Published in final edited form as:

Am Sociol Rev. 2016 February ; 81(1): 85–106. doi:10.1177/0003122415621910.

From Patrick to John F.: Ethnic Names and Occupational Success in the Last Era of Mass Migration

Joshua R. Goldstein and
University of California-Berkeley

Guy Stecklov
Hebrew University

Abstract

Taking advantage of historical census records that include full first and last names, we apply a new approach to measuring the effect of cultural assimilation on economic success for the children of the last great wave of immigrants to the United States. We created a quantitative index of ethnic distinctiveness of first names and show the consequences of ethnic-sounding names for the occupational achievement of the adult children of European migrants. We find a consistent tendency for the children of Irish, Italian, German, and Polish immigrants with more "American"-sounding names to have higher occupational achievement. About one-third of this effect appears to be due to social class differences in name-giving, and the remaining two-thirds to signaling effects of the names themselves. An exception is found for Russian, predominantly Jewish, immigrants, where we find a positive effect of ethnic naming on occupational achievement. The divergent effects of our new measure of cultural assimilation, sometimes hurting and sometimes helping, lend historical empirical support to more recent theories of the advantages of different paths to assimilation. The effects of first names are robust to controls for the ethnic recognizability of last names, suggesting that immigrants' success depended on being perceived as making an effort to assimilate rather than hiding one's origins.

Keywords

assimilation; migration; ethnicity; identity; names; naming

Assimilation and discrimination have long been at the center of discussion about immigration in the United States (Gordon 1961; Massey 1981; Waters and Jiménez 2005). Nearly a century ago, Theodore Roosevelt (1915) famously declared that an immigrant deserved equal treatment only if he were to become "American and nothing else." Today, Barack Obama (2013) calls for national unity in his second inaugural address, saying that being American should not depend on "the origins of our names."

Corresponding Author: Guy Stecklov, Department of Sociology and Anthropology, Hebrew University, Jerusalem, Israel, stecklov@huji.ac.il.

Replication: A data file (STATA v.13) is provided online along with code to enable replication of the four main tables in the paper. A complete version of the dataset can be obtained from the IPUMS website: https://usa.ipums.org/usa/.

In this article, we take on the empirical question of how relevant cultural assimilation was for economic success during the last great wave of immigration at the beginning of the last century. Taking advantage of the millions of newly available historical census records (Ruggles et al. 2010), we analyze the choice of first names to investigate the causes and consequences of immigrant assimilation at the individual level. We ask whether immigrants who assimilated were better able to climb up the economic ladder than those who retained their ethnic identity. An analysis of last names allows us to explore whether effects of the Americanization of names were due to the benefits of hiding one's ethnic origins or to a perceived effort to "act American."

Despite the extensive discussion of assimilation, surprisingly little empirical research examines the consequences of cultural assimilation for economic achievement. One reason for this conspicuous absence is that culture is difficult to measure. In small or specialized surveys, detailed measures of cultural identity can be explored in great detail (Portes and Rumbaut 2012; Waters 1994), but it is quite difficult to measure how "American" an immigrant or an immigrant's children "feel" in the large-scale surveys used to measure economic attainment.

The second challenge is the difficulty of disentangling measures of cultural assimilation from those of socioeconomic success. For example, use of intermarriage as a measure of cultural assimilation is widespread, but individuals who intermarry are different in consequential ways from those who do not. In an analysis of socioeconomic success, it would be difficult to identify the net effect of intermarriage.

Using first names given to children as an index of cultural assimilation offers potential solutions to both of these problems. In terms of measurement, first names represent an opportunity to get at how parents felt about the cultural identity they wished for their children. Furthermore, thanks to the public availability of historical census microdata, this volitional measure is available in surveys of enormous scale, including the decennial census.

Better causal identification is also possible through analysis of names. Because names are (largely) constant over a lifetime, the control of socioeconomic gradients in name-giving allows us to reveal something closer to the causal effect of cultural assimilation indicated by the identity aspect of a name. This removes the effects of other influences, like parental status, that correlate with children's economic success.

First-name choice is a valuable indicator of assimilation because it can be used to measure the loss of distinctiveness between populations and integration into the mainstream (Alba and Nee 2005). In choosing names for their children, parents face an important trade-off between the desire to transmit their traditions to the next generation and providing children with identities that will maximize their chances of success in their new home. The freedom of choice of first names makes them a potentially powerful indicator of this subjective aspect of assimilation.[1] As Sue and Telles (2007:1385) point out, first names provide a "window

[1]In this respect, first-name choice differs from other measures of assimilation that are either highly constrained (e.g., occupation or neighborhood) or serve some other instrumental function (e.g., language acquisition). First names are distinguished from last names, which may also contain important information on ethnic identity but are often inherited.

into parental visions of the ethnic identity of their children" and act as signals of this identity to outsiders.

A large body of literature on name-giving emphasizes the sociological importance of names. Names are important symbols of group affiliation, with connotations of class and ethnicity (Elchardus and Siongers 2011; Rossi 1965; Sue and Telles 2007). The choice of names has been used to explore diverse forms of cultural expression (Lieberson and Mikelson 1995; Obukhova, Zuckerman, and Zhang 2014), and cycles in name popularity have also attracted study (Berger and Le Mens 2009; Li 2012; Lieberson 2010). Immigrant groups have received attention for the period we are studying in the United States (Moser 2012; Watkins and London 1994), as well as historical and contemporary France (Besnard and Desplanques 2001; Coulmont 2011) and Germany (Gerhards and Hackenbroch 2000; Gerhards and Hans 2009).

In economics, there is an emerging literature on the consequences of first names, particularly in the labor market. A number of U.S. studies have examined "black"-sounding names. Audit studies, using the ruse of sending fake résumés to prospective employers, show that "white"-sounding names result in more interview requests than "black"-sounding names (Bertrand and Mullainathan 2004)—with similar findings for disadvantaged ethnic groups in Sweden (Carlsson 2010). Such audit studies, of course, are unable to shed light on historical patterns of discrimination. Using observational data in California Fryer and Levitt (2004) analyze the causes and consequences of distinctively black names.[2] Other recent research on the signaling consequences of names in labor markets includes analysis of name changes by migrants in Sweden (Arai and Thoursie 2009).

Observational studies like ours, which try to look at the causes and consequences of ethnic naming, must contend with the fact that name choice is part of a diverse set of social and cultural behaviors and intentions, many of which are unobserved. We find it useful therefore to consider names as an indirect, or "proxy," measure of this broader set of intentions. This approach allows us to go beyond name choice itself, addressing fundamental theories and processes of immigrant cultural assimilation.

First names can be seen as cultural markers that parents use to express their own desires and traditions. These cultural markers are then perceived by friends, family, teachers, and others, including potential employers. First names can trigger discrimination depending on the extent to which they are recognized as belonging to one group or another. From this perspective, a name is "distinctive" if it is disproportionately held by members of a particular group. For the purposes of this study, we operationalize groups as the country of birth of one's father. For example, a name is distinctively Irish if it is disproportionately held by the children of fathers of Irish origin. A name is less distinctive, and more "American," to the extent it is held disproportionately by those with native-born fathers. Conversely, a name is more distinctive if it is held disproportionately by individuals with fathers from a particular ethnic background. Walter, for example, is scored as a German name, not because

[2]An advantage of the current study over Fryer and Levitt (2004) is that the census gives economic and social characteristics at the individual level, whereas Fryer and Levitt's use of birth certificates relies on area-level estimates of economic covariates.

it necessarily sounds German, but because it is disproportionately found among people with German-born fathers. Our approach emulates the cognitive process whereby people form associations between names (words) and certain social connotations through repeated exposure (Reber 1996; Seger 1994).

In this article, we develop quantitative measure of ethnic distinctiveness applicable to the names of children of immigrants in the United States in the first part of the twentieth century. We limit our consideration to male names because, in the era under study, occupational attainment—our measure of economic success—is more readily available for men. Our measure is based on the statistical frequency of names by parent's place of birth and thus does not require independent raters to classify names by ethnic origins (Gerhards and Hans 2009). This also means our approach is potentially replicable with large modern datasets, using numeric codes to replace actual first names to protect data confidentiality.

Using this new measure, we are able to answer the question of whether cultural assimilation had a causal effect on economic success. We interpret the causal effect of ethnic names as indexing the role of cultural assimilation on labor market outcomes. This may be due to the signaling effect of names in job markets, schools, and other contexts and thus may be attributed to the names themselves. This may also be due to other dimensions of cultural assimilation, such as language, identity, and values, with which names are correlated. To some extent, these two interpretations are not separable, because people who interpret the signaling effect of names (i.e., individuals who discriminate) are not doing so because of the name per se, but rather because of a broader set of characteristics that the name represents. This set of characteristics is part of the multidimensional concept of cultural assimilation.

