

The Economic Payoff of Name Americanization

Costanza Biavaschi, *University of Reading*

Corrado Giulietti, *University of Southampton*

Zahra Siddique, *University of Reading*

(forthcoming in the *Journal of Labor Economics*, 2017)

Accepted version: August 2016

Abstract

We provide the first evidence on the magnitude and consequences of the Americanization of migrants names in the early 20th century. We construct a longitudinal dataset of naturalization records, tracking migrants and their naming choices over time. We consistently find that migrants who Americanized their names experienced larger occupational upgrading than those who did not. Name Americanization embodies an intention to assimilate among low-skilled migrants and reveals the existence of preferences for American names within the labor market. We conclude that the trade-off between individual identity and labor market success was present then, as it is today.

The authors would like to thank participants at the 10th IZA Annual Migration Meeting, the 4th TEMPO conference on international migration, the Essex Fresh meeting, the 2014 annual conference of the Royal Economic Society, the 2014 SOLE conference, the 2014 Cliometrics conference, the 2014 International Economic Association conference, the 26th EALE conference, the Irish Quantitative History workshop, as well as seminar participants at IZA, World Bank, Reading, Tsinghua, Southampton, Potsdam, Surrey, Brown, NYU, Southern Denmark, Georgetown, Oxford, Milano Bicocca, and Marco Caliendo, Rajeev Dehejia, Benjamin Elsner, Lidia Farré, Jesús Fernández-Huertas Moraga, Oded Galor, Daniel Hamermesh, Carolyn Moehling, Andrew Oswald, Daniele Paserman, Núria Rodríguez-Planas, Todd Sorensen, Derek Stemple, Christopher Taber, Konstantinos Tatsiramos and David Weil for their helpful comments. Lena Knops and Eleni Kostopoulou provided excellent research assistance. We are also grateful to John R. Logan who shared the crosswalk between enumeration districts and census tracts with us. The authors acknowledge financial support from IZA and a RETF fund from the University of Reading. An earlier version of this paper appeared as IZA Discussion Paper 7725. Contact the corresponding author, Corrado Giulietti, at giulietticorrado@gmail.com. Information concerning access to the data used in this article is available as supplementary material online.

I. Introduction

Americanization, the process by which immigrants strive to assimilate into American society, encompasses several dimensions. One such dimension is the Americanization of migrants' first names, a key aspect of the desire – or need – to conform to the American norm. The importance of first names has long been stressed by sociologists (Lieberson, 2000), and could serve as a crucial marker to understand the returns and trade-offs from assimilation. In this paper, we provide the first evidence on the economic consequences of the Americanization of first names during a pivotal period in American history.

Most Americans have heard stories of migrant ancestors Americanizing their names in the early half of the 20th century. However, such anecdotes are typically stored solely in familial memories and to date no research has measured the extent and drivers behind name Americanization. Table 1 provides preliminary evidence on the magnitude of this phenomenon.

Table 1
Name Americanization and Popular American Names

A. Name Americanization			B. Popular American Names		
Country of Origin	% Americanized	N	Name	% in U.S.-Born Population	% in Americanized Migrants
Italy	19.86	886	John	6.82	8.10
Russian Empire	57.54	749	William	5.74	2.49
Central Europe (excl. DE)	53.06	686	Joseph	3.91	6.93
Southern Europe (excl. IT)	37.69	130	Charles	3.69	2.57
Germany	24.71	437	George	3.64	2.73
Ireland	1.33	376
U.K.	5.00	160	Patrick	0.25	0.08
Northern Europe	19.00	279
Americas	10.89	202	Moishe	0.00	0.00
Other	38.76	178
Total	31.47	4083	Giulio	0.00	0.00

NOTE.—Panel A. Own tabulations from sample of naturalizing immigrants in 1930 New York City, Ancestry.com. Name Americanization is defined as the custom of adopting a first name that was more frequent in the U.S.-born population than the migrant's name at birth. For details, see the definition on page 7 and 8. For a definition of countries of origin, see the notes in Table 2. Panel B. Own tabulations from IPUMS census, 1930 and from the sample of naturalizing immigrants in 1930 New York City. % in U.S.-Born Population indicates the percentage of the U.S.-born male population in 1930 New York having a specific name. % in Americanized Migrants indicates the percentage who chose a particular name among the migrants who Americanized their names.

Defined throughout as the custom of adopting a first name that was more frequent in the U.S.-born population compared with a migrant's own name at birth, name Americanization was a widespread practice. We find that nearly one third of naturalizing immigrants

abandoned their first names by 1930 to acquire names that were more frequent among the U.S.-born population. Panel A of Table 1 shows substantial variation in name Americanization by country of birth, highlighting that migrants from Italy, Russia and Germany were all very likely to abandon their “foreign-sounding names”. Panel B shows popular names in the U.S.-born population and the percentage of migrants who chose those names. For instance, John was the most common name among the U.S.-born, and in our sample about 8% of migrants who Americanized their names chose to be named John.

Widespread name Americanization prompts the questions of whether it was associated with migrants’ economic success and whether specific segments of the migrant population benefited from it. Figure 1 shows that name Americanization into the most popular names – e.g., the top tercile – was associated with an average occupation-based earnings increase of about 8%. These gains were larger than those experienced by migrants Americanizing into less popular names – e.g., the bottom tercile – and even more so than those experienced by migrants who kept their original name.¹

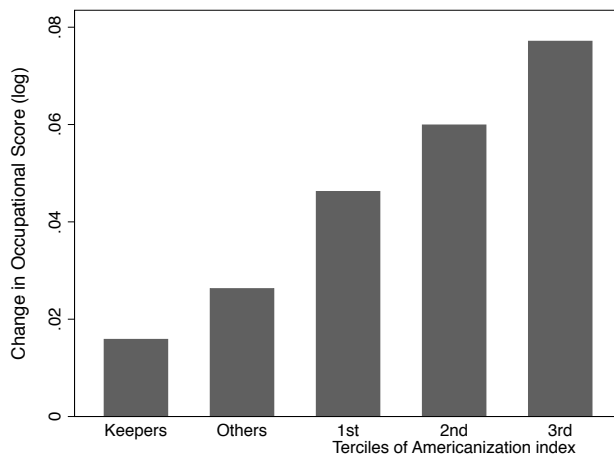


FIG. 1.—Change in log-occupational score and name Americanization. The y-axis represents the change in log-occupational score. The x-axis represents different values of the change of the Americanization index (for details, see the definition on page 7 and 8). Name Americanization is defined as the custom of adopting a first name that was more frequent in the U.S.-born population than the migrant’s name at birth. The Americanization index varies between 0 and 1, and its change over time varies between -1 and 1. “Keepers” are migrants whose index did not change over time (i.e., they did not change name or changed into equally frequent names). “Others” include migrants whose index change was negative (i.e., changed into less frequent names) and migrants whose index changed twice over the period of observation. The “Terciles of Americanization index” refer to positive changes of the index.

¹Our analysis focuses on first names. Surname Americanization was less common, with only 7.03% of migrants Americanizing surnames. Section V includes a model that controls for surname Americanization.

We devote much of our analysis to verifying that the payoff from name Americanization persists once we compare similar migrants. For this purpose, we have digitalized a unique dataset in which we are able to follow the same individual over time without resorting to record linkage. We collect a random sample of migrants who completed their naturalization papers by 1930 in New York City. We are able to follow individuals over time due to the nature of the naturalization process and documentation procedure, which required migrants to first file a declaration of intention to become citizen and later a petition for naturalization. By exploiting this two-step process, we obtain information on names and a wide range of individual characteristics over a period longer than 10 years, from the time of arrival – on average 1918 – until the time of naturalization in 1930.

We estimate the economic payoff of name Americanization using several empirical models. We start by controlling for time-varying individual characteristics that are often unavailable in standard datasets – such as the census, and even modern sources – allowing us to compare migrants with similar socio-demographic traits, who faced analogous nationality and local labor market-specific earnings trajectories. Next, we exploit the longitudinal nature of our data. By examining *changes* in individual occupational-based earnings of the same migrant over time, we are able to compare individuals with the same time invariant unobservables affecting economic success, such as family background and individual ambition. Next, we exploit variation in the timing of name Americanization by focusing only on migrants who Americanize their names. This strategy allows us to compare labor market trajectories of individuals who are equally eager to succeed and equally willing to invest in U.S.-specific skills. Finally, by comparing the labor market trajectories of individuals holding the same name at birth, some who Americanize their name and some who do not, we are able to pin down differences stemming from name-specific perceptions within the U.S. labor market.

