

Do immigrants assimilate more slowly today than in the past?

By RAN ABRAMITZKY, LEAH BOUSTAN, AND KATHERINE ERIKSSON*

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.

* Abramitzky: Department of Economics, Stanford University, Stanford, CA 94305 and NBER (email: ranabr@stanford.edu); Boustan: Princeton University, Department of Economics, Princeton, NJ 08544 and NBER (email: lboustan@princeton.edu); Eriksson: University of California-Davis, Department of Economics, Davis, CA 95616, NBER, and Stellenbosch University (email: kaeriksson@ucdavis.edu). 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, Chris Becker, 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 Pascoli, Santiago Perez, Alain Pineda, 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.

The cultural assimilation of immigrants is a divisive political issue in the United States. Many politicians and voters hold the nostalgic view that European immigrants assimilated quickly in the past, in contrast to immigrants today.¹ However, this passionate debate about cultural assimilation is not new. Writing in 1891, Senator Henry Cabot Lodge asserted that immigration from Southern and Eastern Europe “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 compulsory schooling and English-only laws (Lleras-Muney and Shertzer, 2015; Bandiera, et al., 2015). Eventually, swayed by this growing coalition, Congress passed strict immigration quotas in the early 1920s, ending the Age of Mass Migration (Higham, 1955; Goldin, 1994; King, 2009 and Abramitzky and Boustan, 2017).

Our paper provides the first quantitative comparison of the cultural assimilation of immigrants in the past and the present. We focus on two waves of migration to the United States, the first mostly from Europe (1850-1913) and the second from Asia and Latin America (1965-present). Measuring cultural assimilation is challenging because information on cultural practices – things like food, dress, and accent – are not systematically collected.² The premise of this paper is that we can trace cultural assimilation by examining shifts in the names that immigrants give their children as they spend time in the US. We draw on a rich literature in sociology and economics suggesting that names are signals of cultural identity.³ One benefit of name selection as a measure of cultural assimilation is that it is a pure choice, unlike intermarriage, which reflects both a choice

¹ 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.”

² One counterpoint is Census data on self-reported ability with spoken English (Ward, 2019).

³ Lieberman (2000) is the classic reference in the sociology of naming. 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 Macrailld (2009), and Gerhards and Hans (2009) on immigrants to Europe; Rubinstein and Brenner (2014) on ethnic groups in Israel, and Olivetti and Paserman (2015) on intergenerational mobility in the United States between 1850 and 1940.

(who you want to marry) and a constraint (who wants to marry you).⁴ Furthermore, name choices will capture assimilation that can take place within co-ethnic couples, as well as between couples of different backgrounds.

We construct “Foreignness Indices” for each name in the past and present indicating the relative probability that a given name was held by immigrants versus the US-born. We then compare the name choices that immigrants selected for their children after they spent more time in the US. For the past, we use the 1920 Population Census and for the present, we use California birth certificate records for 1989-2015. Using different data sources raises obvious comparability issues. The 1920 census covers the entire US population, records names as reported to the enumerator, and includes information about immigrants’ year of arrival to the US. In contrast, the birth certificate records cover only the state of California, contain an infant’s legal name, and most importantly do not include direct information about years of arrival. We try to standardize the historical and modern data, and Section V further investigates sensitivity to these differences.

We find that the *rate* of name-based assimilation was similar in both periods. Although immigrants did not completely converge with the US-born in name choices in one generation either historically or today, immigrants reduce the naming gap with the US-born 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, with groups that started with high foreign sounding names switching more rapidly to more American sounding names (Southern and Eastern Europeans in the past, 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.

Despite perceived economic returns to having a native-sounding name,⁵ immigrants may give their children ethnic names in order to retain their original cultural identity. The fact that

⁴ Intermarriage 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) study intermarriage in economics.

⁵ Immigrants who adopt native-sounding first names experience wage growth after the name change (Arai and Thoursie, 2009; Biavaschi, Giulietti and Siddique, 2016; Carneiro, Lee and Reis, 2015). Relatedly, Oreopoulos (2011) finds that resumes with more foreign-sounding last names receive fewer call backs in Canada. Goldstein and Stecklov (2016) and Abramitzky, Boustan and Eriksson (2016) both document differences in outcomes between men with more/less foreign-sounding first names in the cross-section, but we see much smaller differences between pairs of brothers with more/less foreign names.

parents do not fully adopt native naming patterns within the first generation highlights the value that immigrants place on maintaining their cultural identity, 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 may reflect learning about US culture, a greater identification with US culture over time, or making a final decision to stay in the US rather than return home.⁶

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 US-born, with only minor differences associated with their parents' home country.⁷ 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 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 what immigrants do in practice.