Our main goal here is to add new empirical evidence documenting the link between cultural assimilation and economic achievement, but our work also makes four theoretical contributions to the literature on immigrant assimilation. First, we connect theories of population-level immigrant assimilation with socially constructed aspects of ethnic identity. Our study adds nuance and heterogeneity (Brubaker 2001) to the classical studies of assimilation of white ethnic immigrant groups (Gordon 1961, 1964; Warner and Srole 1945).[3] We demonstrate empirically that overall cultural assimilation (the Americanization of names) was indeed linked to structural assimilation (in this case measured by occupation), providing support for assimilation hypotheses. However, we also find considerable variation between, and more importantly within, groups, depending on individuals' choices (Hout and Goldstein 1994; Waters 1990). Individual agency—how American parents wanted their children to be—varied a great deal between parents, and these choices about how much to assimilate had long-term consequences.

Second, our study defines ethnic distinctiveness in a continuous and dynamic way, which makes assimilation and "Americanization" a process aimed at a moving target. Building on

[3]The current article follows in the spirit of Warner and Srole's (1945) quantitative study of immigrant community assimilation in a New England town, adding a direct measure of cultural assimilation and economic outcomes. As the 1940 census microdata become available, it will be possible to incorporate detailed small area community measures and, using our new metric of acculturation based on names, to gain further understanding of the interplay between geographic, economic, and cultural paths of integration and assimilation.

Alba and Nee's (2005) idea of the "mainstream" as the portion of society for which ethnic distinctions are not important, we operationalize non-ethnic names as names held by individuals born in the United States, regardless of ethnic origin. As more immigrants arrive over time, the ethnic composition of the native born changes, and so does the definition of "non-ethnic" naming. This dynamic vision of the mainstream allows us to account for interesting phenomena, like a formerly "foreign" name becoming "American." To take a recent example, the name Michael, which was extremely Irish in the 1930s, was the second most popular name among newborn boys in the United States in the past decade (2000 to 2009).[4] The reverse also occurs, with non-English groups adopting English names. For example, in the 1930s, Russian-Jewish immigrants disproportionately gave their children the English names Harry, Max, and Morris.

Third, our study distinguishes cultural assimilation as a choice from cultural assimilation as an unintended side-effect. For example, Alba and Nee (2005) emphasize that groups acculturate (e.g., move out of a neighborhood densely populated with co-ethnics) to achieve some structural goal (e.g., send their kids to better schools). Here, however, the acculturation of names and other visible cultural characteristics help achieve the structural goal of fitting into U.S. society and increasing economic mobility. We view names as a proxy for intent to assimilate, and not as a by-product of other forms of adaptation.

Finally, our analysis specifies an important conceptual distinction in the discrimination process relevant to children of immigrants who carry ethnic markers. The most direct disadvantage of carrying an ethnic name is that it makes a person ethnically recognizable; indirectly, an ethnic name shows that a person's orientation is not toward the mainstream. The distinction here is between signals that allow classification of a person's origin versus signals that allow classification of a person's orientation. We can reveal the mechanism at play by combining an analysis of first names with an analysis of recognizably ethnic last names. If the advantages of being John versus Patrick are found only among people with last names that are not recognizably Irish, then this is evidence of discrimination against origins. But if we find the same effects for the Kennedys and the O'Donnells—last names that are recognizably Irish—then this suggests "orientation"-based discrimination rather than "origin"-based.

Our article provides one of the first empirical analyses of the relationship between ethnic identity and economic success. The literature on assimilation (exemplified by Gordon [1964]) has dissected different forms of assimilation into distinct dimensions (acculturation, structural assimilation, and identity assimilation) arguing for the causal necessity of one kind of assimilation process over another. However, little work investigates the consequences of assimilation on an individual level. Does the choice to assimilate make a difference for individuals, or are the choices parents make entirely bounded by existing discrimination? Is there something indelible about ethnic origins that makes it difficult for individuals to have any influence on how they are perceived, or can parents' decisions, evident in our analysis of

[4]The name Michael had an ethnic name index value of .97 in the 1930s, which means it was found among the children of Irish fathers but almost entirely absent among the native population.

names, contribute to the different pathways children take in life? These are the kind of questions that an analysis of naming and economic achievement can help answer.

DATA

Our analysis uses the publicly available IPUMS micro-samples of the U.S. censuses of 1880 to 1930. These samples, no longer covered by confidentiality safeguards, include full transcriptions of first and last names. In addition to names, the census records also include individual-level information on age, birthplace, father's and mother's birthplace, individual occupation, and basic individual demographics, such as age and sex.

Our primary dataset is the census of 1930, for which a 5 percent sample is available. We limit our analysis to native-born males age 20 to 59 in 1930, whose fathers were either native-born or born in Ireland, Italy, Germany, Poland, or Russia, the five largest immigrant origins for these cohorts. This gives us a total sample size of about 1.3 million native-born men, with sample sizes, according to father's birthplace, of 33,037 for Ireland, 14,516 for Italy, 73,644 for Germany, 12,683 for Poland, and 11,795 for Russia.

To measure occupational attainment, we use the synthetic IPUMS variable, OCCSCORE. OCCSCORE is based on the average income observed in the 1950 census, in hundreds of dollars, and adjusted to account for changing occupational categories over time (Ruggles 2010). Advantages of OCCSCORE include comparability over time, ease of interpretation, and ranking of occupations along a single dimension. A disadvantage, shared with all occupational measures, is that it cannot detect variation within occupations. We thus use OCCSCORE not as a proxy for direct measurement of income at the individual level, but more broadly as a measure of occupational attainment, in line with the long tradition of research on occupational status (Duncan 1961; Stevens and Featherman 1981).

To prepare the transcribed census names for analysis, we performed the following preprocessing steps. First, we took only first names and dropped middle names; second, we dropped all cases in which first names were single characters; third, we dropped all cases in which the first name was unique in the population; and fourth, we dropped all cases with first names that could not be read by transcribers. These procedures led to a loss of 2 percent of the cases.

METHODS

Measurement of Ethnic Name Index (ENI)

We compute the ethnic distinctiveness of names based on the relative frequency of each name in a given ethnic immigrant group compared to the frequency in the native population. The ethnic group is defined by father's birthplace. We call this measure the ethnic name index (ENI). For name *j* from the perspective of ethnic group *k,* ENI is defined as

$$ENI\left(name_j, ethnicity_k\right) = \frac{p\left(name_j \middle| ethnicity_k\right)}{p\left(name_j \middle| ethnicity_k\right) + p\left(name_j \middle| native\right)}, \quad (1)$$

where p is the fraction holding given name j among people in ethnicity k. For example, the name Patrick is held by about 5 percent of native-born sons of Irish-born fathers and about .05 percent of native-born sons of native-born fathers, which gives a value of ENI(Patrick,Irish) = .05/(.05 + .0005) = .99.

This formulation, adapted from the measure introduced by Fryer and Levitt (2004) to study distinctively black names, ranges from 0 to 1. A value of zero for the index means the name is absent from children of parents in group j but present in the native population. A value of one means the name is exclusive to the ethnic group and absent in the native population. We calculate the ENI separately for each ethnicity. This allows the valence of native names, like John, to vary by group. As Figure 1 shows, John represents more Americanization for the children of Russian or Italian immigrants than it does for Irish or Polish immigrants.

ENI measures differences within ethnic groups, not between them, because each group arrives in the United States with its own distribution of names that varies in its dissimilarity with the native born. The history of immigration also influences between-group differences in ENI, because we classify all native-born fathers as "native," without distinguishing more distant ethnic origins. ENI is thus a dynamic measure that changes over time with the composition of the native born.[5] Figure 2 presents kernel density plots, which facilitate visualization of the probability distribution of a continuous variable for each ethnic group's ENI. Broadly speaking, groups that have been in the United States the longest appear to have the lowest ENI.

Models Relating Names and Occupational Achievement

Occupational attainment of the children of immigrants is a complex process with many causal influences. We use a diagram to describe a simplified model of how the economic status and cultural assimilation of immigrant parents influences their children's occupation, after removing contextual factors such as birth cohort, geography, and rural/urban residence, which potentially influence both name choice and occupational attainment.

Figure 3 shows the model with occupational attainment as the outcome of interest. One influence on occupation is the status of one's family of origin, intergenerationally transmitted through children's education, wealth, and aspirations, and signaled in part by name choice. Another influence is a family's degree of cultural assimilation. Among the mechanisms that influence occupational attainment are ethnic signaling, the availability of ethnic networks, and the influence of a child's occupational aspirations. Ethnic assimilation can be signaled through ethnic naming. But it is also important to recognize that ethnic assimilation can be signaled by other characteristics, such as language ability, accent, mode of dress and appearance, and other ethnic-cultural markers. In Figure 3, cultural assimilation thus influences occupation via signaling through ethnic naming and other mechanisms.

The statistical models we estimate allow us to see several different aspects of the relationship between assimilation and occupation. We begin with some descriptive evidence

[5]ENI is potentially unstable when names are rare. To check the robustness of our findings, we reanalyzed the data restricting ourselves to names that are represented more than 10 times (rather than more than once, as in the analysis we present); we found no substantive differences in the results.

on the link between ethnic naming and occupation, identified by a bold arrow between naming and occupation in Figure 3. Absent our controls, this bold arrow identifies the unconditional effect of ethnic naming, which includes how naming makes one recognizably ethnic, how naming signals one's orientation toward the mainstream, and other determinants of occupation that are correlated with naming.