All our empirical models reveal the existence of a sizeable payoff from name Americanization. Across specifications, average returns to name Americanization are between three and five percent. This payoff is much larger for migrants who change from purely ethnic names to the most popular American names. For instance, when comparing the labor market trajectories of two migrants both named Guido at birth, one who Americanizes his name to John and one who keeps his name, John’s occupational-based earning growth is 22% higher than Guido’s occupational-based earning growth. To contextualize these findings, we show that the largest effects of name Americanization are comparable to average occupational upgrades experienced by individuals ageing from 20 to 24 years during that time.

The discrepancy between unconditional correlations and regression estimates reveals that name Americanization masks an intention to assimilate. The direction of the bias suggests the presence of pre-existing constraints to occupational mobility among low-skilled migrants.

Our results are compatible with name Americanization being effective in overcoming these constraints and allowing migrants with otherwise scarce means of economic success to climb up the occupational ladder. Consistent with this interpretation, we show that name Americanization had more pronounced effects for sub-groups who were poor, unskilled and for whom the barriers to labor market success were arguably higher. In addition, name assimilation might also be a response to external pressures. The persistent effect estimated by comparing occupational trajectories of migrants with the same name at birth is strongly suggestive of the presence of preferences for American names or statistical discrimination.

The quantification of name Americanization, its effects and the plausible mechanisms behind it all broaden our understanding of immigrant assimilation throughout the 20th century. Unlike established literature on the economics of names focusing on recent periods, our study examines an era that was pivotal in laying the foundations of modern America.² The existence of a personal trade-off between maintaining individual identity and labor market success indicates that the process of cultural assimilation at the dawn of the modern “melting pot” was instrumental in migrants’ economic advancement. Therefore, such a trade-off was as important during the early-1900s as it is today.

Our data and results are also informative for historical investigations based on record linkage that aim to create longitudinal datasets.³ In these settings, where record linkage is often accomplished through matching by names (besides age and country of birth), Americanization might be a source of a failed match. We show that name Americanization was prevalent among migrants facing stronger barriers in the labor market; consequently immigrants who Americanized their names differed from name keepers in terms of both observable and unobservable traits. Hence, our study can help to understand the representativeness of linked samples, especially for nationalities among whom name Americanization was widespread.

To the best of our knowledge, only two studies have explored the decision to change one’s name, with both focusing on renouncing surnames.⁴ The first relates names to gender

²More common names have been shown to result in better educational outcomes (Figlio, 2005) and labor market success due to reduced discrimination (e.g., Bertrand and Mullainathan, 2004). The relationship might also stem from unobserved factors correlated with naming choices made by parents and economic success (e.g., Fryer and Levitt, 2004). Furthermore, psychologists have shown that first names closer to those of host societies are associated with positive attitudes among host populations (Kang, 1971, Drury and McCarthy, 1980) and in particular among employers, co-workers and customers (Laham et al., 2012).

³The use of longitudinal data to study immigrant assimilation has recently become the gold standard in the migration literature to account for the non-randomness in migrant samples caused by selective return migration (Bandiera et al., 2013 show that return migration was substantial in the early 1900s). Prominent examples are Abramitzky et al. (2012, 2014). Abramitzky et al. (2012) estimate the return to migration by matching Norway-to-U.S. migrants with their brothers who stayed in Norway in the late 19th century. Abramitzky et al. (2014) analyze assimilation by linking migrants and natives in censuses over time.

⁴Other precedents on a similar topic are the sociological study by Broom et al. (1955) and the more recent historical study by Fermaglich (2015). Broom et al. (1955) explores the characteristics of 1,107 petitions for

identity, and studies women who decide to keep their maiden name rather than acquiring their husband’s surname upon marriage (Goldin and Shim, 2004). While this work analyzes surname choice rather than its effect on outcomes, it provides compelling evidence on reverting to the custom of taking the husband’s name during the second half of the 20th century. It also documents that low-skilled women are more likely to change their surname, which is consistent with our finding that name Americanization is more prevalent among low-skilled individuals. The second study explores the effect of surname changes made by immigrants from Asian and Slavic countries living in Sweden in the 1990s (Arai and Thoursie, 2009). In this study, in which only 0.4% of migrants changed surnames, fixed-effects estimates show a 141% increase in earnings. The authors conclude that their result is likely driven by reduced labor market discrimination associated with having a Swedish-sounding surname.

The scant literature is due to a scarcity of datasets containing sufficient information to empirically test whether changing names improves economic outcomes. Furthermore, in a historical setting like ours, pinning the consequences of name Americanization is challenging due to the non-availability of longitudinal data sources, together with difficulties in tackling endogenous name choices. Our data and estimation techniques overcome these limitations and contribute to a better understanding of the pressures that migrants faced to conform to American norms.

In the next section we provide a description of our data. We then present our empirical strategy and discuss the estimated effects of name Americanization, before providing robustness checks and concluding.

II. Data

A. Naturalization Documents

To measure the extent and consequences of name Americanization, we exploit the rich records stemming from the naturalization procedure.⁵ To become citizens, free white aliens residing in the U.S. for at least two years – one spent in the state or territory in which the application

name changes filed at the Los Angeles Superior Court in 1946-1947. However, the study is very descriptive in nature, lacks clarity about how petitions have been selected, does not focus on migrants and does not consider labor market outcomes. Fermaglich (2015) focuses on Jewish-American petitioners and studies their name change patterns, although not the consequences of this behavior. While petitions for official name change are an interesting source of information, they do not provide comprehensive information about the migrants, are not longitudinal in nature and capture a very different phenomenon, involving primarily second-generation, middle-class Americans.

⁵Naturalization was considered the ultimate act of identifying with the American culture. In fact, before the 1924 restrictions – as well as after the imposition of quotas – direct benefits from citizenship were rather limited per se. Immigrants during this time were permanent residents and, before the New Deal, social benefits were too small to provide a motivation for naturalization (Lleras-Muney and Shertzer, 2015).

was made – were required to first file a declaration of intention (also known as first papers). The next step involved filing a petition for admission to citizenship (second papers), which could be undertaken a minimum of five years following the initial declaration. Upon filing citizenship papers, the Bureau of Immigration and Naturalization checked ship manifests and issued a certificate of arrival, which included the name held at birth.⁶ A petition number works as the unique identifier across all documents and allows following the same migrant over time.⁷ Figure 2 shows the three documents and the information they contain. We have extracted all available information from these documents.

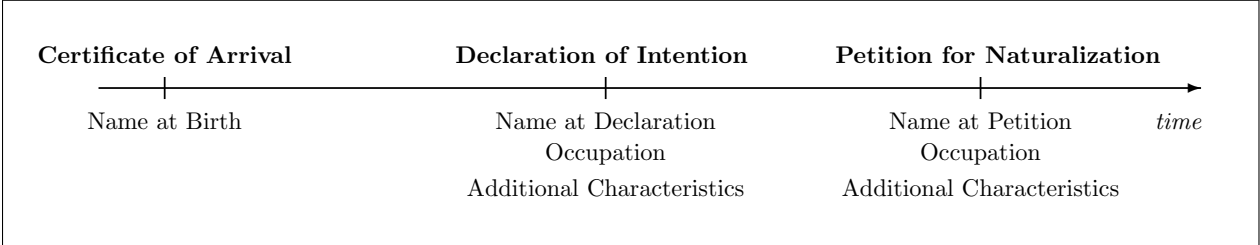


FIG. 2.–Naturalization documents and data extracted.

We obtained the naturalization records from Ancestry.com, a website providing genealogical and family history records. We accessed the complete naturalization records filed at the U.S. District Court for the Southern and Eastern districts of New York City in 1930. The Southern District Court jurisdiction comprises the counties of New York (Manhattan), the Bronx, Westchester, Putnam, Rockland, Orange, Dutchess, and Sullivan. The Eastern District court’s territorial jurisdiction includes the counties of Kings (Brooklyn), Queens, Richmond (Staten Island), Nassau, and Suffolk.⁸ We manually transcribed a 25% random

⁶To be precise, this document contains the name at arrival of the migrant. Contrary to popular belief, names were not changed at Ellis Island. Clerks never wrote down the names of arriving migrants. Instead, they refer to migrants using the ship manifests, which were created at departure by ship pursers who usually spoke the immigrant’s language (Cannato, 2009). Since the only mistakes in compiling passenger lists are likely to be transcription or spelling errors, in the paper we refer to the name at arrival as name at birth.

⁷It should be noted that historical data sources such as the census do not include identifiers that allow for tracking individuals over time. This is the reason why much of the existing studies resort to linking individual records (e.g., Ferrie, 1996, Abramitzky et al., 2014). Such a strategy is clearly not an option for us, as we would be unable to match individuals who have changed name.