I. Data and Definitions

A. Datasets with information on names

For the past, we use the 1920 Census to reconstruct the names and ages of children living in the household with immigrant and US-born mothers. We also observe mother's current age, place of birth, and arrival year in the US.⁸ Because modern Census data does not reveal names,

⁶ Gould (1980) and Bandiera, Rasul and Viarengo (2013) document high rates of return migration to Europe in the past. Dustmann (2016) finds similar rates of return migration today.

⁷ 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.

⁸ 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 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.

our modern sample is the universe of births occurring in the state of California between 1989 and 2015 (excluding 2011, which is incomplete).⁹ California housed one-third of the foreign-born population in 1990. Birth certificates report the names of children, their mothers, and often their fathers. California birth certificates do not contain information on age of arrival, which is a key limitation.

Our research design compares multiple births from the same mother. In the 1920 Census, we directly observe all children living in the household. For the modern data, we instead link mothers across births occurring in California using mothers' first name, maiden name, birth place, and exact birth date. Overall, we match 80 percent of US-born mothers who report having two or more children (as recorded in the parity measure). We match a lower share of foreign-born mothers (67 percent) because a larger proportion of their births occurred outside California. Appendix Figure 1 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. Panel A of Appendix Table 1 compares the characteristics of matched and unmatched mothers. Matched mothers have higher levels of education than unmatched mothers. Panel B documents that these educational differences are driven primarily by the lower match rates for immigrants and mothers born outside of California, suggesting that the matching process itself does not select on the basis of educational attainment.

B. Measuring the foreignness of given names

We construct an index of name foreignness for every given name, based on the relative probability that the name is held by someone born in the US or abroad. The formula is given by:

$$R_{name} = \frac{\frac{\# \text{foreigners}_{name}}{\text{total} \# \text{foreigners}}}{\frac{\# \text{natives}_{name}}{\text{total} \# \text{natives}}}.$$

This measure has a natural interpretation: for example, $R = 2$ implies that a name is twice as likely to be used by foreigners as by natives.

⁹ 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.

The formula for relative probability is sensitive to outliers. We thus adopt the normalized index used by Fryer and Levitt (2004) to measure 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 distinctively native names and a value of 100 reflecting distinctively foreign names. Note that the F-index is a simple function of R , equivalent to $R/(1+R)$.¹⁰

Calculating the F-index is data intensive because it requires having counts of foreign- and US-born residents with each given name. This is straightforward in the historical data because the 1920 census contains information on the first name and birth country of every resident of the US. However, in the modern birth records, we can only use the names of birth parents to compute the modern F-index. A further 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-born mothers were paired with foreign-born fathers, and likewise for US-born mothers.¹¹ 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; parents born between 1985-1995 for the present).¹² Appendix Table 3 lists the most foreign, neutral and native names for boys and girls for our main F-index in the past and present.

Section V presents results with alternative versions of the Foreignness Index. Our preferred historical index uses the names of the entire population (both parents and non-parents), but we show below that results are nearly identical when using the names of parents only (the correlation

¹⁰ Appendix Table 2 presents strong pairwise correlation between alternative measures of cultural assimilation (intermarriage, English fluency, and application to US citizenship) and our names-based measure for the historical data. However, we acknowledge that we cannot compare any of the observable indicators of cultural assimilation with unobserved, home-based practices.

¹¹ 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). Yet, imputing father's birthplace will introduce some measurement error into the F-index values of boys' names.

¹² Parents born between 1985 and 1995 are observed having children primarily after 2005.

between the two indices is 0.954). For the historical data, we 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 . We address measurement error in boys' names in the modern data by calculating an F-index from the names of the children themselves, rather than from birth parents. In this case, we classify children of foreign-born mothers as “foreign” and children of US-born mothers as “native.”

II. Estimating the relationship between time in the US and name choice

Conceptually, we are interested in comparing the two periods of mass migration in both the *level* of assimilation (gap) and the *rate* of assimilation (closing the gap) between the US-born and foreign born. We start by describing our empirical exercise and then we discuss the identification assumptions required to compare both the level and rate of assimilation across the two periods.

We first consider the *rate* of assimilation. A natural approach to estimating the effect of years spent in the US on name foreignness of children would be to consider the following specification for foreign-born mothers:

$$FIndex_{ijt} = \alpha_j + \beta_1 \sum_{m=1}^{20} \mathbf{1}[YearsUS_{ijt} = m] + \beta_2 \sum_{k=1}^K \mathbf{1}[Birthorder_{ijt} = k] + \gamma_{ijt} + \varepsilon_{ijt} \quad (1)$$

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 a set of dummy variables for the number of years a mother spent in the US at the child's year of birth t . The omitted category is women who give birth in their year of arrival to the US ($YearsUS = 0$). By including mother fixed effects (α_j), we identify the effect of time spent in the US using name differences between siblings born in different years. We control for the child's rank in the birth order with a set of dummy variables to account for traditions in some cultures to name oldest sons and daughters after relatives.¹³ Note that child's birth year dummy variables would be collinear with mother's years in the US and the mother fixed effect,

¹³ Appendix Figure 2 shows results that interact birth rank with mother's country of origin to account for differences in naming traditions. Results are similar.

which absorbs mother’s year of arrival. We thus add child’s birth year (γ_{ij}) in five-year bins to absorb secular trends in naming.¹⁴ We later show that results are robust to using three year instead.