To analyze the *consequences* of ethnic names, we first estimate two multivariate models, separately by ethnic group, in which we relate the occupational earnings of adult men to their own first names. The first model aims to estimate the overall effect of first names on occupational attainment, adjusting for observable demographic characteristics:

$$OCCSCORE_i = \alpha + \beta ENI_i + \gamma X_i + e, \quad (2)$$

where $OCCSCORE_i$ is the occupational income of person I; X_i is a vector of control variables noted earlier, including age, urban/rural residence, region of residence, and nativity of mother; and ENI_i is the ethnic distinctiveness of person i's name. Controlling for age in our cross-sectional data eliminates this factor as a possible confounder, since both ethnic naming and occupational attainment may be correlated across cohorts. We adjust for mother's nativity because it may also be correlated with name choice and occupation. We adjust for urban residence because it may be an important determinant of occupational attainment. We adjust for the effect of region of residence by including a series of eight dummy variables in the model.[6] With these additional predictors, the coefficient β in Model 2 tells us about the extent to which assimilation—depending on the family of origin's socioeconomic position and degree of cultural assimilation—goes hand-in-hand with occupational attainment.[7]

The results of Equation 2 can be interpreted as the joint effect of the social gradient in name-giving and the causal effect of cultural assimilation as indexed by carrying a particular name. For instance, if Patricks are in lower-earning occupations, this may be partly because they are perceived as Irish, and partly because children coming from families who named their children Patrick are disadvantaged. To disentangle these effects, we estimate a second model that controls for the correlation between name-giving and class background. This step helps rule out the alternative explanation that occupational attainment is driven primarily by parent's class, and not name choice and other cultural factors that are proxied by name choice. Ideally we could do this using linked pairs of father–sons—a strategy we undertake on a limited subsample based on IPUMS linked-sample files (see Part A of the Appendix). However, father's occupation is not observed for the great majority of our 1930 sample of adults. As a result, we use a name-level measure of father's occupation, taking the average OCCSCORE of fathers by name of son in the 1900 census, which we call NAMEFOCC. Although this measure is approximate for the individual, it is unbiased at the level of first names. Unlike the usual errors-in-variable model, which leads to a downward bias of

[6]The regional categories include New England (reference category), Middle Atlantic, East North-Central, West North-Central, South Atlantic, East South-Central, West South-Central, Mountain, and Pacific.

[7]Some households include multiple members. Testing showed that correcting standard errors for within household correlations did not alter our findings.

regression estimates, the use of group-level means, as in father's occscore by firstname of son, has been shown to give unbiased estimates (Berkson 1950; Durbin 1954). The resulting model is then

$$OCCSCORE_i = \alpha + \beta ENI_i + \gamma NAMEFOCC_i + \delta X_i + e \quad (3)$$

This final set of analyses enables us to estimate the effect of ethnic naming, controlling for family's class background (and other contextual factors) as well as observable individual characteristics. In this formulation, because we do not observe other features of cultural assimilation that influence occupational attainment, ethnic naming functions as an indirect measure of cultural assimilation. The counterfactual is that a person would not only have had a different name, but would have come from the kind of family that would have given this name. Thus, while we might use the shorthand of referring to the causal effect of naming, our approach goes beyond estimating the causal effect of the name itself. Instead, we capture the causal effect of the broader sense of immigrant cultural assimilation that is indexed by the ethnic distinctiveness in names.

RESULTS

Before analyzing multivariate models, we begin with a graphic illustration of the relationship between ethnic distinctiveness of first names and occupational attainment as measured by OCCSCORE. The five panels of Figure 4 show the relationship between the ENI and OCCSCORE for the 20 most popular names given to the sons of immigrants from each country.[8] For Ireland, Germany, Italy, and Poland, the most "ethnic" names, such as Patrick for the Irish, Herman for the Germans, and Stanley for the Polish, are all among the lowest earning. More Anglo-American names, such as Charles, William, and Edward, are all among the highest earning. For sons whose fathers were born in Italy, the pattern is less clear for the top names. For sons with Russian-born fathers, those with "American"-sounding names, such as William and John, earn less than sons with ethnic biblical names, such as Samuel and Joseph, and less than those with non-biblical but high-ENI names, such as Harry and Louis.[9] (The lines shown in the figures are unweighted, bivariate regression lines.)

The bivariate plots shown in Figure 4 suggest a strong association between the degree of ethnic distinctiveness of names and occupational earnings. They are also reminders of the substantial differences in the average occupational income and the variable gap between richest and poorest across groups. We now adjust for the other factors mentioned earlier, such as rural or urban residence, region of residence, age, and mother's nativity, all of which might be jointly correlated with both naming and occupational earnings.[10] Table 1 shows the

[8]The top-20 names for each group include 74 percent of the Irish sample, 62 percent of the Italian sample, 64 percent of the German sample, 60 percent of the Polish sample, and 55 percent of the Russian sample.

[9]The tendency of nicknames like Fred, Joe, and Tony to have lower earnings is likely also due to class differences in the reporting of names in the census: people of higher status likely report Frederick, Joseph, or Anthony, and those of lower status overreport Fred, Joe, and Tony. Many nicknames also have an ethnic valence, as indicated here by the ENI scores of Joe and Tony. Tony Bennett may have changed his name from "Anthony Dominick Benedetto," but his use of "Tony" retained his ethnic distinctiveness.

[10]We restrict age to control for temporal trends in naming, an approach we implement in our regression analysis by controlling for age.

results of multivariate analyses that include all names with more than one occurrence (not just the most popular).

Table 1 shows regression coefficients of ENI (and control variables) on occupational earnings. For the Irish, for example, being named Patrick (ENI = .99) rather than William (ENI = .60) is associated with occupational earnings 29 dollars less per year, or about 1.5 percent of annual earnings.[11] The results show that the ethnic distinctiveness of names has a negative effect on occupational earnings for sons of Irish, Italian, German, and Polish fathers. For Russians, however, the effect is positive.

It is important to adjust the estimate for urban residence, which is strongly tied to occupational category and our measure of occupational earnings. Indeed, ENI tends to be higher for urban than for rural residents, accordingly, controlling for urban status reduces the apparent effect of ENI on occupational earnings.

As discussed above, the estimates in Table 1 should not be interpreted as the causal effect of ethnic naming, because we know that there are class patterns in ethnic naming and that class persists across generations (Elchardus and Siongers 2011; Lieberson 2010). To adjust for this effect, we estimated the average occupational earnings of fathers, by name of their sons. We did this for sons age 10 and younger in the 1900 census, so that gradients in names would be most relevant for the grown men age 30 to 40 we observe in 1930.[12] Table 2 shows results of our second regression, based on Model 3, providing better identification of causality, in the sense that we have differenced out the effect of family of origin.[13]

As Table 2 shows, the effects of ENI, once we adjusted for average occupation of fathers by name of sons, show the same general pattern we saw in our bivariate analysis of top names in Figure 3 and in the multivariate analysis shown in Table 1. The Irish, Italians, Germans, and Poles all show negative effects of more ethnic names on occupational achievement, whereas the Russians show a positive effect.[14] Overall, adjusting for the social gradient of naming reduces the effect of ENI by roughly a third. For example, the Irish coefficient falls from about −.8 to about −.6, the Italian coefficient from −.6 to −.5, the German coefficient from −1.0 to −.5, and the Russian coefficient from +2.1 to +1.5.

Accordingly, we interpret the overall pattern to be robust to the inclusion of family background: two children who had the same socioeconomic status as measured here would earn more or less depending on the level of parental cultural assimilation transmitted to them by their parents and identified with their given names.[15] We find further support for this

[11] $-.755 \times (\text{eni}_{\text{Patrick}} - \text{eni}_{\text{William}}) \times 100 = 29$ dollars less per year.

[12] Table B1 in the Appendix shows results of a simple regression of ENI on father's occupation.

[13] This controls for differences between families but does not account for potential differences within families. Our analysis cannot exclude the possibility that some other influence, such as parental ambition, drives both naming and outcomes *within* families. For example, parents might have higher ambitions for one son and thus give him a more "American-sounding" name than his brother. But this effect would have to be wholly within families, because any systematic differences in ambition and naming between families would be controlled for using NAMEFOCC.

[14] Although the names William and John appear to be influential for Russians in Figure 4, separate models estimated without these names also produced very positive and significant coefficients for the effect of ENI on occupational income.

[15] We also conducted a falsification test using infant's month of birth; this was suggested by an anonymous reviewer to help reassure against spurious causality. We estimated models using the 1900 IPUMS sample to test whether month of birth predicted ENI and found no association among any of our five ethnic groups.

result in a smaller analysis of German father–son pairs that could be linked across the two census files (see Part A of the Appendix).

The study of the relationship between naming and occupation can also be carried out at a finer geographic resolution. Doing so would confirm that our national results are not driven by a spurious small-scale correlation between naming patterns and economic structures. To test this, we fit county-level fixed-effect models (see Table 3). County-level fixed effects enable us to model the relationship between ENI and occupational income for each group without requiring us to assume there are no unmeasured (fixed) differences between counties that might bias our estimates. These results, which are very similar to our national estimates, support our interpretation of assimilation being at the individual (family) level and not a spurious effect of small area variation.