⁸More than 30% of all certificates granted in the U.S. pertain to the district of New York (Annual Report of the Commissioner of Naturalization, 1930, p.15). Ancestry.com provides access to 26,113 of the official 30,361 petitions that were filed in the naturalization district of New York (Annual Report of the Commissioner of Naturalization, 1930, p.15), corresponding to more than 85% of all records believed to be available for that year. The remaining records are likely to have been granted by other district courts of New York City or within the state, although they are not available in electronic form. The overwhelming share of naturalizations granted by the Eastern and Southern courts highlights their importance, making the records representative of almost the entire population of naturalizing immigrants residing in the state of New York. See Appendix C for a discussion about the representativeness of our sample.

sample of the available records for 1930. While our randomization procedure involved the collection of naturalization records for males and females, we use the records for male immigrants only.⁹ We restrict the sample to migrants aged 20 years or older at the time of declaration and for whom we have addresses. This gives us a final sample of 4,083 migrants.¹⁰

While our sample covers only part of the immigrants living in the U.S., this group has particular relevance for several reasons. First, it captures immigrants who arrived on average in the late-1910s, during the last surge of migration before U.S. doors were shut. Second, these migrants settled permanently, truly contributing to ‘the making of modern America’. Third, the cultural and economic trade-offs faced by naturalized migrants in New York City might be less pronounced than those faced by non-naturalized migrants in the city or migrants in other states who were less exposed to the new migration waves.¹¹

In the following, we describe the steps taken to derive our key variables.

B. American norm and name Americanization

We measure conformability with American norms by exploring how immigrants’ names compare with American ones. The first step is to rank American names by the frequency of their appearance in the U.S.-born population using the 5% Integrated Public Use Microdata Series (IPUMS) of the 1930 census (Ruggles et al., 2010). We focus on American individuals living in the state of New York at the time of the census. The reason for focusing on New York is to narrow down the pool of names to which migrants in the sample were likely exposed. Panel A of Table 1 shows common American names. Names such as John, William, Joseph, Charles and George were most common in the American-born population, which we consider representative of the American norm.

We then derive a simple metric to capture name Americanization. For a name held by individual i and observed at time t , we define our metric as a normalized frequency of this name in the American-born population living in the state of New York:

$$A_{it} = \frac{W_{it}^k}{\max(W^1, \dots, W^K)} \quad (1)$$

⁹The reasons for doing this are twofold. First, females could naturalize through marriage and not file papers. Therefore, the female subsample is likely to be systematically different from the population of all female migrants. Second, after the Act of 1922, females married to a U.S. citizen were no longer required to file first papers; hence, we cannot observe them over time.

¹⁰Note that immigrants only appear in our sample if their naturalization procedure has been completed – irrespective of whether citizenship was granted. For about 2.6% of migrants in the sample citizenship was refused. We keep these records in the sample. In addition, in 0.2% of cases we could not find the whole set of documents for a particular petition number.

¹¹In Appendix C we compare the characteristics of our sample with samples of natives, foreign-born and naturalized migrants from the Integrated Public Use Microdata Series (IPUMS) of the 1930 census (Ruggles et al., 2010).

where $W_{it}^k = \sum_j \mathbb{1}(\text{Name}_{it} = \text{Name}_j)$ for each name $k \in \{1, \dots, K\}$ among j U.S.-born.

A_{it} is our metric, $\mathbb{1}(\text{Name}_{it} = \text{Name}_j)$ is an indicator variable that takes the value one if the name of a native individual j is the same as the name of individual i at time t ; thus, W_{it}^k counts the number of natives j holding the same name as individual i at time t . The denominator represents the maximum frequency across all names k held by U.S.-born individuals in New York, and hence A_{it} measures how frequent individual i 's name at time t is in the American-born population on a scale from 0 to 1.

Names that are unique to migrants are observed with a frequency equal to 0, whereby the metric associated with these names will also be 0. On the other hand, migrants called John and George will have a metric of 1 and 0.529, respectively (given by $0.036/0.068$, see Table 1). Name Americanization occurs when a migrant changes his name to one that occurs more frequently in the U.S. population, corresponding to an increase in A_i over time. As an example, a migrant called Giovanni who changes into John would be Americanizing his name, with an initial value of A_i equal to zero and a subsequent value equal to 1. On the other hand, A_i would take a value of 1 at any t for a British migrant called John who does not change name. Finally, while Americanizing one's name corresponds to an increase in A_i , the index could also decrease if migrants were to change to more distinctive foreign names.¹²

C. Occupations and Earnings

The lack of earning measures prior to the 1940 census represents a challenge in studying historical labor market outcomes. As is standard in this literature (e.g., Abramitzky et al., 2012, 2014, Olivetti and Paserman, 2015), we rely on a measure of occupational standing available from the IPUMS, which is based on median total income in hundreds of 1950 dollars. It should be noted that although this occupational score has a well-established use in economics, its limitations are also well recognized (e.g., Abramitzky et al., 2012, 2014, Olivetti and Paserman, 2015). For instance, we are unable to measure within-occupation changes in earnings related to the Americanization of a migrant's name. To find an effect, changing one's name must push a migrant into a different mean-wage occupation, which annihilates the actual variation in wages.¹³

¹²The purpose of our index is to capture a distribution of names that is not contaminated by migration. It differs from indices such as Fryer and Levitt (2004) in one aspect: while the aforementioned authors are interested in a relative index that is invariant to name popularity across minority groups, we aim to exclusively measure the popularity of American names. When we construct an index along the BNI of Fryer and Levitt (2004), we find results similar to our baseline estimations, as shown in Table B1, Appendix B.

¹³In section V we perform sensitivity checks on the categorization of occupations. Additional checks are also available in Appendix B.

D. Timing and Payoff of Name Americanization

We show the timing of name Americanization in Figure 3. Migrants could Americanize their names at different points in time: early on, namely between arrival and declaration, or later on, between declaration and petition. We refer to these two groups as “early Americanizers” and “late Americanizers”, respectively. Migrants could also change into more distinctive names (i.e., a negative change in the Americanization index) or change A_{it} twice. We group these latter two into the category “Others”.¹⁴ We refer to migrants who keep their name or change to an equally frequent name as “Keepers”. This latter group accounts for 62.33% of the sample. The “Others” represent 6.20% of the observations. The remaining 31.47% are the name Americanizers. The majority (28.68%) Americanized their name between arrival and declaration, and the rest (2.79%) Americanized their name between declaration and petition.

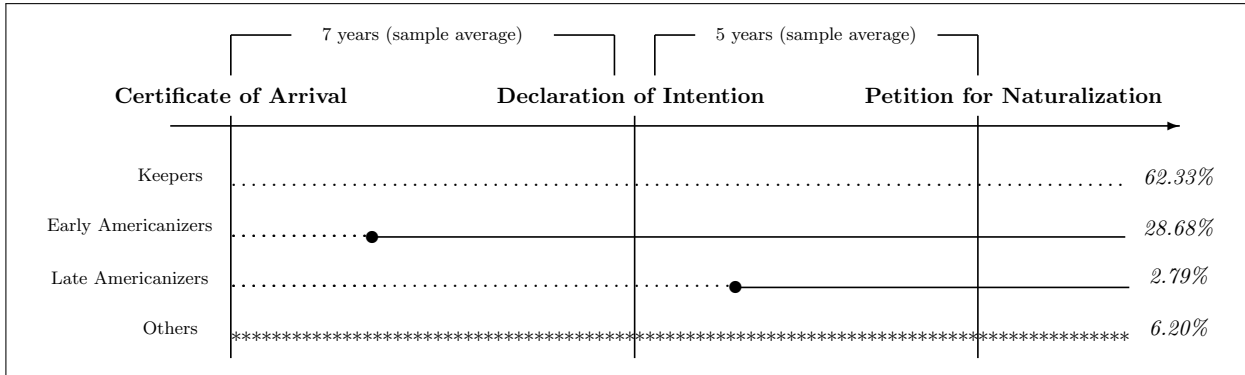


FIG. 3.—Timing of name Americanization. “Keepers” are migrants whose index did not change over time (i.e., they did not change name or change into equally frequent names). “Early Americanizers” are individuals who Americanized their name between arrival and declaration. “Late Americanizers” are individuals who Americanized their names between declaration and petition. A continuous line indicates a name change. “Others” include migrants whose index change was negative (i.e., changed into less frequent names) and migrants whose index changed twice in the period of observation.

Figure 4 shows the average occupational score (in logs) at declaration and at petition across groups of migrants. At the time of declaration, all groups exhibit similar outcomes, although labor market trajectories diverge over time. The groups “Keepers” and “Others” exhibit little change in occupational scores between the two points in time, with the difference in occupational score between declaration and petition being 0.016 log points (s.e. 0.011) and 0.026 (s.e. 0.038), respectively. By contrast, we observe economically important and statistically significant earnings growth for both groups of name Americanizers. For “early Americanizers”, the difference in occupational score between declaration and petition is 0.051

¹⁴The majority in this group (67%) is comprised of individuals who change into more distinctive names.

log points (s.e. 0.018), while for “late Americanizers” is 0.164 (s.e. 0.074). This suggests a persistent earnings increase following name Americanization. Since both the early and late Americanizers changed name by the time of petition, we analyze the two groups together in our baseline regressions. We also perform checks separately for the two groups in Section V.