We cannot use equation 1 for a past/present comparison because we do not observe mother’s arrival year in the CA birth records. In addition, equation 1 does not allow us to compare the name choices of US-born and foreign-born mothers. Instead, our main analysis uses mother’s age in birth year t as a proxy for mother’s time spent in the US. This proxy relies on a demographic equation relating mother’s years spent in the US by birth year t to her age in birth year t and in her year of arrival to the US (*mother’s years in US by birth year t = mother’s age in birth year t – mother’s age at arrival in 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 birth. Controlling for a mother fixed effect absorbs a mother’s age at arrival in the US, allowing us to use mother’s age at birth year t as a proxy for mother’s years in the US (net of measurement error).¹⁵ The key assumption underlying our interpretation of this proxy is that mother’s age at birth does not directly affect the rate or level of assimilation.

For our past/present comparison, we thus estimate the relationship between name foreignness and mother’s age at the time of birth for mothers between the ages of 20 and 40:

$$FIndex_{ijt} = \alpha_j + \pi_1 \sum_{m=21}^{40} \mathbf{1}[AgeMother_{ijt} = m] + \pi_2 \sum_{k=1}^K \mathbf{1}[Birthorder_{ijt} = k] + \gamma_{ijt} + \varepsilon_{ijt} \quad (2)$$

The main right-hand side variable here is a set of dummy variables for the mother’s age at the child’s year of birth t . The omitted category is mothers who were 20 years old at a child’s birth.

We provide two pieces of evidence to support the assumption that mother’s age at birth does not have a *direct effect* on name choice beyond its effect on years spent in the US. First, we consider the name choice of US-born mothers as they age. Our assumption implies that there should be a decline in name foreignness with mother’s age at birth (= years in the US) for foreign-

¹⁴ A common approach to addressing the standard age-period-cohort collinearity problem is to group one of these categories into bins (e.g., Card and Lemieux, 2000 impose that the effect of some cohorts are the same).

¹⁵ 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. 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), rather than 1.0, demonstrating that mother’s years in the US and age at birth are closely but not perfectly related.

born mothers due to cultural assimilation but not for US-born mothers, which is indeed what we find (see Figure 1). Second, we consider the names given to the children of immigrants who themselves were born abroad before their mothers moved to the US. 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. We indeed find a substantially smaller reduction in name foreignness with mother's age for children born abroad than for children born in the US (see Appendix Figure 3). Note that we can only observe children born abroad in the historical data.

Unless assimilation proceeds at a constant pace for each year spent in the US, we need to additionally assume that the spacing between the mother's age at arrival in the US and her age at first birth in the US is the same in the past and in the present in order to compare the rates of assimilation across these two time periods. Imagine that immigrant mothers in the past had their first child after two years in the US, whereas immigrant mothers today wait five years before having their first child. If the rate of cultural assimilation is fastest in the first few years in the country, then we would overstate assimilation in the past relative to the present. However, we find a very linear relationship between mother's age and name foreignness in both periods, suggesting that differential spacing between age at arrival and first birth would not likely affect our conclusions.

Without information on age of arrival in the modern data, we must proceed with caution in making statements about the *levels* of the naming gap between foreign and native born mothers upon first arrival or by years spent in the US. Mothers arrive in the US at somewhat older ages today (age 24 on average) than they did in the past (age 20 on average).¹⁶ Our estimation approach uses mother's age at birth as a proxy for mother's years in the US, which implicitly considers age 20 to be the first year in the US in both time periods. Thus, we potentially *overstate the naming gap in levels* between immigrant and native mothers in the modern data at any given number of years spent in the US. For example, at age 40, mothers will have had only 16 years in the US in which to assimilate (on average), but we will treat them as if they have been in the US for 20 years. Given we find that that immigrants erase about -0.5 F-index points per year in the US, a back of an envelope calculation that accounts for differences in age of arrival suggests that we overstate

¹⁶ Calculations from 1920 and 2000 Censuses, respectively.

the level of the naming gap in the modern data relative to the historical data by 2 points ($= -0.5 * 4$ years) after 20 years in the US.¹⁷

III. The name choice of immigrant mothers, past and present

A. Core results

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 in both the past and the present. Figure 1 graphs the implied Foreignness Index of US-born and foreign-born mothers by mother's age at the time of birth for the past (Panel A) and the present (Panel B). In particular, we report the constant plus the estimated effect of the individual mother's age at birth indicators from equation (2), allowing us to show both the initial level and rate of assimilation.¹⁸ Appendix Figure 4 instead compares the *rate* of assimilation in both periods by directly graphing the coefficients on each mother's age at birth indicator relative to the omitted category of mothers who gave birth at age 20.