Does the effect we find for names vary by the composition of the local population? For example, ethnic enclaves may protect individuals with ethnic names (Portes 1987; Wilson and Portes 1980). We tried to evaluate this hypothesis by interacting the proportion of the county from each ethnic background with the individual ENI scores. However, these analyses were very sensitive to different specifications of the density variables and not very robust—unsurprising given the small sample sizes available within many counties and the highly skewed distribution of ethnic composition. However, analysis of the 100 percent data for 1940 is a good topic for future research on the role of ethnic concentration.

Name Changes

First names recorded in the census are not necessarily identical to those given at birth. We know that immigrants frequently changed their names. The second generation, particularly those with very ethnic names, may have undergone its own process of name-changing and Americanization. To what extent might cultural assimilation, proxied by names, be a consequence rather than a cause of economic success? There are several reasons to believe this kind of reverse causality is not driving our main results. First, although name changes figure frequently in the history of Jewish immigrants, the reverse pattern we find for sons of Russian immigrants, in which less ethnic names have lower achievement, shows that Americanization among the high achieving is not a universal empirical pattern across all groups.

Second, although we cannot determine all name changes, it is possible to identify nicknames in the census by using a standard dictionary of nicknames,[16] mapping, for example, Jim and Jimmie to James. Nearly 8 percent of respondents in our sample gave names that can be coded as nicknames. These nicknames tend to sound more “American” than their full-name forms. (For example, among sons with Irish-born fathers, the name Jimmie has an ENI of .16, whereas James has an ENI of .74.) Despite the fact that these nicknames are more “American,” the positive effect of cultural assimilation appears to be overwhelmed by the status effect of listing a nickname rather than a full name. People who identified with a nickname on the census reported occupations associated with income levels that were lower

[16]Through personal correspondence, we obtained a dictionary file of nicknames from the IPUMS project.

by some 10 percent (details available by request). Americanization via nicknames thus does not appear to be driving our main results.

Finally, although name changes may in some cases be a consequence of economic success, this does not preclude our main interpretation: the name change itself contributed to further success. Indeed, individual choices to Americanize names may have been motivated by the belief that sending a less ethnic signal would be of value.

Distinguishing Mechanisms Using Last Names

Our analysis shows that for four of our ethnic groups—everyone except the Russians—greater cultural assimilation is consistent with higher achievement. However, last names, which are also ethnic markers, may in some cases be equally or more relevant signals of ethnic identity. Yet, in contrast to first names, last names are generally not chosen but are passed on from one generation to the next (Clark 2014; Clark and Cummins 2015). This means that both first and last names can be public expressions of ethnic identity, yet first names also reflect parental intentions. A combination of information on the ethnic salience of *both* first and last names enables us to distinguish between the effects of discrimination and the effects of individual and parental intentions that underlie assimilation.

Table 4 presents a series of models that estimate the effects of ENI while controlling for last names in various ways. We are interested in whether controlling for the ethnic valence of last names removes or lessens the impact we observe for the effect of ethnic first names on occupational income. If the ENI of the first name no longer matters for those with an ethnic last name, then we interpret this as supporting the idea that the effect of names is to signal ethnic *origins*, and that once these origins are clear, the first name has little impact. On the other hand, if the effect of ENI of the first name persists even when combined with an ethnic last name, then we interpret the effect of names as signaling the ethnic *orientation* of individuals.

The difference between “origins” and “orientation” revealing names is important. From the point of view of discrimination, this difference echoes that between ascribed and achieved status so central in the sociological literature. Here, the ascribed status is the relatively fixed nature of “origins” and the achieved status is the assimilative ambition or “orientation.” There is also a difference with regard to the agency of parents concerning first names. If hiding origins is the key to success, then the choice of first names is insufficient, and last names would also need to be changed. However, if displaying orientation is important, then parents can influence perceptions of their children by their choice of names, giving them a degree of agency over their children’s success.

This analysis requires us to define the ethnic content of last names. Rather than using a statistical approach like the ENI,[17] we build on typical ethnic markers associated with last names to create a dummy indicator for recognizably ethnic last names within each group (a list of the top-10 ethnic and non-ethnic last names for each group is shown in Table S1 in the

[17]Constructing an ENI index for last names proves problematic because the degree of diversity of last names (and their spelling) is so much greater than for first names. Nonetheless, when such an index is created we obtain similar results to those shown in Table 4.

online supplement [http://asr.sagepub.com/supplemental]).[18] The first column in Table 4 replicates the earlier result of ENI, controlling for father's occupational income but without accounting for last names (repeated from Table 2); the second column includes a dummy variable to control for whether the last name is recognizably ethnic. In the third column, we estimate a model including only the subsample of people from each ethnic group with last names classified as recognizably ethnic. Finally, the last column presents last name fixed-effects models. This specification removes all factors that are similar for people holding the same last name and estimates the effect of ENI based on the association of ENI and occupational earnings among people sharing the same last name.

Our first set of results, shown just to the right of the main model, shows that inclusion of a simple control for whether a last name is recognizably ethnic does not alter the main results. Thus, our finding that the effect of ethnic identification of first names affects occupational income remains unchanged after adjusting for the ethnic content signaled by last names. The next column provides further support for this claim, showing the impact of ENI on the subsample that has recognizably ethnic last names. Despite large reductions in sample size (and increases in standard errors) for most groups, the main ENI effects are substantively very similar. The most obvious changes are a substantial decline in the estimated effect of ENI among Germans and Poles and a large increase for Russian Jews. Finally, the fixed-effect results provide additional indication of the robustness of our results. Again, these estimates are substantively very similar to those from the main model.

The evidence in Table 4 shows that including the last name ethnicity indicator has little substantive impact, regardless of modeling approach, on the effect of ENI for any of the five groups in our analysis. It is thus unlikely that overt discrimination based on ethnic recognition is driving our main ENI findings. Instead, the signaling effect of names appears to result from their expression of cultural orientation, a kind of "acting mainstream," even by people who are recognizably of foreign origin.

DISCUSSION

For most major groups in the last era of immigration, carrying an ethnically distinct name was associated with lower occupational attainment. Native-born sons of Irish, Italian, German, and Polish immigrant fathers who were given very ethnic names ended up in occupations that earned, on average, $50 to $100 less per year than sons who were given very "American" names. This represented 2 to 5 percent of annual earnings. To put this effect in perspective, it would take about five to ten years of labor market experience to produce an effect of similar magnitude. We estimate that about one-third of the effect of

[18]For the Irish, we include names beginning with an "O" or one of several popular Irish last names, such as Kelly or Walsh. For Italians, we use last names that end in vowels. For Germans, we use last names that have typical consonant (e.g., *sch*) or vowel (e.g., *ue* or *ie*) combinations. Polish ethnic last names are identified based on suffixes such as *ski* or beginning with "Z" or "J." Finally, Russian last names we define as ethnic are actually Jewish last names and are based on typical Jewish strings, including *gold* and *stein,* or traditional, biblical last names, such as Cohen and Levy. For all groups, we aim to capture last names that mark those groups in their respective ethnic category, except for Russians where the marker is for Jewish. Our ability to identify these last names is not perfect and varies considerably by immigrant group, but we adopt a conservative approach in that we aim to reduce the rate of false negative attribution. According to our categorization, 88 percent of Italian last names were recognizably ethnic, but we defined as recognizably ethnic only 41 percent of German names, 33 percent of Polish names, and 8 percent of Irish names. Among Russians, only 26 percent of last names were identified as typically Jewish.

ethnic naming was due to the correlation of naming with class background. We interpret the remaining effect as being due to the level of cultural assimilation associated with and the signaling effect of ethnic naming.

That "American"-sounding names went hand-in-hand with occupational advances for immigrants is consistent with the broad narrative of assimilation advanced by historians and sociologists (Perlmann and Waldinger 1997; Watkins and London 1994). The measures and methods presented here provide new empirical support for a key claim in the existing literature on the role of assimilation in immigrant economic success. Furthermore, because our measures involve choice at the individual level, they allow us to directly capture the role of individual decisions within the broader social process of group assimilation, as shown in the pioneering group-level work by Warner and Srole (1945).

Our work also produces several new findings. First, we find that for many groups there appears to be a causal effect of having been assigned an ethnic name—an effect due to both the name and other cultural signals correlated with having been given this name. This effect has eluded earlier observational studies, such as Fryer and Levitt's (2004) research on black names, and the vast number of studies using other measures of assimilation. Our control for family characteristics is a powerful control for the possible influences on name-giving, strengthening the causal interpretation.

Second, further analyses taking into account last names allowed us to gain insight into the mechanism behind the effects of cultural assimilation on achievement. Assimilation may operate by hiding foreign origins or by displaying an American orientation, even for people who have recognizably foreign origins. Our finding that American-sounding first names were an advantage even for people with recognizably ethnic last names suggests that the signal being sent was one of mainstream orientation rather than origin. Our study supports the idea that U.S. society shared Roosevelt's perspective: only immigrants who abandoned their foreign affiliations deserved "exact equality."