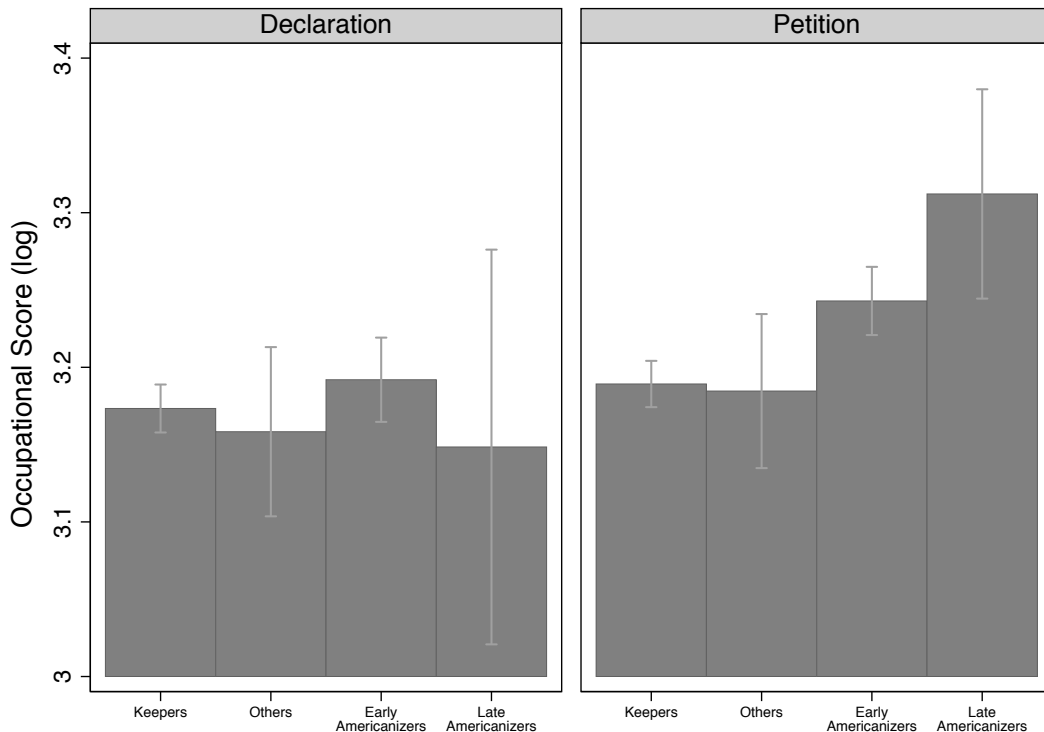


FIG. 4.—Occupational score (logs) over time, for name keepers and name Americanizers. “Keepers” are migrants whose index did not change over time (i.e., they did not change name or changed into equally frequent names). “Others” include migrants whose index change was negative (i.e., changed into less frequent names) and migrants whose index changed twice in the period of observation. “Early Americanizers” are individuals who Americanized their name between arrival and declaration. “Late Americanizers” are individuals who Americanized their names between declaration and petition. T-test for the (log) occupational score being equal over time show the following absolute mean differences and standard errors: 0.016 (s.e. 0.011) for “Keepers”, 0.026 (s.e. 0.038) for “Others”, 0.051 (s.e. 0.018) for “Early Americanizers” and 0.164 (s.e. 0.074) for “Late Americanizers”.

E. Summary Statistics

Table 2 shows summary statistics for selected variables at the time of declaration (top panel) and changes in average characteristics between declaration and petition (bottom panel). Table E1 in the Appendix reports statistics for all variables used in the analysis. At declaration, the average occupational score across name Americanizers and name keepers was similar. The average Americanization index was 0.14, partly reflecting that several migrants have

already Americanized their name by the time of declaration.¹⁵ The bottom panel shows that there are substantial differences in the occupational scores across groups over time. While the change between declaration and petition is essentially zero for “Keepers” and “Others”, it is large and positive for migrants who Americanize their name (0.06 log points). Finally, the Americanization index increased by 0.08, i.e., by half the level at declaration. Among those who Americanized their names, the index increased by 0.25, i.e., by about one standard deviation.¹⁶

Table 2
Characteristics by Level of Americanization

Variable	All	Keepers	Others	Americanize	1 st	2 nd	3 rd
	At Declaration						
Log Occupational Score	3.177 (.436)	3.173 (.399)	3.158 (.444)	3.188 (.499)	3.174 (.510)	3.220 (.495)	3.172 (.492)
Americanization Index	.145 (.261)	.113 (.237)	.066 (.187)	.225 (.297)	.024 (.054)	.104 (.071)	.549 (.311)
Age	31.166 (8.651)	31.484 (8.584)	30.352 (7.962)	30.696 (8.883)	30.886 (9.322)	30.156 (8.614)	31.028 (8.674)
Years Since Migration	7.036 (7.258)	5.956 (6.826)	6.470 (6.584)	9.284 (7.696)	8.806 (7.748)	9.550 (7.778)	9.514 (7.556)
	Difference Petition-Declaration						
Log Occupational Score	.031 (.478)	.016 (.435)	.026 (.485)	.061 (.552)	.046 (.652)	.060 (.491)	.077 (.493)
Americanization Index	.077 (.210)	.000 (.000)	-.037 (.113)	.251 (.304)	.020 (.019)	.114 (.055)	.621 (.255)
Age	5.159 (1.741)	5.202 (1.712)	5.466 (1.818)	5.013 (1.771)	5.100 (1.797)	4.849 (1.764)	5.082 (1.744)
Years Since Migration	4.702 (1.716)	4.743 (1.689)	5.028 (1.763)	4.556 (1.748)	4.645 (1.800)	4.368 (1.711)	4.650 (1.719)
N	4083	2545	253	1285	439	418	428

NOTE.—Standard deviations in parentheses. “Keepers” are migrants whose index did not change over time (i.e., they did not change name or changed into equally frequent names). “Others” include migrants whose index change was negative (i.e., changed into less frequent names) and migrants whose index changed twice in the period of observation. “Americanize” refers to migrants who changed names into more frequent names. 1st, 2nd and 3rd refer to terciles of the change in the Americanization index as defined in the text.

¹⁵The top panel of Table E1 in the Appendix shows that across the characteristics reported at the time of declaration, significant differences arise primarily in three traits: first, at the time of declaration migrants who Americanized their names had stayed longer in the U.S. compared with name keepers (9 years versus 6 years, respectively); second, they had a higher probability of having a U.S.-born child; and third, the distribution by country of birth varied substantially. In particular, migrants from Russia, Southern and Eastern Europe were more likely to Americanize their names compared with other migrants. This is likely driven by the names of these migrants being less common in the U.S. than those of migrants from the U.K. or Ireland. These patterns persist when examining the terciles of the Americanization index separately.

¹⁶At the same time, the bottom panel in Table E1 in the Appendix shows that there are very few differences across groups in terms of changes over time of the remaining characteristics.

III. Empirical Strategy

We apply several estimators to a general model, which takes the form:

$$y_{it} = \beta_0 + \beta_1 A_{it} + \mathbf{x}'_{it} \gamma + c_i + \epsilon_{it}, \quad (2)$$

where y_{it} is the log-occupational score of individual i observed at time t and A_{it} is our key explanatory variable, as defined in equation (1), representing the normalized frequency of individual i 's name at time t in the U.S.-born population.

\mathbf{x}'_{it} includes time-varying socio-economic characteristics such as marital status, a binary variable taking the value one if the spouse is U.S.-born, number of children, a binary variable taking the value one if there are any U.S.-born children, years since migration and a binary variable taking the value one for arrival prior to 1921.¹⁷ The naturalization documents also contain the residential address (street name and house number) of the migrant and his dependents at declaration and at petition. We use addresses in two ways: first, to create indicators for local labor markets, defined as residence in one of the 59 community districts of New York City; and second, to create an indicator for migrants who reside outside New York City.¹⁸ We also include interactions between these variables and years since migration. These regressors capture socio-economic factors that may be correlated with the decision to Americanize one's name. Years since migration, country of birth, arrival cohort, residence in New York City, local labor market indicators and the interactions terms capture assimilation patterns in the U.S. and changes in cohort quality over time, as well as cohort-, country-of-birth- and local-labor-market-specific assimilation trends.¹⁹

The term c_i is an unobserved, time-invariant individual effect, while ϵ_{it} captures unobserved time-varying effects. Both c_i and ϵ_{it} may be potentially correlated with A_{it} although

¹⁷We selected 1921 as the preferred threshold since the first quota system was implemented during that year, although the results do not change if we modify the definition of arrival cohort.