In the past, the children of recently-arrived immigrant mothers (mothers who are 20 years old at birth) were given names that averaged 55 points on the F index, and children of US-born mothers who gave birth at age 20 were given names that averaged 34 points. This initial disparity amounted to a 21 point gap between young immigrant and US-born mothers. As mothers age, the US-born continued to give equally foreign sounding names, but foreign-born mothers reduced the average F-index by around 10 index points over 20 years, erasing half of the naming gap between immigrants and natives. At the end of 20 years, foreign-born mothers chose names that averaged 45 points on the F-index and US-born mothers chose names that averaged 35 points on the index, a difference of only 10 points. Similarly, in the modern period, the initial gap in F-index between

¹⁷ Note that concerns arising from the missing information on age at arrival affects our interpretation of the estimates, but does not bias the estimates themselves. If one is interested in the naming gap by mother's age at birth, then our estimates can be used directly without considering differences in age at arrival.

¹⁸ An alternative way to measure the initial naming gap is to pool native- and foreign-born mothers into a single regression without mother's fixed effects. In this case, we can include an indicator for being foreign-born and interact this indicator with mother's age at birth dummy variables. The coefficient on the foreign-born indicator reveals the names gap for mothers who are 20 years old at birth. The coefficient was 22.33 in the past and 20.13 in the present. A gap of this size is also apparent in Figure 1 for mothers in their young 20s.

young foreign- and US-born mothers was 24 points (48 points on the F-index for the foreign born and 24 points for the US-born).¹⁹ After 20 years, the foreign born mothers reduced the average F-index by a similar degree (around 10 index points), erasing just under half of the naming gap with the US-born. The foreign-native gap at the end of 20 years was 12 points (37 points in the F-index for the foreign born and 25 points for the native born). One notable difference between the two periods is that both immigrant and US-born young mothers chose less-foreign names today.²⁰

We return to the more natural specification for cultural assimilation that uses mother's years in the US, rather than mother's age at birth, in Figure 1 (Panel C). We can only implement this specification (equation 1) in the historical data. 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.²¹

B. Heterogeneity by country of origin and education

Immigrants from sending countries that are culturally distant from the US or that face high levels of discrimination may enjoy 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 differs by country of origin in both past (Panel A) and present (Panel B). For brevity, we report estimates from a modified version of equation (2) that replaces indicators for individual years of age with a linear effect of mother's age in birth year.

In the past, immigrants from Western European countries like England and Denmark gave their children less foreign-sounding names on average (F-index of 39-40, compared to 35 for the US-born) but exhibited *little to no decline* in name foreignness with time spent in the US. In contrast, immigrants from countries like Portugal, Finland, Austria and Russia gave their children highly-foreign names on average (F-index = 59) but also exhibited faster rates of name assimilation

¹⁹ We replicate modern results for the subsets of mothers who had all of their reported births in California or who had all reported births after 1998; these restrictions help to balance the matched and unmatched mothers. Results are very similar in this case (see Appendix Figures 5 and 6).

²⁰ For both past and present, we show results using 3-year bands for children's year of birth to control for changes in name trends, rather than 5-year bands in Appendix Figure 7.

²¹ 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 8).

as they spent more years in the US (declines of 8-12 index points after 20 years 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.²²

Today, 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).²³ However, Mexican immigrants shift away from foreign-sounding names most quickly, reducing the F-index by 10 points after 20 years spent in the US. Vietnamese immigrants and immigrants from other countries (“rest of world”) also demonstrate rapid assimilation (a reduction of 7 points after 20 years in the US), whereas immigrants from China and the Philippines exhibit little name shifting.²⁴

Immigrants of all literacy and education levels shift 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 (Panel A, historical data) or education (Panel B, modern data); as above, we report the coefficient on a linear effect of mother’s age at birth year on name foreignness. Panel A reports these estimates for the six countries that have at least 2,000 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. We find substantial assimilation in all cases, with no clear pattern by literacy status. Illiterate mothers from Finland, Germany, and Portugal assimilated more rapidly than their literate counterparts, whereas for mothers from Austria, Italy and Russia this pattern was reversed.

Panel B subdivides mothers in the modern data by country of origin and highest degree obtained in the data. For mothers from Mexico and Vietnam, the two countries that exhibit

²² 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).