This distinction between signaling "origin" and "orientation" is useful for the study of other forms of group differentiation and discrimination. For example, one interpretation of the effect of distinctive black names is that they reveal "blackness" that would otherwise remain hidden. This might play out at the job search stage in which only paper applications are being considered. But another interpretation is that the disadvantage of distinctive naming is not so much in revealing origins (in this case, skin color) but in revealing orientation (in this case, a cultural orientation away from mainstream white society). With this latter interpretation, the effect of distinctive names would persist even in situations, like job interviews, when skin color is known.

Third, our finding that for some groups—notably the Russians, who were primarily of Jewish origin—having an ethnic name has a positive effect on occupational achievement has important implications. Scholars are currently challenging the applicability of the lessons of the past century to recent waves of immigrants, arguing that the potential for downward mobility and the advantages of ethnic networks and enclaves could make it advantageous for some groups to maintain their culture of origin (Portes and Zhou 1993; Rumbaut 1997). Our

findings suggest that this kind of differentiated assimilation is not a purely contemporary phenomenon. Instead, it has a strong historical precedent among at least one group, Russian Jews. This finding is surprising given the attention found in scholarship, biography, and literature on the importance for Jews of Americanizing first and last names. In show business, for example, Jews considered name changes a crucial ingredient for success (Bial 2005; Buhle 2004; Lieberson 2010). Yet we find this was not true for the general population. Being named Moses or Mordechai did not confer disadvantage—quite the opposite.[19]

On the other hand, the literature on the Eastern European Jewish immigrant experience attributes advantages to displaying a strong ethnic identity. Scholars have described the important role of ethnic aid societies in Jews' economic success (Kahan 1978; Kasinitz 2008). In a climate of discrimination, the use of ethnic networks are advantageous for occupational advancement, particularly for a minority group that tends toward entrepreneurship and self-employment (Rischin 1977). Facing discrimination from non-Jews, the potential benefits of an "American"-sounding first name are small in comparison to strong identification with one's own *landsleit.*

The recently released IPUMS 100 percent sample from the 1940 census will allow researchers to advance beyond the present work.[20] First, it will make it possible to further explore the role of ethnic enclaves, which requires more refined geographic measurements beyond those available in the 1930 5 percent sample. Second, as noted earlier, interpretation of the results for Russian immigrants requires a caveat. Although the vast majority of these immigrants were Jewish, there were also Christian immigrants. Part of the positive gradient we observe for ethnic names may be driven by the lower occupational earnings of non-Jewish immigrants, who had less "ethnic" first names, like John. However, Table 4 shows that the positive association between ENI score and occupational income is stronger among Russians with recognizably Jewish last names, providing support for the protective effect suggested by the ethnic enclave thesis. The 1940 sample will allow researchers to further explore this issue using a larger sample size. Additional data sources, such as shipping registers from the turn of the last century, which sometimes include an indicator for Jewish ethnicity, may also prove helpful in future work (Spitzer 2015).

CONCLUSIONS

Amid the current wave of interest in using big data for social science research (Bond et al. 2012; Golder and Macy 2011; Savage and Burrows 2007), from twitter feeds, to electronic purchase records, to social media data, it is easy to forget that hundreds of millions of historical census records have already been transcribed and are awaiting exploitation. Ruggles (2014) proposes using census microdata as a variant of "big data." Here, we build

[19]A recent study by Zhang, Zuckerman, and Obukhova (forthcoming) offers a detailed analysis of name choices among Jews in the early twentieth-century United States. They suggest that Jews chose names that were less fashionable yet established in an effort at "acculturative conservatism."

[20]As this article went to press, we learned of a new, independent analysis of first names using the complete count 1940 census (Abramitzky, Boustan, and Eriksson 2014). This paper also finds negative effects of ethnic naming on certain aspects of economic achievement. Their approach uses different groups to measure ENI (foreign-born versus native-born rather than father foreign-born versus father native-born) and different economic outcomes. But their general findings are reassuringly consistent with those found here.

on this idea by taking advantage of the textual information contained in census microdata. The massive transcription of names in existing census data is an opportunity to work with large-scale textual information that is both publicly available and representative (to the limits of the census) of the population.

Our article shows that it is possible to translate names into quantitative indices of cultural assimilation for children of immigrants, information that can easily be incorporated into models of occupational attainment. We find that for most groups, cultural assimilation as measured by first names is strongly positively correlated with occupational achievement. Our controls for the social class gradient of name-giving strengthens a causal interpretation, wherein parents who were able to and chose to make their children “more American” confer an advantage in terms of occupational achievement. The one exception to this pattern are Russian, predominantly Jewish, immigrants: for this group, sons with ethnic names ended up, on average, in higher-earning occupations. Considering only the children of immigrants with recognizably ethnic last names enabled us to distinguish between the possible mechanisms via which the cultural assimilation effect works. We find that American-sounding first names are an advantage even for people with ethnic last names. We interpret this as indicating that cultural assimilation operates by revealing an “American” orientation among the children of immigrants, rather than by concealing foreign origins.

Using historical datasets with names can be useful for related research topics of substantive interest, including (1) the historical effects of distinctively black names (Cook, Logan, and Parman 2014), (2) the effect of ethnic names on marriage and ethnic intermarriage, and (3) the effect of religion on fertility, as proxied by the use of biblical names (Hacker 1999). The present study can be extended to women by studying the names of daughters of immigrants to assess the hypothesis that cultural assimilation via ethnic naming is related to higher rates of upward class mobility in the marriage market. In addition, it may be possible to study systematic changes in last names using the type of demographic reconstruction that have been used to separate the demographic and cultural factors in the growth of white ethnic populations (Hout and Goldstein 1994).

Our methods are also potentially well suited for application to contemporary immigration and naming. Because we measure ethnic distinctiveness statistically by relative frequencies, our approach can be applied to data having only coded numeric strings for distinct names, with no need for the spelling out of the name itself. This opens up the possibility of names research with modern sources still limited by confidentiality restrictions to study topics of contemporary interest, such as the new immigrant second generation (Portes 1996).

Acknowledgments

Funding

Goldstein: NIH Population Center Grant (#NICHD R24 HD073964-01A1).

Stecklov: Israel Science Foundation Grant (#1509/14).

The authors benefitted from helpful discussions with Dan Lichter, Seth Sanders, and Ezra Zuckerman as well as feedback from our presentations at the Population Association of America (PAA) Meeting in New Orleans, LA, in 2013 and the Alpine Population Meetings (AlpPop) in La Thuile, Italy in 2013. We also acknowledge the

thoughtful input from reviewers and editors at *ASR* and particularly the suggestion of a reviewer to incorporate last names into our study.

References

Abramitzky, Ran; Boustan, Leah; Eriksson, Katherine. Cultural Assimilation during the Age of Mass Migration. 2014. Retrieved January 15, 2015 (http://wpcarey.asu.edu/sites/default/files/uploads/department-economics/abramitzkyboustanasu.pdf)

Alba, Richard D.; Nee, Victor. Remaking the American Mainstream: Assimilation and Contemporary Immigration. Cambridge, MA: Harvard University Press; 2005.

Arai, Mahmood; Thoursie, Peter Skogman. Renouncing Personal Names: An Empirical Examination of Surname Change and Earnings. Journal of Labor Economics. 2009; 27(1):127–47.

Berger, Jonah; Le Mens, Gaël. How Adoption Speed Affects the Abandonment of Cultural Tastes. Proceedings of the National Academy of Sciences. 2009; 106(20):8146–50.

Berkson, Joseph. Are There Two Regressions? Journal of the American Statistical Association. 1950; 45(250):164–80.

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

Besnard, Philippe; Desplanques, Guy. Temporal Stratification of Taste: The Social Diffusion of First Names. Revue française de sociologie. 2001; 42(1):65–77.

Bial, Henry. Acting Jewish: Negotiating Ethnicity on the American Stage & Screen. Ann Arbor: University of Michigan Press; 2005.

Bond, Robert M.; Fariss, Christopher J.; Jones, Jason J.; Kramer, Adam DI.; Marlow, Cameron; Settle, Jaime E.; Fowler, James H. A 61-Million-Person Experiment in Social Influence and Political Mobilization. Nature. 2012; 489(7415):295–98. [PubMed: 22972300]

Brubaker, Rogers. The Return of Assimilation? Changing Perspectives on Immigration and Its Sequels in France, Germany, and the United States. Ethnic and Racial Studies. 2001; 24(4):531–48.

Buhle, Paul. From the Lower East Side to Hollywood: Jews in American Popular Culture. New York: Verso; 2004.

Carlsson, Magnus. Experimental Evidence of Discrimination in the Hiring of First- and Second-Generation Immigrants. Labour. 2010; 24(3):263–78.

Clark, Gregory. The Son Also Rises: Surnames and the History of Social Mobility. Princeton, NJ: Princeton University Press; 2014.

Clark, Gregory; Cummins, Neil. Intergenerational Mobility in England, 1858–2012. Wealth, Surnames, and Social Mobility. Economic Journal. 2015; 125(582):61–85.