¹⁸We implement a geocoding procedure through which we use migrants' addresses to derive geographical coordinates, subsequently assigning local labor markets to each individual. We flagged addresses for which some judgement was made on our side. We identified local labor markets with the New York City community districts, motivated by the finding that community districts best capture individuals' labor market outcomes across geographic entities (e.g., enumeration districts, tracts, neighborhoods). New York City comprises 59 Community Districts. Our data also contain migrants with addresses located in the 12 Joint Interest Areas, which are nowadays not destined to residence purposes (e.g., major parks and airports). At the time of declaration, 481 migrants lived outside New York, and 61 migrants lived in New York but outside New York City. We constructed an indicator for these two groups. Further details about the construction of the local labor markets are given in Appendix D.

¹⁹We prefer using the years since migration as a continuous variable because we want to employ it to estimate a variety of interactions and at the same time keep the econometric model as parsimonious as possible, given our sample size. Nonetheless, in unreported regressions, we corroborated that results remain virtually the same if years since migration were introduced as a categorical variable.

the sign of this correlation and the net impact on β_1 is unclear. For instance, c_i will be positively correlated with A_{it} if migrants who Americanize their names are more ambitious and from better socio-economic backgrounds. On the other hand, more ambitious migrants from higher social classes might also pursue alternative strategies for socio-economic improvement rather than name Americanization. In such a case, c_i will be negatively correlated with A_{it} , as migrants more likely to Americanize their names would be those with lower ability and facing stronger barriers in the labor market. In principle, ϵ_{it} might include time-varying unobserved components of occupational success such as language ability, motivation to invest further in U.S.-specific skills or job-search activities. Since migrants might Americanize their names in response to their occupational achievements, ϵ_{it} might also be potentially correlated with A_{it} – an issue that our identification strategy will tackle.

To pin down the economic payoff of name Americanization and disentangle some of the plausible explanations behind it, we adopt several estimation methods. We start by pooling the data and apply an OLS estimator to equation (3). Accordingly, we compare migrants with similar observable characteristics such as personal and labor market attributes, but disregard the potential correlation of A_{it} with c_i and ϵ_{it} . Next, we apply a first-difference estimator (FD). By looking at *changes* in log-occupational scores and changes in other covariates between the time of declaration and the time of petition, we purge individual heterogeneity (c_i) and all time-invariant characteristics out of the model. Therefore, we are able to assess the importance of factors such as family background and individual ambition as key drivers for name Americanization.²⁰

Due to the structure of our data, we slightly modify the pure first-difference model. A first-difference equation correctly measures the association between changing names and occupational scores if all migrants change their names between declaration and petition. As shown in Figure 3, in our dataset most individuals who Americanized their names had done already so by the time of declaration. A pure first-difference model would then consider as name keepers not only migrants who never Americanize their names (the “Keepers” in Figure 3) but also the “early Americanizers”. To address this issue and include all the migrants who have Americanized their names after arrival, we relate occupational changes between petition and declaration to the change in the Americanization index between petition and name at birth. Looking back at Figure 4, this estimation strategy examines whether changes in earnings of “Keepers” differs from changes in earnings of all other groups, other things being equal. Using name at birth is arguably exogenous to unobserved shocks that might

²⁰It should be noted that in the FD model, the inclusion of two-way interactions between years since migration and cohort, country of birth and labor market indicators still allows capturing assimilation patterns that are specific to these groups. The FD estimator also accounts for time-invariant factors that affect occupational trajectories such as birth place and socio-economic background at birth.

occur after the migrant’s arrival and provides a more conservative estimate of the relationship of interest if the payoff of name Americanization fades over time. We discuss the robustness of our results to this modification in Section V.

The next two empirical strategies tackle the potential bias that might arise from time-varying unobserved components (ϵ_{it}) being correlated with A_{it} .

Our third model focuses on the name changers only (NC). We exclude from the estimation sample migrants who do not Americanize their names throughout the period (“Keepers” in Figure 4). In other words, we compare migrants who already Americanized their name (“early Americanizers”) with those who subsequently did so between declaration and petition (“late Americanizers”) and those who changed into more distinctive names or changed into several names of different popularity twice (“Others”). Accordingly, we abstract from selection into name Americanization and compare the outcomes of migrants whose name popularity varies over time, with the objective of reducing the differences in outcomes stemming from time-varying unobservables.²¹ This strategy also serves two additional purposes. First, it highlights whether one’s name per se might induce further investments in U.S.-specific skills. If name Americanization induced acquisition of U.S.-specific skills, the group of “early Americanizers” should exhibit a higher skill level and higher earnings growth compared to the “late Americanizers”. In this case, a comparison of the two groups should show little or perhaps even a negative impact of name Americanization. Second, if name Americanization was associated with increased job search, we would expect “late Americanizers” to converge to the occupational standing of “early Americanizers”. If this is the case, as before, the comparison of the two groups should reveal little effect of name Americanization.

In a fourth specification, we report a first-difference model in which we control for name-at-birth specific time trends (NB). Hence, we compare the labor market trajectories of two individuals with the same name at birth – e.g., Giovanni – one who Americanizes his name and one who does not. These individuals should have very similar labor market trajectories in the U.S. and this specification should control for name-specific assimilation patterns.²²

The empirical models described above aim to identify many of the explanations that might

²¹A similar strategy is adopted by Arai and Thoursie (2009).

²²We have replicated Figure 4 for the subsample of observations for which we have at least two distinct individuals holding the same name at birth – the source of our identification in the NB model. The payoff of name Americanization persists and remains economically and statistically important for migrants who Americanize their names. T-test for the (log) occupational score being equal over time show the following absolute mean differences and standard errors: 0.015 (s.e. 0.011) for “Keepers”, 0.020 (s.e. 0.042) for “Others”, 0.034 (s.e. 0.022) for “Early Americanizers” and 0.196 (s.e. 0.086) for “Late Americanizers”. In addition, at the time of declaration, the average log-occupational score of “Keepers” is economically and statistically similar to that of the name Americanizers, with the absolute difference being 0.015 (s.e. 0.015). By contrast, at the time of petition, name Americanizers have occupational-based earnings that are 0.059 log points higher than “Keepers” (s.e. 0.013).

drive the positive association between name Americanization and earnings growth. If other motives remain – for instance, name Americanization occurs in response to occupational achievement – the estimated return will capture the payoff from name Americanization as well as the influence of these other motives.²³

IV. The Economic Payoff of Name Americanization

Table 3 shows the key results. For notational simplicity, we use A_{it} throughout to indicate the Americanization index irrespective of the estimator used. We begin by discussing our findings, and then provide additional evidence consistent with our main conclusions.

A. Main Results

We start by pooling the data and reporting OLS estimates in columns (a) and (b). Name Americanization is associated with approximately a 3% increase in occupational score, an estimate that remains stable after including additional controls. Being married and having a U.S.-born spouse are both associated with a 4% increase in occupational scores, whilst having children has a negative effect on occupational standing. However, having U.S.-born children positively influences occupational upgrading. Residents outside New York City do not exhibit any particular pattern of occupational change. Additional (unreported) estimates show that migrants who arrived prior to 1921 under perform the late-comers. This is unsurprising as the introduction of the quota system might have changed the selection patterns of migration at origin with arrivals in the early-1900s consisting primarily of low-skilled migrants. In column (b), we include indicators for nationality-specific trends to capture different labor market patterns across origin groups. We find that while our results become statistically weak and point estimates slightly decrease, the association between name Americanization and occupational upgrading persists. This indicates that the impact of local labor market conditions is moderate and does not influence our main result.

Next, we apply a first-difference estimator to equation (3) and show the results in columns (c) and (d) of Table 3. The positive association between name Americanization and occupational-based earnings persists. The estimates are larger and the effects statistically stronger compared to OLS. Despite not being very common, changing from a very distinctive name to the most popular American name had a payoff of about 12%. The last two rows of the

²³In Section V we report a test that relates *past* name Americanization with *future* outcomes. The results from this test are consistent with name Americanization inducing earnings growth. Furthermore, we also estimate a model where we instrument for changes in A_{it} . This empirical strategy and the associated estimation results are alluded to in Section V and described in greater detail in Appendix A. The instrumental variable estimates do not reverse our main finding.