²³ 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.

²⁴ 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.

substantial name-based assimilation, 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.

IV. Sensitivity and Data Limitations

A. Data comparability issues

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 replicate our historical analysis in Appendix Table 4 (row 1) for the state of California only; we still find substantial assimilation but the pace was somewhat slower than in the national sample.²⁵ Alternatively, we try reweighting the modern 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 -10 points to -8 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. Individuals may also change their names after birth in order to assimilate. This possibility introduces the potential role of measurement error in birth names in the 1920 Census. In this case, particularly foreign names given by young mothers may be changed later on, thereby masking some of the actual assimilation that took place. Although we cannot fully address this issue, we restrict the historical data in Appendix Table 4 (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. Patterns are similar to the full sample. 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.

²⁵ California was the state with the 32nd fastest pace of assimilation in the past, out of 48 states.

B. Heterogeneity by gender

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 11 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 12 reproduces results by gender using an alternative F-index constructed from the names of the children of foreign-born versus US-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.

C. Alternative measures of the F-index

Appendix Table 5 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. 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).

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 the US-born by about one-half both 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 slow assimilation assimilating most rapidly.

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 by which newcomers learned US culture, made a commitment to build roots in their adoptive country, and came to identify as Americans. Naming patterns also highlight the tradeoff that immigrant families face between maintaining their cultural identity and assimilating into society at large.

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.²⁶ The evidence suggests that over time immigrant populations come to resemble the US-born, and that new generations form distinct identities as Americans.

²⁶ Fouka (2015) shows that forced assimilation policies can even generate backlash.

References

- Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122, no. 3.
- Abramitzky, Ran and Leah Platt Boustan. 2017. "Immigration in American Economic History." *Journal of Economic Literature*, 55, no. 4: 1311-45.
- Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2016. "Cultural Assimilation during the Age of Mass Migration." NBER Working Paper 22381.
- Akerlof, George A. and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics*: 715-53.
- Alba, Richard D. and Reid M. Golden. 1986. "Patterns of Ethnic Marriage in the United States." *Social Forces* 65, no. 1.
- 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. 2013. "The Economic Incentives of Cultural Transmission: Spatial Evidence from Naming Patterns across France." Manuscript.
- Arai, Mahmood, and Peter Skogman Thoursie. 2009. "Renouncing Personal Names: An Empirical Examination of Surname Change and Earnings." *Journal of Labor Economics* 27, no. 1: 127-147.
- Bandiera, Oriana, Imran Rasul, and Martina Viarengo. 2013. "The Making of Modern America: Migratory Flows in the Age of Mass Migration." *Journal of Development Economics* 102: 23-47.
- Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. 2015. "Nation-Building Through Compulsory Schooling During the Age of Mass Migration." Unpublished, University College London.
- Benabou, Roland and Jean Tirole. 2011. "Identity, Morals and Taboos: Beliefs as Assets." *Quarterly Journal of Economics* 126: 805-855.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94, no. 4: 991-1013.
- Biavaschi, Costanza, Corrado Giulietti, and Zahra Siddique. 2016. "The Economic Payoff of Name Americanization." *Journal of Labor Economics*.

- Bisin, Alberto and Thierry Verdier. 2000. "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits." *Quarterly Journal of Economics*: 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. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *Journal of Labor Economics*: 463-489.
- California Vital Records Data and Statistics. 1989-2015. *California Department of Public Health*.
- 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. 1978 . "The Effect of Americanization on the Earnings of Foreign-born Men." *Journal of Political Economy*: 897-921.
- Currie, Janet and Enrico Moretti. 2007. "Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight." *Journal of Labor Economics* 25:2.
- 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. 1999. "Did Irish Marriage Patterns Survive the Emigrant Voyage? Irish-American Nuptiality, 1880-1920." *Irish Economic and Social History* 26: 15-34.
- Fouka, Vasiliki. 2015. "Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I." Manuscript.
- Figlio, David N. 2005. "Names, Expectations and the Black-White Test Score Gap." National Bureau of Economic Research Working Paper No. w11195.
- Fryer, Roland G., and Steven Levitt. 2004 . "The Causes and Consequences of Distinctively Black Names." *Quarterly Journal of Economics* 119, no. 3: 767-805.
- Gerhards, Jürgen, and Silke Hans. 2009. "From Hasan to Herbert: Name-Giving Patterns of Immigrant Parents between Acculturation and Ethnic Maintenance." *American Journal of Sociology* 114, no. 4: 1102-1128.
- Goldin, Claudia. 1994. "The Political Economy of Immigration Restriction in the United States, 1890 to 1921." *The Regulated Economy: A Historical Approach to Political Economy*, pp. 223-258. University of Chicago Press.