Cook, Lisa D.; Logan, Trevon D.; Parman, John M. Distinctively Black Names in the American Past. Explorations in Economic History. 2014; 53:64–82.

Coulmont, Baptiste. Sociologie Des Prénoms. Paris: La Découverte; 2011.

Duncan, Otis Dudley. A Socioeconomic Index for All Occupations. In: Reiss, AJ., editor. Occupations and Social Status. New York: Free Press; 1961. p. 109-138.

Durbin, James. Errors in Variables. Revue de l'Institut International de Statistique/Review of the International Statistical Institute. 1954; 22(1/3):23–32.

Elchardus, Mark; Siongers, Jessy. First Names as Collective Identifiers: An Empirical Analysis of the Social Meanings of First Names. Cultural Sociology. 2011; 5(3):403–422.

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

Gerhards, Jürgen; Hackenbroch, Rolf. Trends and Causes of Cultural Modernization: An Empirical Study of First Names. International Sociology. 2000; 15(3):501–531.

Gerhards, Jürgen; Hans, Silke. From Hasan to Herbert: Name-Giving Patterns of Immigrant Parents between Acculturation and Ethnic Maintenance. American Journal of Sociology. 2009; 114(4): 1102–28.

Golder, Scott A.; Macy, Michael W. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength across Diverse Cultures. Science. 2011; 333(6051):1878–81. [PubMed: 21960633]

Gordon, Milton M. Assimilation in America: Theory and Reality. Daedalus. 1961; 90(2):263–85.

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

Hacker, J David. Child Naming, Religion, and the Decline of Marital Fertility in Nineteenth-Century America. History of the Family. 1999; 4(3):339–65.

Hout, Michael; Goldstein, Joshua R. How 4.5 Million Irish Immigrants Became 40 Million Irish-Americans: Demographic and Subjective Aspects of the Ethnic Composition of White Americans. American Sociological Review. 1994; 59(1):64–82.

Kahan, Arcadius. Economic Opportunities and Some Pilgrims' Progress: Jewish Immigrants from Eastern Europe in the U.S., 1890–1914. Journal of Economic History. 1978; 38(1):235–51.

Kasinitz, Philip. Becoming American, Becoming Minority, Getting Ahead: The Role of Racial and Ethnic Status in the Upward Mobility of the Children of Immigrants. ANNALS of the American Academy of Political and Social Science. 2008; 620(1):253–69.

Li, Wentian. Analyses of Baby Name Popularity Distribution in U.S. for the Last 131 Years. Complexity. 2012; 18(1):44–50.

Lieberson, Stanley. A Matter of Taste: How Names, Fashions, and Culture Change. New Haven, CT: Yale University Press; 2010.

Lieberson, Stanley; Mikelson, Kelly S. Distinctive African American Names: An Experimental, Historical, and Linguistic Analysis of Innovation. American Sociological Review. 1995; 60(6): 928–46.

Massey, Douglas S. Dimensions of the New Immigration to the United States and the Prospects for Assimilation. Annual Review of Sociology. 1981; 7:57–85.

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

Obama, Barack. Inaugural Addressd by President Barack Obama. 2013. Retrieved November 12, 2015 (http://www.whitehouse.gov/the-press-office/2013/01/21/inaugural-address-president-barack-obama)

Obukhova, Elena; Zuckerman, Ezra W.; Zhang, Jiayin. When Politics Froze Fashion: The Effect of the Cultural Revolution on Naming in Beijing. American Journal of Sociology. 2014; 120(2):555–83.

Perlmann, Joel; Waldinger, Roger. Second Generation Decline? Children of Immigrants, Past and Present – A Reconsideration. International Migration Review. 1997; 31(4):893–922. [PubMed: 12293209]

Portes, Alejandro. The Social Origins of the Cuban Enclave Economy of Miami. Sociological Perspectives. 1987; 30(4):340–72.

Portes, Alejandro, editor. The New Second Generation. New York: Russell Sage Foundation; 1996.

Portes, Alejandro; Rumbaut, Rubén G. Children of Immigrants Longitudinal Study (CILS), 1991–2006. 2012. Retrieved November 15, 2015 (http://doi.org/10.3886/ICPSR20520.v2)

Portes, Alejandro; Zhou, Min. The New Second Generation: Segmented Assimilation and Its Variants. Annals of the American Academy of Political and Social Science. 1993; 530(1):74–96.

Reber, Arthur S. Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious. New York: Oxford University Press; 1996.

Rischin, Moses. The Promised City: New York's Jews, 1870–1914. Cambridge, MA: Harvard University Press; 1977.

Roosevelt, Theodore. Roosevelt Bars the Hyphenated. New York Times. 1915 Oct 13.

Rossi, Alice S. Naming Children in Middle-Class Families. American Sociological Review. 1965; 30(4):499–513.

Ruggles, Steven. Big Microdata for Population Research. Demography. 2014; 51(1):287–97. [PubMed: 24014182]

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

Rumbaut, Ruben G. Paradoxes (and Orthodoxies) of Assimilation. Sociological Perspectives. 1997; 40(3):483–511.

Savage, Mike; Burrows, Roger. The Coming Crisis of Empirical Sociology. Sociology. 2007; 41(5): 885–99.

Seger, Carol A. Implicit Learning. Psychological Bulletin. 1994; 115(2):163–96. [PubMed: 8165269]

Spitzer, Yannay. Pogroms, Networks, and Migration. 2015. Retrieved November 15, 2015 (https://yannayspitzer.files.wordpress.com/2014/11/spitzer_pogromsnetworksmigration_150529.pdf)

Stevens, Gillian; Featherman, David L. A Revised Socioeconomic Index of Occupational Status. Social Science Research. 1981; 10(4):364–95.

Sue, Christina A.; Telles, Edward E. Assimilation and Gender in Naming. American Journal of Sociology. 2007; 112(5):1383–1415.

Warner, W Lloyd; Srole, Leo. The Social Systems of American Ethnic Groups. New Haven, CT: Yale University Press; 1945.

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

Waters, Mary C. Ethnic and Racial Identities of Second-Generation Black Immigrants in New York City. International Migration Review. 1994; 28(4):795.

Waters, Mary C.; Jiménez, Tomás R. Assessing Immigrant Assimilation: New Empirical and Theoretical Challenges. Annual Review of Sociology. 2005; 31(1):105–25.

Watkins, Susan Cotts; London, Andrew S. 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; 18(2):169–209.

Wilson, Kenneth L.; Portes, Alejandro. Immigrant Enclaves: An Analysis of the Labor Market Experiences of Cubans in Miami. American Journal of Sociology. 1980; 86(2):295–319.

Zhang, Jiayin; Zuckerman, Ezra W.; Obukhova, Elena. A Lack of Security or of Cultural Capital? Acculturative Conservatism in the Naming Choices of Early 20th-Century U.S. Jews. Social Forcess. Forthcoming.

Biographies

Joshua R. Goldstein is the Chancellor's Professor of Demography at the University of California-Berkeley. His research interests include fertility, family, race and ethnicity, immigration, aging, and mathematical demography. This is his third article in *ASR*, following "How 4.5 Million Irish Immigrants Became 40 Million Irish Americans: Demographic and Subjective Aspects of the Ethnic Composition of White Americans" (with Michael Hout) and "Marriage Delayed or Marriage Forgone? New Cohort Forecasts of First Marriage for U.S. Women" (with Catherine T. Kenney).

Guy Stecklov is a Professor of Sociology and Anthropology at Hebrew University. His research focuses on the relationship between social and economic resources and demographic behavior as well as on research methodology. He is currently studying the impact of migrant ethnic identity on patterns and consequences of assimilation in Israel (with Ahmad Hilhel) and in the United States (with Joshua Goldstein). Additional ongoing work includes explorations of the use of survey-based field experiments to understand social interaction (with Alex Weinreb and Mariano Sana) and analysis of the impact of cash transfer programs on fertility and migration (with Paul Winters and Alex Weinreb).

APPENDIX

A. Linked File Analysis

In the main analysis, we adjust for class background effects on name-giving by calculating the average occupational income of the households of origin by esach child's first name. In theory, this analysis at the level of the name produces unbiased coefficients (Berkson 1950; Durbin 1954). However, this approach may also introduce statistical error, because grouped estimates are, in general, less precise, and because households of origin may not match perfectly to the actual households of origin given that the 1900 and 1930 data are essentially drawn from two independent samples. In this appendix, we provide a reanalysis based on longitudinally linked samples across census years.

The IPUMS USA project provides files linking individuals across the microdata samples from various censuses to the complete census of 1880. These samples are too small for our main analysis but are useful to help validate the use of name-grouping methods. Because of restriction in sample size, we were able to generate sufficient samples for linked father–son pairs only for German immigrants and their sons. We further combined results from 1910, 1920, and 1930, each of which includes some persons matched once and some potentially matched more than once across the samples. We calculated ENI values based on children younger than 10 years of age in 1880 using the 10 percent sample of the 1880 census.