Table 3
The Effect of Name Americanization on Log-Occupational Score

	OLS			First Differences			Name Changers Only			Name-at-Birth		
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)				
A_{it}	.035** (.017)	.031* (.017)	.098*** (.038)	.117*** (.039)	.223* (.118)	.239** (.114)	.188*** (.064)	.219*** (.066)				
Married	.041*** (.012)	.032*** (.012)	-.015 (.020)	-.019 (.021)	.013 (.041)	.013 (.043)	-.011 (.032)	-.019 (.032)				
Has U.S.-born Spouse	.041*** (.015)	.044*** (.016)	.062 (.040)	.070* (.041)	.031 (.060)	.024 (.061)	.070 (.059)	.081 (.061)				
Number of Children	-.012*** (.004)	-.008** (.004)	-.009 (.010)	-.008 (.010)	-.031 (.019)	-.036 (.024)	-.003 (.014)	-.002 (.015)				
Has U.S.-born Child(ren)	.035*** (.012)	.032*** (.012)	.030 (.019)	.034* (.020)	.020 (.033)	.040 (.036)	-.003 (.031)	.009 (.032)				
Resides outside N.Y. City	-.044* (.023)	-.020 (.028)	-.008 (.025)	-.019 (.089)	-.032 (.059)	.113 (.203)	.021 (.032)	-.074 (.149)				
R^2	0.01	0.03	0.01	0.03	0.01	0.06	0.36	0.38				
N	8166	8166	4083	4083	1538	1538	4083	4083				
Years since Migration (YSM)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes				
Arrival Prior 1921	Yes	Yes	-	-	-	-	-	-				
Arrival Prior 1921 \times YSM	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes				
Country of Birth	No	Yes	-	-	-	-	-	-				
Country of Birth \times YSM	No	Yes	No	Yes	No	Yes	No	Yes				
Labor Market	No	Yes	-	-	-	-	-	-				
Labor Market \times YSM	No	Yes	No	Yes	No	Yes	No	Yes				
Pred. Occ. Score whole sample	3.192	3.192	0.028	0.028	0.048	0.046	0.031	0.031				
Pred. Occ Score Americanizers	3.197	3.204	0.044	0.045	0.097	0.103	0.064	0.062				

NOTE.—Robust standard errors in parenthesis. For OLS, N refers to the number of individual \times period observations; for other models, N refers to the number of individuals. A_{it} = Americanization index, which varies between 0 (names with the lowest frequency) and 1 (names with the highest frequency). See text for explanation. In column (c) to (h), the model is estimated in first differences and A_{it} is interpreted as the difference in the Americanization index between petition and birth. Name Changers Only refers to a sample in which Keepers (see Figure 3) are dropped. Labor Market are dummy variables for each of the N.Y. City community districts. Married, Has U.S.-born Spouse, Has U.S.-born Child(ren), Moved into N.Y. City are all indicators. In column (c) to (h), arrival cohort, country of birth and labor market indicators are dropped when taking first-differences, although their interaction with YSM is identified.

* $p < .10$.
 ** $p < .05$.
 *** $p < .01$.

table report the average predicted gains, which are around 3% across the whole sample and around 5% among the name Americanizers. By contrast, changes in household characteristics have little impact on occupational upgrading, apart from a statistically weak effect of having a U.S.-born spouse and U.S.-born children. Unreported estimates for the “Country of Birth \times Years since Migration” interactions suggest low occupational convergence across nationalities, aside from the Russians, Irish and the “Other” category.

Columns (e) and (f) show the results for the specification in which we only include migrants who changed their names over time. Changes into the most popular American names are associated with a 24% increase in occupation-based earnings, with average payoffs being 5% across the whole sample and 10% in the subsample of name Americanizers.

Finally, we introduce name-at-birth fixed effects in columns (g) and (h). The payoff of name Americanization persists even after controlling for assimilation patterns within individuals who are from the same country of birth and hold the same name at birth. This specification indicates that returns to name Americanization can be as large as 22% for migrants named, e.g., Hans, Carmine or Johann who Americanized their names into, e.g., John, compared with migrants who retained their names. Someone named Francesco or Franz who chose to be named Frank would have gained about 9% more than someone who remained Francesco or Franz. Finally, migrants with Jewish names such as Moische who Americanized into Morris, would have gained about 2.6% more than those who kept their name at birth.

Figure 5A shows the average payoff for individuals who Americanized into the names listed. The full set of predicted earning changes as a function of name Americanization is shown in Figure 5B.

To put these magnitudes into perspective, we show that the estimated returns are within the range of existing occupational mobility from that time. For this purpose, we use the 1930 census to understand the age-earnings profile of males living in the U.S.. Figure 6 shows the average log-occupational score by age in 1930. The average payoff of name Americanization across the whole sample estimated using our last specification (3%) is comparable to the average occupational upgrade of individuals ageing from 26 to 30 years at that time. If we consider the subset of individuals who Americanize their names, the average return of 6% compares with the occupational upgrade occurring as individuals age from 24 to 26 years. Finally, the largest estimated effect of 22% – which refers to the subset of individuals changing from a very distinctive to the most popular American name – is close to the average occupational upgrade observed as individuals age from 20 to 24 years. Hence, the largest payoffs that we estimate are comparable to the existing labor market mobility patterns experienced by labor market entrants in the U.S. in 1930.

To summarize, by studying changes in occupational-based earnings over time, we ex-

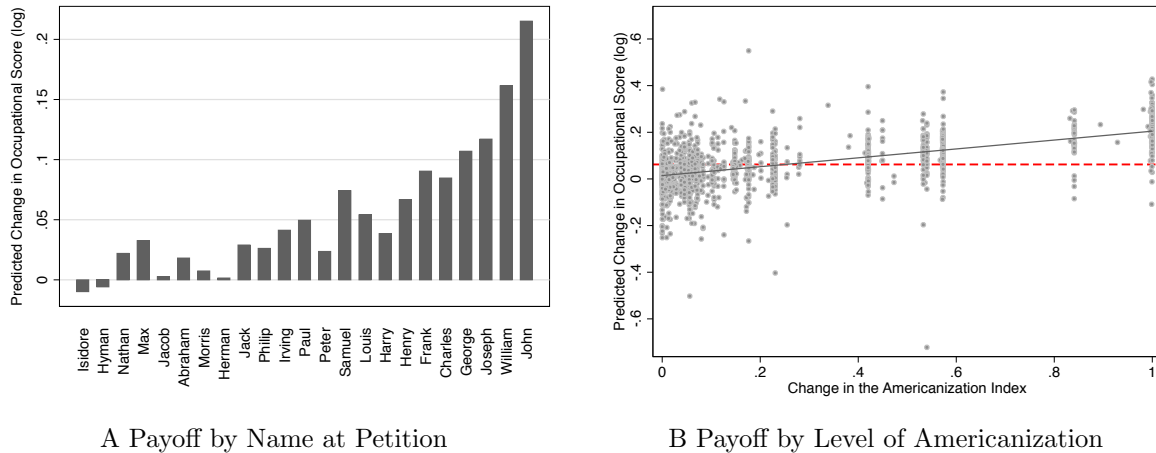


FIG. 5.—Predicted changes in log-occupational score, name-at-birth estimates. Predictions are based on the estimates in Table 3, column (h). Panel A shows the predicted changes in occupation-based earnings by name at petition, for the subset of name Americanizers. Panel B shows the scatter plot of predicted changes in occupation-based earnings as a function of the change in the Americanization index, for the subset of name Americanizers. Dotted line: average prediction in the sub-sample. Continuous line: fitted values.

cluded the possibility that the payoff to name Americanization is fully explained by family background and ambition. By exploiting the different timing of name Americanization, we also excluded that this payoff is fully driven by investments in skills or job search activities. Finally, by exploiting name-at-birth variation, we adjust for name-specific trajectories in the labor market. The persistent payoff that we find is therefore consistent with employers, co-workers, customers and neighbors having a taste for American names or employers using either prejudice or a form of statistical discrimination in their hiring decisions.

B. Heterogeneity of Relationship Between Name Americanization and Labor Market Success

Several factors contribute to make the estimated effects both credible and relevant. First, we find a persistent payoff to name Americanization across a battery of estimation strategies. Second, across all models we consistently find a downward bias in the OLS estimate of β_1 . In other words, migrants who faced the worst occupational trajectories found it beneficial to Americanize their names. This evidence is consistent with Goldin and Shim (2004), who find that renouncing maiden names is more common among low-skilled women.