- 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. 1964. *Assimilation in American Life: The Role of Race, Religion, and National Origins*. New York: Oxford University Press.
- 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. 2006. "The Fertility of the Irish in the United States in 1910." *Explorations in Economic History* 43, no. 3: 465-485.
- Hatton, Timothy J. 1997. "The Immigrant Assimilation Puzzle in Late Nineteenth-Century America." *Journal of Economic History* 57, no. 01: 34-62.
- Hatton, Timothy J. and Jeffrey G. Williamson. 1998. "The Age of Mass Migration: Causes and Economic Impact." *New York: Oxford University Press*.
- Higham, John. 1955. "Strangers in the Land: Patterns of American Nativism, 1860-1925." *Rutgers University Press*.
- King, Desmond. 2009. "Making Americans: Immigration, Race, and the Origins of the Diverse Democracy." *Harvard University Press*.
- Liebersohn, Stanley. 2000. "A Matter of Taste: How Names, Fashion, and Culture Change." *New Haven: Yale University Press*.
- Liebersohn, Stanley, and Mary C. Waters. 1988. "From Many Strands: Ethnic and Racial Groups in Contemporary America." Russell Sage Foundation.
- Liebersohn, Stanley, and Eleanor O. Bell. 1992. "Children's First Names: An Empirical Study of Social Taste." *American Journal of Sociology*: 511-554.
- Liebersohn, Stanley, and Kelly S. Mikelson. 1995. "Distinctive African American Names: An Experimental, Historical, and Linguistic Analysis of Innovation." *American Sociological Review*: 928-946.
- Limbaugh, Rush. 2014. "Hyphenated Americans and Immigration," *The Rush Limbaugh Show*.
- Lleras-Muney, Adriana, and Allison Shertzer. 2015. "Did the Americanization Movement Succeed? An Evaluation of the Effect of English-Only and Compulsory Schooling Laws on Immigrants." *American Economic Journal: Economic Policy* 7, no. 3: 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. 2005. "Intermarriage and the Economic Assimilation of Immigrants." *Journal of Labor Economics* 23, no. 1: 135-174.
- Moehling, Carolyn, and Anne Morrison Piehl. 2009. "Immigration, Crime, and Incarceration in Early Twentieth-Century America." *Demography* 46, no. 4: 739-763.
- Moehling, Carolyn M., and Anne Morrison Piehl. 2014. "Immigrant Assimilation into US Prisons, 1900–1930." *Journal of Population Economics* 27, no. 1: 173-200.
- Olivetti, Claudia, and Daniel Paserman. 2015. "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940." *American Economic Review*, 105(8): 2695-2724.
- Oreopoulos, Philip. 2011. "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes." *American Economic Journal: Economic Policy* 3, no. 4: 148-171.
- Pagnini, Deanna L., and S. Philip Morgan. 1990. "Intermarriage and Social Distance among US Immigrants at the Turn of the Century." *American Journal of Sociology*: 405-432.
- Rossi, Alice S. 1965. "Naming Children in Middle-Class Families." *American Sociological Review*: 499-513.
- Rubinstein, Yona, and Dror Brenner. 2014. "Pride and Prejudice: Using Ethnic-Sounding Names and Inter-Ethnic Marriages to Identify Labour Market Discrimination." *Review of Economic Studies* 81, no. 1: 389-425.
- Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. "Integrated Public Use Microdata Series: Version 5.0." [Machine-readable database]. *Minneapolis: University of Minnesota*.
- Smith, Malcolm T., and Donald M. MacRaild. 2009. "Paddy and Biddy 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: 595-608.
- Sue, Christina A., and Edward E. Telles. 2007. "Assimilation and Gender in Naming." *American Journal of Sociology* 112, no. 5: 1383-1415.
- US Census Bureau. Fourteenth Census of The United States, Population 1920. Retrieved from the National Bureau of Economics Research.
- Ward, Zachary. 2019. "The Low Return to English Fluency during the Age of Mass Migration," Working paper.
- Watkins, Susan Cotts, ed. 1994. "After Ellis Island: Newcomers and Natives in the 1910 Census." *New York: Russell Sage Foundation*.

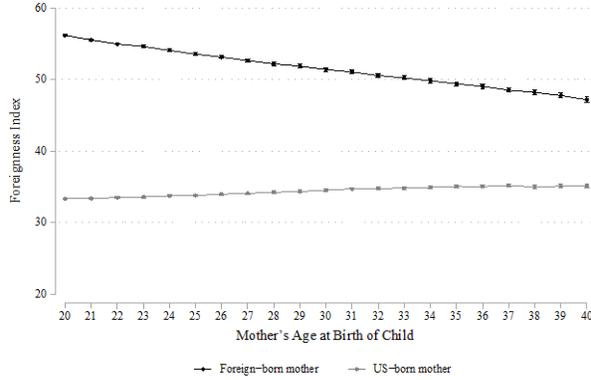
Watkins, Susan Cotts, and Andrew S. London. 1994 . “Personal Names and Cultural Change: A Study of the Naming Patterns of Italians and Jews in the United States in 1910.” *Social Science History*: 169-209.