Table A1 shows results of this analysis. The first two columns show regressions on the son's ENI score. The results show that higher-earning fathers tend to give their sons slightly less ethnic names. The first column shows the results using father's occupation averaged over the first name of sons. The second column uses father's occupation at the individual level, which is possible for the linked file. In this case, use of the average actually produces a similar negative result, with both coefficients being small and negative. The next three columns show regressions on the occupation of sons. Column three shows that changing from a completely "American" name to a completely German name is negative and substantively significant, decreasing occupational earnings by about 13 percent. The right-most columns show what happens after including father's occupation. In both cases, the effect of father's occupation is statistically significant and positive and the effect of ENI also remains statistically significant. The estimated coefficients and their standard errors are nearly the same in the two analyses (with average and individual-level measurement of father's occupation).

In this case, class background has a weak effect on naming so adjusting for this factor does little to alter the effect of ENI. Therefore, our overall conclusion is that use of individual-level and name-averaged father's occupation produce substantively identical results.

Table A1

Analysis of German Father–Son Linked Pair Samples 1880 to 1910, 1880 to 1920, and 1880 to 1930 IPUMS Files

	Avg. OCC	Indiv. FOCC	No FOCC	Avg. OCC	Indiv. FOCC
	ENI			Son OCCSCORE	
ENI			−3.517** (−1.071)	−3.127** (−1.069)	−3.294** (−1.044)
Avg. FOCCS	−.005*** (−.001)			.303*** (−.066)	
Indiv. FOCCS		−.001 (−.001)			.263*** (−.026)
Urban	−.002 (−.01)	−.004 (−.011)	11.211*** (−.492)	1.806*** (−.497)	9.226*** (−.518)
Age	0 (−.001)	0 (−.001)	−.001 (−.043)	.013 (−.043)	.011 (−.042)
American Mother	−.073*** (−.013)	−.074*** (−.013)	.847 (−.61)	.747 (−.607)	.77 (−.594)
1920 Census Year	.016 (−.014)	(.015 (−.014)	.639 (−.684)	.562 (−.681)	.747 (−.667)
1930 Census Year	.02 (−.021)	.017 (−.021)	.577 (−.991)	.327 (−.988)	.61 (−.966)
Constant	.732*** (v.058)	.629*** (−.05)	9.333*** (−2.457)	2.319 (−2.883)	5.690* (−2.422)
R-squared	.028	.022	.255	.263	.293
No. of cases	1,988	1,988	1,904	1,904	1,904

Note: Standard errors are in parentheses. Authors' analysis of U.S. native-born men age 20 to 59 with foreign-born fathers in IPUMS linked samples. Regional controls (eight dummies) included in all models but not shown.

* $p < .05$;

** $p < .01$;

*** $p < .001$(two-tailed tests).

B. How Ethnic Naming Depends on Father's Occupational Income

Here we show results from our analysis of the effect of father's occupational income on the ENI score. Results of the regression in Table B1 show how ENI in 1900 depends on father's occupational score. We estimated separate regressions for each group. For the children of Irish, Italian, German, and Polish immigrants, higher-earning fathers tend to give less ethnic names. For the children of Russian fathers, the gradient is in the opposite direction: father's with higher occupational scores give more ethnic names. The standard deviation of occupational earnings for Russian-born fathers in our sample is about 4. According to these results, a one standard deviation increase in father's occupational score would be associated with an increase in ENI of about .08, which is about an increase of one-third of a standard deviation in ENI for the sons of Russian-born fathers.

Table B1

How Ethnic Naming Depends on Father's Occupational Income

	Ireland	Italy	Germany	Poland	Russia
Father's Occupational Score	(−.025*** (−.001	(−.003*** −.001	−.023*** 0	−.009*** (−.001)	(.019*** −.001
Constant	(1.309*** (−.014	(.798*** −.017)	1.166*** (−.007)	.923*** (−.014)	(.210*** −.014

	Ireland	Italy	Germany	Poland	Russia
R-squared	.056	.001	.072	.016	.089
No. of cases	34824	15597	76241	13380	12739

Note: Standard errors are in parentheses. Authors' analysis of foreign-born fathers in 1900 U.S. IPUMS sample with U.S.-born sons. ENI (ethnic name index) is calculated separately for each group for sons under 10 years in 1900. Column headings refer to father's birthplace.

* $p < .05$;

** $p < .01$;

*** $p < .001$(two-tailed tests).

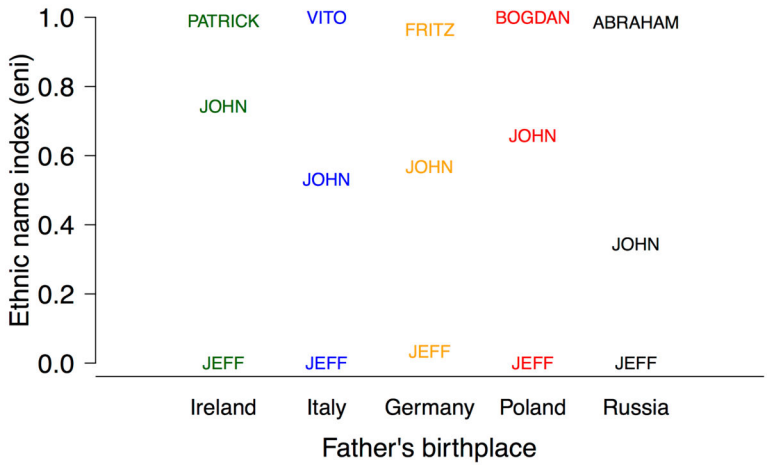

Figure 1.
Examples of Ethnic Distinctiveness Index (ENI) by Birthplace of Father

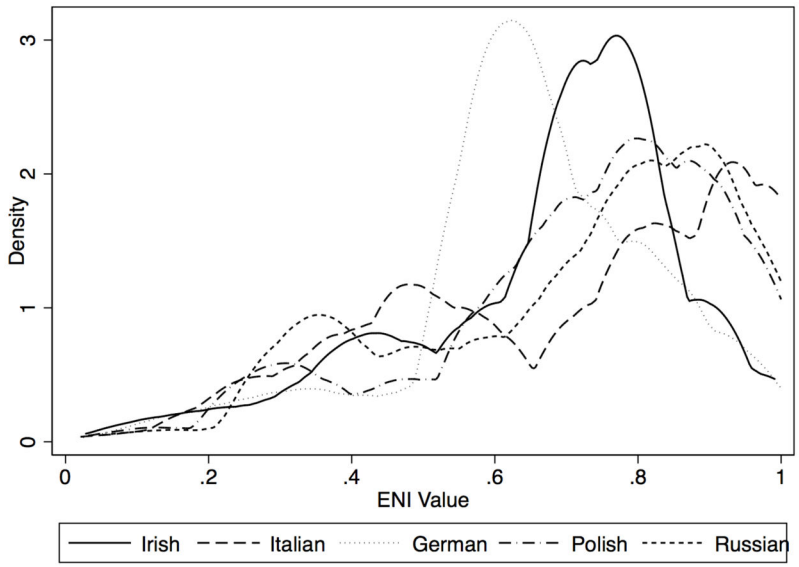

Figure 2.
Kernel Density Distributions for Ethnic Name Index (ENI) Calculated for Five Ethnic Groups in 1930 (IPUMS)