Therefore, we claim that the economic payoff uncovered here masks assimilation patterns of individuals who are from lower socio-economic backgrounds, more discriminated against and facing stronger barriers to upward mobility. By contrast, high achieving individuals probably faced high costs of identity change or did not receive substantial benefits from it,

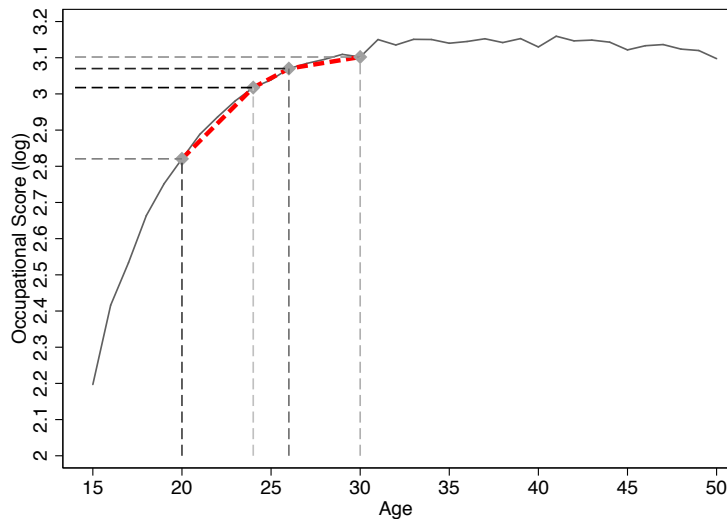


FIG. 6.—Average log-occupational score in 1930, by age. Source: 1930 IPUMS census. Plot of average log-occupational scores in the U.S., in 1930.

as they were subject to less discrimination or they could more easily access alternative means for socio-economic improvement. Consequently, we examine the hypothesis that groups that are “poorer” should experience a larger effect of name Americanization on occupation-based earnings growth.

Estimation results exploring this possible heterogeneity are given in Table 4. We first identify proxy measures for low socio-economic background. In the absence of direct measures of pre-migration income and literacy, we use two proxies: country of birth and height.

In the top panels, we distinguish between “old migrants” (i.e., who belong to nationalities whose migration to the U.S. had started in the mid-1800 century, e.g., Germans, Scandinavians, Irish, British) and the “new migrants” (i.e., who belong to nationalities whose migration had started during the late-1800s, e.g., Italians, Eastern Europeans, Russians). Migration from the latter nationalities was in fact perceived as an invasion of unskilled labor. In the bottom panels, we use height as a proxy for socio-economic background. This approach is inspired by a large body of literature establishing that the average stature of a group is related to many economic aspects such as skills, education, income, wealth, and health (see Komlos and Meermann, 2007 for a review of the introduction and use of anthropometric indicators in labor and development economics). We look at migrants whose height was below the average height of American-born individuals.²⁴

Table 4 shows that the effect is larger for new migrants and shorter migrants. By contrast,

²⁴The average height of American-born white males was 5’9” at that time (Costa, 2015).

the magnitude of the estimates is smaller and the effects statistically insignificant in most specifications for old migrants. Similarly, we cannot find statistically significant effects for taller migrants.

Overall, our analysis indicates that name Americanization had more pronounced effects for migrants belonging to sub-groups of the population for whom the barriers to labor market success were arguably higher. This could be driven by stronger discrimination in the labor market for these individuals or by their relatively low level of skills. Indeed, historical accounts confirm that there was widespread resentment and discrimination against these groups (see e.g., Higham, 2002).

Table 4
The Effect of Name Americanization on Log-Occupational Score, by Group

	Old Migrants				New Migrants			
	OLS	FD	NC	NB	OLS	FD	NC	NB
A_{it}	-.006 (.025)	.095* (.055)	.080 (.103)	.078 (.082)	.044* (.025)	.122*** (.047)	.273* (.160)	.233*** (.087)
N	2502	1252	231	1252	5660	2831	1307	2831
	Tall				Short			
	OLS	FD	NC	NB	OLS	FD	NC	NB
A_{it}	-.028 (.039)	.109 (.095)	.193 (.232)	.033 (.207)	.051*** (.020)	.115*** (.043)	.200* (.104)	.183** (.074)
N	1846	924	300	924	6116	3059	1202	3059

NOTE.—Robust standard errors in parenthesis. All models refer to the specification with all covariates in Table 3. OLS: pooled regression; FD: first-difference estimator; NC: name changers only; NB: name-at-birth fixed effects. For OLS N refers to the number of individual \times period observations; for other models, N refers to the number of individuals. A_{it} = Americanization index, which varies between 0 (names with the lowest frequency) and 1 (names with the highest frequency). See text for explanation. Old Migrants refers to migrants from Germany, Ireland, U.K. and Northern Europe. New Migrants refers to migrants from Italy, Russian Empire, Central and Southern Europe, Americas and Other. Tall refers to migrants with a height above 5 feet 9 inches, which corresponds to approximately the 3rd quartile of the height distribution. Height is reported only by 3,982 migrants.

* $p < .10$.
** $p < .05$.
*** $p < .01$.

V. Robustness checks

We conclude our analysis by carrying out a series of checks aiming to ascertain the robustness of our results with respect to reverse causality concerns, as well as numerous data aspects and definitions.

One possible criticism of our analysis is that since names and occupation-based earnings are observed at the same time, we cannot disentangle whether occupational changes are in response to name Americanization or the other way around. To rule out reverse causality, we report additional checks in the first four columns of Table 5. In the estimation sample, we only include “Keepers” and “early Americanizers” (see Figure 4); consequently, by construction the NC model is not estimated. We then estimate the following:

$$y_{it} = \beta_0 + \beta_1 A_{it-1} + \mathbf{x}'_{it} \gamma + c_i + \epsilon_{it}. \quad (3)$$

In these settings, the name-at-birth model identifies the different *future* earning trajectories (between declaration and petition) between a Giovanni who remains Giovanni and a Giovanni who Americanized his name into John (between arrival and declaration). Hence, while this specification allows for dynamic effects, it also assures that occupational changes are observed following name Americanization. If name Americanization was a consequence of occupational change (due to reverse causality), further future occupational upgrades are plausibly less likely to occur. The first four columns in the top panel of Table 5 show that the economic payoff of name Americanization persists in this model, suggesting that name Americanization was associated with future improvements.

To further reassure about the absence of reverse causality, in Appendix A we describe an instrumental variable strategy. We calculate the “Scrabble points” for each name at birth by summing the scores attributed to each letter in the popular board game and use these points to predict name Americanization. Scrabble points capture the structure of words, measuring both their length and how uncommon their letters are. Therefore, they provide a measure encapsulating the graphemic and phonemic features of names. Identification relies on the following exclusion restriction assumption: while name popularity (A_{it}) influences labor market outcomes – since names implicitly signal individuals’ socio-economic background (e.g., Bertrand and Mullainathan, 2004, Fryer and Levitt, 2004) – names’ linguistic structure does not have a direct impact in the labor market. This analysis, shown in Table A3 in Appendix A, does not reverse our main conclusions.

Our baseline model considers changes in name between time of arrival and petition to capture all the name changers in the sample. To check the consistency of our results, we estimated a “true first difference model” using name at declaration rather than name at birth. Note that by construction the NC model is the same as the one presented in Table 3, while in the FD and NB models the “early Americanizers” will now be considered as “Keepers”. The results remain strong and statistically significant once a pure first-difference model is estimated, as reported in the last four columns in the top panel of Table 5.

Table 5

The Effect of Name Americanization on Log-Occupational Score, Robustness Checks

	Early Americanizers and Keepers				Using Name-at-Declaration			
	OLS	FD	NC	NB	OLS	FD	NC	NB
A_{it}	.033*	.092**	–	.158**	.031*	.250**	.239**	.325**
	(.020)	(.038)	–	(.068)	(.017)	(.118)	(.114)	(.146)
N	7432	3716	–	3716	8166	4083	1538	4083
	Early Declarants				Occupational Upgrade			
	OLS	FD	NC	NB	OLS	FD	NC	NB
A_{it}	.014	.116*	.269*	.172	.032**	.091***	.171**	.116**
	(.025)	(.066)	(.152)	(.119)	(.013)	(.030)	(.069)	(.048)
N	3666	1833	548	1833	8166	4083	1538	4083
	NYSIIS				Surname Americanization			
	OLS	FD	NC	NB	OLS	FD	NC	NB
A_{it}	.020	.137***	.343**	.294***	.032*	.116***	.242**	.218***
	(.018)	(.052)	(.157)	(.083)	(.018)	(.039)	(.115)	(.067)
N	8166	4083	1538	4083	8166	4083	1538	4083

NOTE.—Robust standard errors in parenthesis. All models refer to the specification with all covariates in Table 3. OLS: pooled regression; FD: first-difference estimator; NC: name changers only; NB: name-at-birth fixed effects. For OLS N refers to the number of individual \times period observations; for other models, N refers to the number of individuals. A_{it} = Americanization index, which varies between 0 (names with the lowest frequency) and 1 (names with the highest frequency). See text for explanation. Early Americanizers and Keepers refers to a subsample of migrants who did not Americanize their names or Americanized their names by the time of declaration. Note, the NC specification cannot be identified for the Early Americanizers and Keepers regression. Using Name-at-Declaration refers to a specification in which name Americanization is defined between petition and declaration. Early Declarants refers to a subsample of migrants who filed the declaration of intention within three years since arriving in the U.S.. Occupational Upgrade refers to regressions where the dependent variable is an indicator that equals one if occupation at petition is ranked higher than at declaration. We categorized occupations into standard broad categories (professional, technical, and kindred; managers, officials, and proprietors; farmers; clerical and sales; craftsmen; operatives; service workers and laborers, including farm laborers). We then define occupational upgrading as any occupational change into a higher skilled job. NYSIIS refers to an Americanization index where all names (of the U.S.-born population, and of migrants at birth, declaration and petition) have been standardized using the New York State Identification and Intelligence System Phonetic Code (NYSIIS). Surname Americanization refers to a model where we interact the A_i index with an indicator for surname Americanization. Surname Americanization is defined as the custom of adopting a surname that was more popular in the U.S.-born population than the original surname.