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

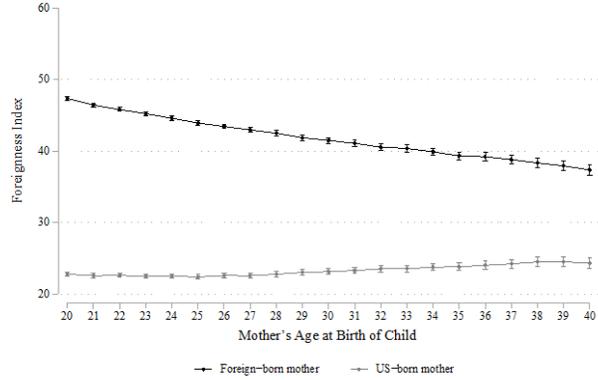
Zelinsky, Wilbur. 1970. “Cultural Variation in Personal Name Patterns in the Eastern United States.” *Annals of the Association of American Geographers* 60, no. 4: 743-769.

Figure 1: Immigrants selected less foreign names after spending time in US, past and present

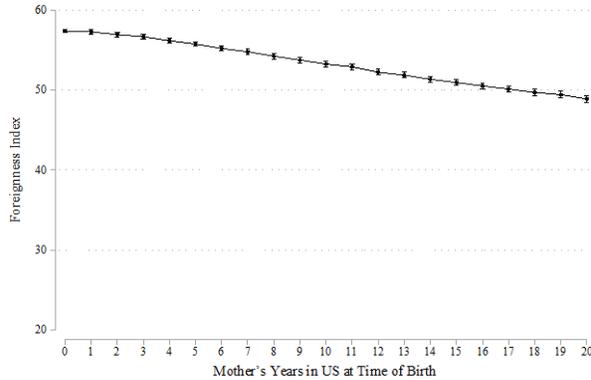
(A) Historical data (1920 Census)



(B) Modern data (California birth certificates)

