503	270,753	121,395	

Notes: This table reports coefficients from estimates of equation 2, a regression of the Foreignness Index of a child's name on a linear measure of mother's age at the time of birth. Regressions control for child's rank in the birth order, indicators for child's age in five-year bands and a set of mother fixed effects, and are conducted separately for each sub-sample. See the notes to Figure 3 for additional details on the sample. Robust standard errors are reported in parentheses. \*\*\* = significant at 1 percent level, \*\* = significant at 5 percent level, \* = significant at 10 percent level.