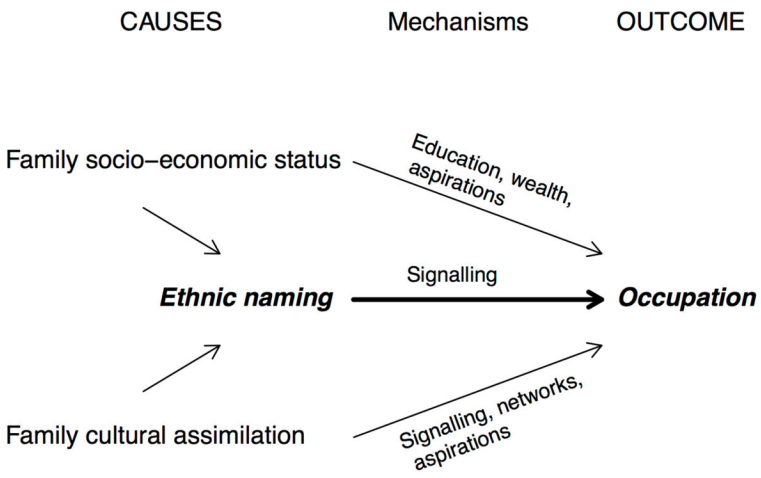

Figure 3.
A Model Relating Ethnic Naming to Children's Occupational Attainment

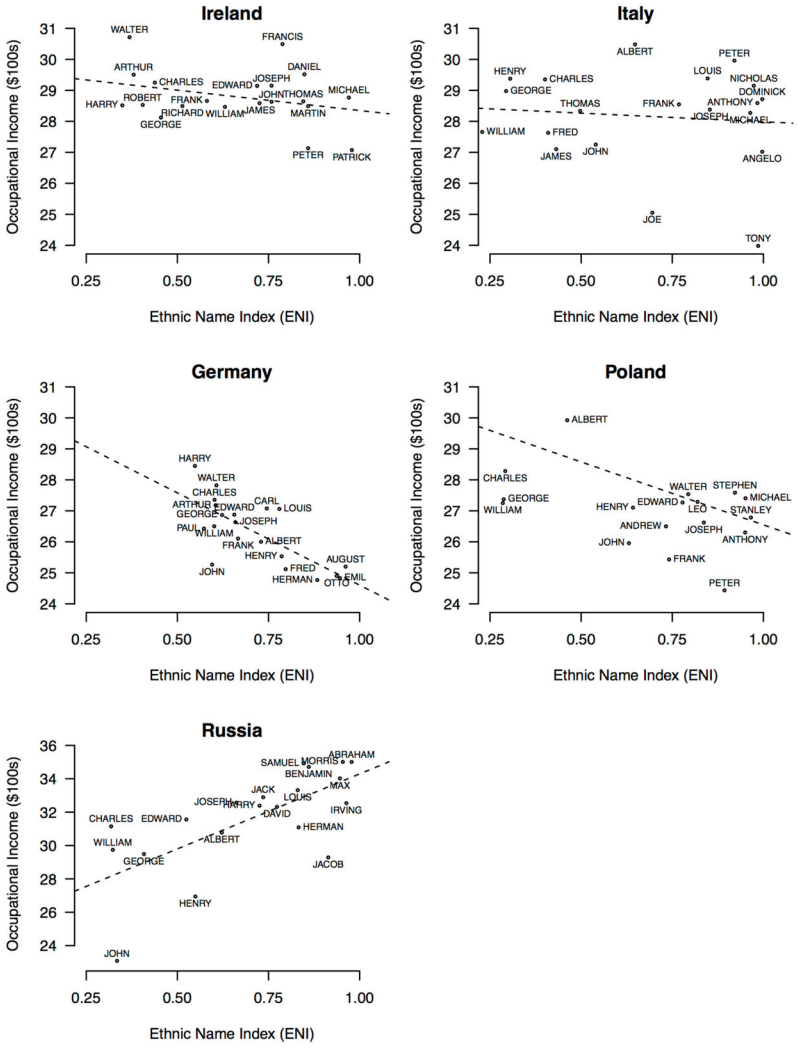

Figure 4.

Relationships between Ethnic Distinctiveness and Occupational Achievement for the Most Popular Names within Each Ethnic Group

Note: The scale on the *y*-axis for OCCSCORE differs across the five figures. We include the 20 most popular names among people age 30 to 49 in each group.

Table 1

Predictors of Annual Occupational Earnings in 1930 (in 100s of dollars)

	(1)	(2)	(3)	(4)	(5)
	Irish	**Italians**	**Germans**	**Polish**	**Russians**
ENI	−.755** (.283)	−.582* (.285)	−1.020*** (.198)	−2.600*** (.363)	2.088*** (.486)
Urban	6.957*** (.155)	3.088*** (.225)	9.724*** (.0770)	7.368*** (.210)	1.78*** (.405)
Age	.0752*** (.00492)	.237*** (.00974)	.0771*** (.00330)	.184*** (.00941)	.284*** (.0141)
American Mother	1.014*** (.130)	.351 (.276)	.709*** (.0776)	.597 (.321)	−.113 (.509)
Constant	18.89*** (.335)	16.79*** (.446)	17.01*** (.295)	15.63*** (.502)	10.79*** (.767)
Observations	33,037	14,516	73,644	12,683	11,795
R-squared	.073	.057	.237	.138	.194

Note: Standard errors are in parentheses. Authors' analysis of U.S. native-born men age 20 to 59 with foreign-born fathers in 1930 U.S. IPUMS 5 percent sample. ENI (ethnic name index) is calculated separately for each group. Column headings refer to father's birthplace. Regional controls (eight dummies) included in all models but not shown.

* $p < .05$;

** $p < .01$;

*** $p < .001$ (two-tailed tests).

Table 2

Predictors of Annual Occupational Earnings in 1930 (in 100s of dollars), Controlling for Average Father's Occupational Earnings by First Name

	(1)	(2)	(3)	(4)	(5)
	Irish	**Italians**	**Germans**	**Polish**	**Russians**
ENI	−.582 * (.291)	−.520 (.286)	−.495 * (.206)	−2.201 *** (.363)	1.496 ** (.510)
Urban	6.953 *** (.155)	3.088 *** (.225)	9.685 *** (.0771)	7.176 *** (.210)	1.58 *** (.408)
Age	.0755 *** (.00493)	.236 *** (.00974)	.0780 *** (.00330)	.181 *** (.00937)	.285 *** (.0141)
American Mother	1.010 *** (.130)	.337 (.276)	.692 *** (.0775)	.496 (.319)	−.213 (.509)
Father's OCC Score	.0787 * (.0311)	.0911 *** (.0249)	.166 *** (.0174)	.268 *** (.0257)	.121 *** (.0316)
Constant	16.80 *** (.892)	14.55 *** (.758)	12.82 *** (.529)	9.426 *** (.778)	8.156 *** (1.031)
Observations	33,037	14,516	73,644	12,683	11,795
R-squared	.073	.058	.238	.145	.195

Note: Standard errors are in parentheses. Authors' analysis of U.S. native-born men age 20 to 59 with foreign-born fathers in 1930 U.S. IPUMS 5 percent sample. ENI (ethnic name index) is calculated separately for each group. Column headings refer to father's birthplace. Father's occupational score (earnings) are based on average earnings of fathers of sons under age 10 in 1900 by son's first name. Regional controls (eight dummies) included in all models but not shown.

* $p < .05$;

** $p < .01$;

*** $p < .001$ (two-tailed tests).

Table 3

County Fixed-Effects Predictors of Occupational Earnings in 1930 (in 100s of dollars)

	(1)	(2)	(3)	(4)	(5)
	Irish	**Italians**	**Germans**	**Polish**	**Russians**
ENI	−.756* (.296)	−.467 (.287)	−.472* (.206)	−1.952*** (.365)	1.267* (.522)
Urban	5.570*** (.207)	2.007*** (.293)	8.078*** (.102)	3.741*** (.291)	7.021*** (.688)
Age	.0826*** (.00503)	.233*** (.00999)	.0766*** (.00334)	.202*** (.00969)	.282*** (.0148)
American Mother	1.104*** (.133)	.448 (.286)	.686*** (.0783)	1.132*** (.330)	−.113 (.538)
Father's OCC Score	.0651* (.0319)	.0909*** (.0252)	.133*** (.0174)	.198*** (.0264)	.0747* (.0328)
Constant	18.57*** (.915)	15.92*** (.760)	14.57*** (.479)	13.48*** (.775)	12.10*** (1.114)
Observations	33,037	14,516	73,644	12,683	11,795
R-squared	.032	.042	.088	.057	.042

Note: Standard errors are in parentheses. Authors' analysis of U.S. native-born men age 20 to 59 with foreign-born fathers in 1930 U.S. IPUMS 5 percent sample. ENI (ethnic name index) and ethnic last name indicators are calculated separately for each group. Column headings refer to father's birthplace. Father's occupational score (earnings) are based on average earnings of fathers of sons under age 10 in 1900 by son's first name. Regional controls (eight dummies) included in all models but now shown.

* $p < .05$;

** $p < .01$;

*** $p < .001$ (two-tailed tests).

Table 4

Estimated ENI Coefficients for Predicting Occupational Earnings in 1930 (in 100s of dollars), Using Three Separate Specifications to Account for Last Name Ethnicity

	Main Model Coefficient Estimates	Controlling for Recognizably Ethnic Last Name	Selection on Recognizably Ethnic Last Name[a]	Last Name Fixed Effects
Irish	−.582 (.291)	−.59 (.291)	−1.24 (1.048)	−.647 (.344)
Italian	−.52 (.286)	−.523 (.286)	−.408 (.309)	−.432 (.442)
German	−.495 (.206)	−.51 (.206)	−.179 (.334)	−.241 (.275)
Polish	−2.201 (.363)	−1.992 (.364)	−1.094 (.541)	−1.424 (.608)
Russian	1.496 (.51)	1.443 (.511)	2.367 (1.04)	1.061 (.747)

Note: Standard errors are in parentheses. Authors' analysis of U.S. native-born men age 20 to 59 with foreign-born fathers in 1930 U.S. IPUMS 5 percent sample. ENI (ethnic name index) and ethnic last name indicators are calculated separately for each group. Row headings refer to father's birthplace. Father's occupational score (earnings) are based on average earnings of fathers of sons under age 10 in 1900 by son's first name. Models include controls for age, estimated father's occupational income, American mother, urban residence, and regional controls (eight dummies).

[a]Sample sizes for models in columns 1 and 2 are similar to those shown in Tables 1 and 2. In the ethnic last name models, our samples sizes are as follows: Irish (n = 2,748), Italian (n = 12,515), German (n = 29,935), Polish (n = 4,922), and Russian (n = 3,354). The samples sizes in the last names fixed-effects models are Irish (n = 33,034), Italian (n = 14,513), German (n = 73,633), Polish (n = 12,683), and Russian (n = 11,793).