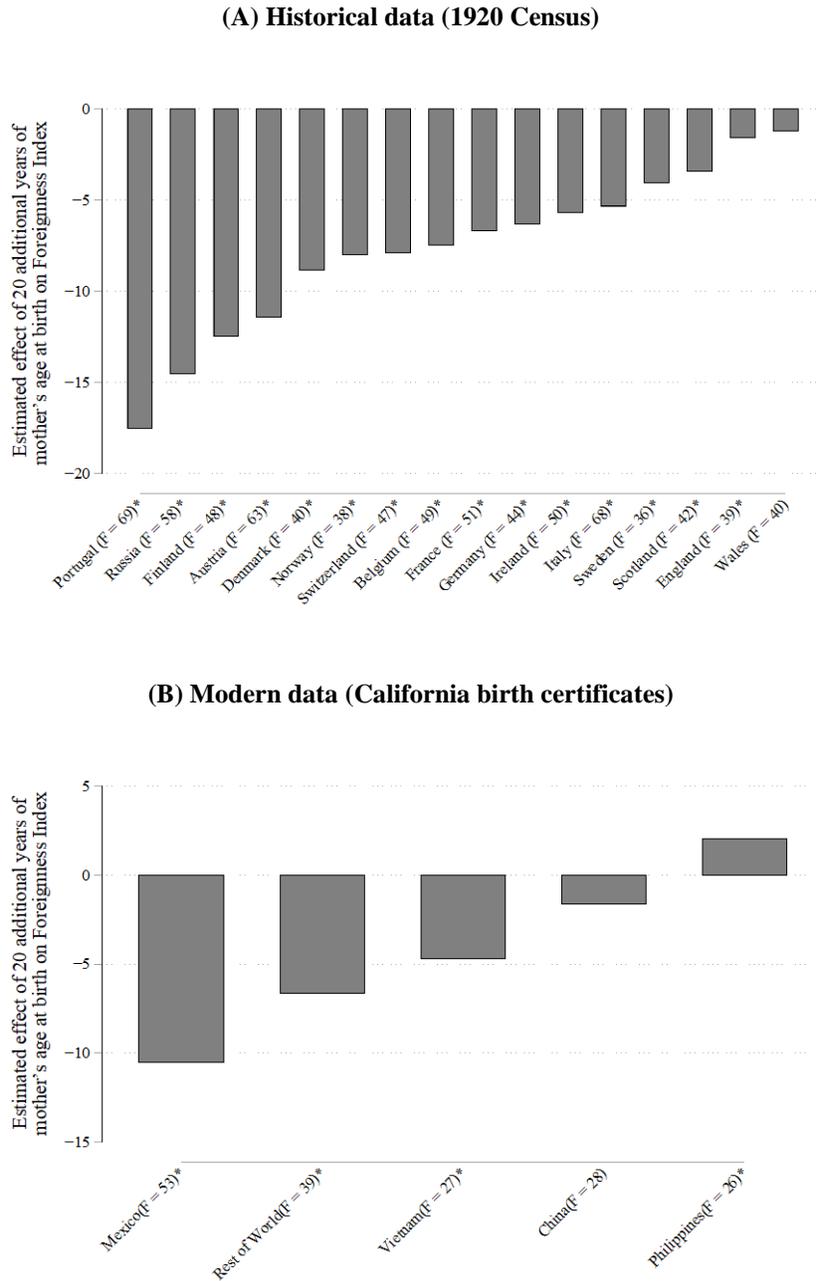


(C) Mother's years in US at time of child's birth (1920 Census)



Note: Panels (A) and (B) show the Foreignness Index of names given to children of native and foreign-born mothers by the mother's age when the child was born. These figures report the constant plus the estimated effect of the individual mother's age at birth indicators from equation (2). Panel (C) shows the foreignness index of names given to children of foreign-born mothers by the mother's years in the US at the time of the child's birth. Specifically, we report the constant plus the coefficients on indicators of mother's years in US when the child was born from equation (1). Regressions also include indicators for child's rank in the birth order, birth year in five-year bands, and mother fixed effects. The complete-count 1920 census data 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. $N(\text{foreign born}) = 4,219,213$; $N(\text{native born}) = 9,337,285$. The California birth certificates data includes all children born to a mother aged 20-40 in California from 1989-2015. $N(\text{native born}) = 2,957,731$; $N(\text{foreign born}) = 2,082,142$.

Figure 2: Effect of time in US on name foreignness by sending country, past and present



Note: In this figure we estimate a regression of the foreignness index of a child's name on a set of interactions between mother's country of birth and a linear term for mother's age at birth. We report 20 x the coefficient value of mother's age at birth for each country of birth. Regressions also include indicators for child's rank in the birth order, birth year in five-year bands, and mother fixed effects. Country labels also report the foreignness index of children born to mothers age 20 from each country. Panel A uses the complete-count 1920 census. Panel B uses California birth certificates from 1989-2015. Sample details in note to Figure 1. Countries marked with * have coefficients that are significantly different from zero at the 5 percent level.

Table 1: Effect of time in US on name foreignness by mother’s characteristics, past and present.**(A) Historical data (1920 Census)**

		Dependent Variable = Foreignness Index	
		Mother is....	
		Literate	Not literate
Italy	Mom age at birth	-0.318 (0.033)	-0.215 (0.034)
	<i>N</i>	405,349	307,335
Austria	Mom age at birth	-0.697 (0.039)	-0.520 (0.062)
	<i>N</i>	259,011	83,297
Germany	Mom age at birth	-0.307 (0.040)	-0.714 (0.245)
	<i>N</i>	257,488	5,812
Russia	Mom age at birth	-0.909 (0.037)	-0.693 (0.060)
	<i>N</i>	447,652	132,141
Finland	Mom age at birth	-0.574 (0.107)	-1.049 (0.386)
	<i>N</i>	48,264	4,037
Portugal	Mom age at birth	-0.633 (0.187)	-0.724 (0.181)
	<i>N</i>	15,845	16,442

(B) Modern data (California birth certificates)

Dependent Variable = Foreignness Index				
	Less than HS	HS grad	Some College	College plus
	(1)	(2)	(3)	(4)
Vietnam	-0.389 (0.319)	-0.393 (0.166)	-0.328 (0.176)	-0.245 (0.138)
<i>N</i>	21,296	44,331	32,784	46,190
Mexico	-0.544 (0.035)	-0.455 (0.045)	-0.488 (0.067)	-0.486 (0.106)
<i>N</i>	1,397,249	625,503	270,753	121,395
China	-0.768 (0.682)	-0.018 (0.331)	-0.094 (0.067)	-0.072 (0.168)
<i>N</i>	10,582	25,109	25,717	98,319

Note: This table reports coefficients from estimates of a modified version 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 include indicators for child’s rank in the birth order, birth year in five-year bands, and mother fixed effects, and are conducted separately by mother’s country of origin and literacy status. Panel (A) uses complete-count 1920 census data. Panel (B) uses California birth certificate data (1989-2015). Sample details in Note to Figure 1. The six countries included in Panel (A) exhibit assimilation in Figure 2 and have at least 2,000 mother-child observations in each category (literate/not).