* $p < .10$.** $p < .05$.*** $p < .01$.

As a further check, we restrict our sample to individuals who declare their intention to naturalize within three years of arrival and exclude everybody else (first four columns in the second panel of Table 5). Accordingly, we limit the effects of new names to a shorter time span and focus on individuals who change name and file declaration papers at early

stages. This specification serves two main purposes: first, it better captures the trajectories immediately after arrival, presumably working with a more exogenous initial condition for the occupational distribution; second, it is plausibly less likely that name Americanization has occurred for this group as a consequence of having gained substantial labor market experience. The estimates are consistent with baseline results, albeit they are somewhat statistically imprecise, due to the smaller sample size.

Next, we implement further checks to ascertain the sensitivity of our results to the definition of the occupational score. Instead of relying on a measure of “earning potential” (such as the occupation score), we categorize occupations into standard broad categories (professional, technical, and kindred; managers, officials, and proprietors; farmers; clerical and sales; craftsmen; operatives; service workers and laborers, including farm laborers). We then define our dependent variable as occupational upgrading, i.e., any occupational change into a higher-skilled job between declaration and petition. This definition is quite restrictive: to find an effect, changing one’s name must push an individual into a different broad occupational group. The last four columns of the second panel of Table 5 show that our results persist and we continue to find that name Americanization is associated with a higher likelihood of occupational upgrading.

One potential problem with our approach so far is that, while we treat any spelling of a name as a separate name, some names simply have multiple spellings or may have been misspelled on one of the documents we use. We therefore implement a robustness check where name Americanization is defined as the change into a phonetically different name, i.e., the change in the “sound” of the name. This definition prevents that potential misspellings are interpreted as name Americanization. To this end, we use the New York State Identification and Intelligence System Phonetic Code (NYSIIS) algorithm, whereby names that sound the same yet are spelled differently such as John and Jon have the same Americanization index. Therefore, we are able to purge possible misspelling errors made in the original record by the court clerks. As can be seen in the first four columns of the third panel of Table 5, even when this restrictive definition is adopted the results remain similar to our baseline analysis.

In the last check we control for surname Americanization. We define surname Americanization along the lines of name Americanization. We then interact the A_i index with an indicator for surname Americanization. The results in the last four columns of the third panel of Table 5 show that the inclusion of the interaction term has no effect on our main estimates and there are no additional returns associated with surname Americanization.²⁵

²⁵Appendix B shows the robustness of our results when excluding imputed occupations, excluding imputed addresses, changing the reference name distribution to the U.S.-born Americans 50+, and using an index on the same lines as that of Fryer and Levitt (2004). The reader might wonder whether some migrants change name for pragmatic reasons. In unreported analyses we show that our results hold across all migrants,

VI. Conclusions

This paper provides the first direct evidence on the magnitude and economic consequences of name Americanization. Previously only known anecdotally, this phenomenon was widespread and had a substantial impact on upward mobility during the first half of the 20th century. Our analysis reveals that immigrants who Americanized their names experienced substantial occupation-based earnings growth. These results persist across numerous specifications, including controlling for individual heterogeneity and labor market trajectories of individuals holding the same name at birth. The effects are largely driven by migrants facing pre-existing constraints to occupational mobility. For these migrants, name Americanization was effective in unleashing economic success.

Our finding provides important insights for current research. First, it is relevant for historical studies that use record linkage to obtain longitudinal data. The results indicate that the representativeness of linked samples could be improved by taking into account that a large fraction of naturalized migrants Americanized their names. Second, our results highlight the existence of a trade-off between maintaining individual identity and enhancing labor market outcomes, suggesting that cultural assimilation was instrumental in economic assimilation. From a historical perspective, this implies that migrants adopted alternative strategies to climb up the occupational ladder despite facing barriers to occupational upgrading (Abramitzky et al., 2014). As one example of such a strategy, we find that returns to name Americanization were quite high. While OLS estimates suggest that the association between name Americanization and earnings might be attributed to changes in the market valuation of migrants' skill endowment, a reduction in discrimination or more rapid human capital accumulation, more elaborate identification strategies rule out many of these reasons, with the exception of discrimination. The consequences of this are twofold: first, low occupational mobility observed in previous studies might have been caused by different attitudes and discrimination levels towards certain ethnic groups; and second, from a broader perspective, the implied trade-off between individual identity and labor market success suggested in several recent analyses (e.g., Bertrand and Mullainathan, 2004, Fryer and Levitt, 2004, Arai and Thoursie, 2009) seems to have been present since the early making of modern America.

including those whose name might be easier to change for purposes such as spelling. To this end, we assigned to each name at birth an indicator which takes the value one if there exists a name in the English language that is etymologically equivalent – i.e., an English cognate – to the migrant's name at birth. We constructed an indicator using data on equivalent given names available from the 2006 Oxford Dictionary of First Names (Hanks et al., 2006). We estimate baseline regressions controlling for the existence of a cognate and interacting this indicator with the Americanization index, finding no heterogeneous effects across groups.

References

- Abramitzky, Ran, Leah Boustan, and Katherine Eriksson. 2012. Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review* 102, no. 5:1832–1856.
- . 2014. A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy* 122, no. 3:467–506.
- 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.
- Bertrand, Marianne and Sendhil Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review* 94, no. 4:991–1013.
- Broom, Leonard, Helen P Beem, and Virginia Harris. 1955. Characteristics of 1,107 petitioners for change of name. *American Sociological Review* 20, no. 1:33–39.
- Cannato, Vincent J. 2009. *American passage: The history of Ellis Island*. Harper Collins.
- Costa, Dora. 2015. Health and the economy in the United States, from 1750 to the present. *Journal of Economic Literature* 53, no. 3:503–570.
- Drury, Darrel W and John D McCarthy. 1980. The social psychology of name change: Reflections on a serendipitous discovery. *Social Psychology Quarterly* :310–320.
- Fermaglich, Kirsten. 2015. "Too long, too foreign... too Jewish": Jews, name changing, and family mobility in New York City, 1917-1942. *Journal of American Ethnic History* 34, no. 3:34–57.
- Ferrie, Joseph P. 1996. A new sample of males linked from the public use microdata sample of the 1850 US Federal census of population to the 1860 US Federal census manuscript schedules. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 29, no. 4:141–156.

- Figlio, David N. 2005. Names, expectations and the black-white test score gap. *NBER Working Paper 11195* .
- Fryer, Roland G. and Steven D. Levitt. 2004. The causes and consequences of distinctively black names. *The Quarterly Journal of Economics* 119, no. 3:767–805.
- Goldin, Claudia and Maria Shim. 2004. Making a name: Women’s surnames at marriage and beyond. *The Journal of Economic Perspectives* 18, no. 2:143–160.
- Hanks, Patrick, Hardcastle Kate, and Hodges Flavia. 2006. *A dictionary of first names*. Oxford University Press.
- Higham, John. 2002. *Strangers in the land: Patterns of American nativism, 1860-1925*. Rutgers University Press.
- Kang, Tai S. 1971. Name change and acculturation: Chinese students on an American campus. *The Pacific Sociological Review* 14, no. 4:403–412.
- Komlos, John and Lukas Meermann. 2007. The introduction of anthropometrics into development and economics. *Historical Social Research/Historische Sozialforschung* :260–270.
- Laham, Simon M, Peter Koval, and Adam L Alter. 2012. The name-pronunciation effect: Why people like Mr. Smith more than Mr. Colquhoun. *Journal of Experimental Social Psychology* 48, no. 3:752–756.
- Lieberson, Stanley. 2000. *A matter of taste: How names, fashions, and culture change*. Yale University Press.
- Lleras-Muney, Adriana and Allison Shertzer. 2015. Did the Americanization movement succeed? An evaluation of the effect of English-only and compulsory schools laws on immigrants’ education. *American Economic Journal: Economic Policy* 7, no. 3:258–290.
- Olivetti, Claudia and M. Daniele Paserman. 2015. In the name of the son (and the daughter): Intergenerational mobility in the United States, 1850-1940. *American Economic Review* 105, no. 8:2695–2724.
- 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] .