Economic and Cultural Assimilation in the Age of Mass Migration

Economics 970: The Economics of Immigration

Crystal Xue

Using a linked representative census data sample from 1880 and 1990, I examine the relationship between economic assimilation and cultural assimilation for immigrants in the United States during the Age of Mass Migration. I do not evidence of such a relationship. However, I find evidence of a weak positive association between lower levels of cultural assimilation and high occupational scores in 1880. In addition, I find evidence of a strong positive association between having a native spouse and high occupational scores in both 1880 and 1900. While it appears that there was not a meaningful relationship between economic and cultural assimilation in the Age of Mass Migration, cultural assimilation should still be valued for its possible impact on other quality of life indicators and policymaking decisions.

I. Introduction

Are economic assimilation and cultural assimilation of immigrants in the United States related? Immigrants to the United States generally face two categories of pressures: those related to economic assimilation and those related to cultural assimilation. In terms of economic assimilation, the goal is clear: earn money at a comparable level to similarly-skilled natives. Doing so allows immigrants to increase utility and provide better opportunities for their families. But in terms of cultural assimilation, the goal is less clear; is the objective to become fully "American"? To maintain the home country's culture? To find a balance that includes both? And even when a goal is ascertained, why should immigrants care? Perhaps one reason they should care is that cultural assimilation is related to economic assimilation. Yet does this relationship even exist? In this paper, I will examine the relationship between economic and cultural assimilation during the Age of Mass Migration (late 19th century through early 20th century) using the change in occupational rank of an immigrant and the change in foreignness index derived from the name an immigrant gives his child.

Because of data constraints, this paper will use rankings and indices derived from prior literature. Although the occupational rankings used in this paper were provided by my data source and represent the median income of an individual with that occupation, my idea of using occupational ranking as a measure of economic assimilation came from Collins and Zimran (2019). For both Collins and Zimran (2019) and this paper, it was important to use occupational ranking as there is no individual wage data in censuses conducted in the Age of Mass Migration. Collins and Zimran (2019) measured social mobility without income data by classifying jobs into six categories based on skill level. Based on these ranks, Collins and Zimran were then able to

track social mobility. The paper found that natives and immigrants who started off at the same occupational rank experienced different rates of social mobility; the natives experienced more. However, because there were more immigrants who started off in these socially mobile positions, it appeared that immigrants in the first wave of migration were more mobile than those in the second wave.

To gauge cultural assimilation, I adopted the name foreignness index that Abramitzky et al (forthcoming) uses for the same purpose. Abramitzky et al. (forthcoming) seeks to compare cultural assimilation during the Age of Mass Migration and the late 20th century. In order to so, the authors measure an immigrant mother's cultural assimilation by the change in how foreign the name she gives her children are. Their index is the ratio between the immigrant population share with the name and the sum of the immigrant population share with the name and the native share with the name. Abramitzky et al. (forthcoming) finds that cultural assimilation occurs more slowly today than in the Age of Mass Migration.

Other papers are notable for similar research questions. Goldstein and Steklov (2016) measures the relationship between cultural and economic assimilation, albeit for the "child", not the parent. The paper uses a similar name index and their measure for occupational achievement is also similar to Collins and Zimran (2019). Goldstein and Steklov (2016) find differing results for the name of different geographical origins. The positive effects, which hold for children of immigrants from Ireland, Italy, German, and Poland, are attributed to both social class difference in name-giving (i.e. immigrants with higher social class give their children more "American" names) and signaling (i.e. "American" names indicate that the first-generation American children come from higher social class). Similar to Goldstein and Steklov (2016), Abramitzky et al.

(2020) studies the impact of perceived cultural assimilation (based on how foreign a name is) and labor market outcomes. Using the same foreignness index found in both Abramitzky et al. (forthcoming) and this paper, the authors find "children of immigrants [in the early twentieth century] who were given more foreign first names completed fewer years of schooling, earned less, and married less assimilated spouses". However, rather than positing that "the negative association between ethnic names and adult outcomes", the authors attributed the difference in outcomes to "household differences associated with cultural assimilation". The authors came to this conclusion because they found "few differences in the adult outcomes of brothers who were given more foreign versus more American-sounding first names". Although this paper, Goldstein and Steklov (2016), and Abramitzky et al. (2020) will use very similar indices to measure cultural assimilation, their target populations are different. Rather than looking at "static" cultural assimilation of immigrants' children as Goldstein and Steklov (2016) and Abramitzky et al. (2020) did, I will be looking at the parents' "dynamic" cultural assimilation.

Reitz and Sklar (1997) also considers the relationship between cultural and economic assimilation. They find that "foreignness" has the largest effect on economic outcomes for select European immigration populations but more broadly for racial minority immigration populations. Reitz and Sklar believe that these impacts are "presumably based on skin color, regardless of a specific culture, identity, behaviors, or network affiliations". Similar to my paper, they separate their subject based on the country of origin. My work (1997) will differ from Reitz and Sklar in the population it studies. Reitz and Sklar (1997) use surveys conducted on Toronto-area residents in 1979; I will use data from the 1880 and 1900 United States Census. As a result, while Reitz and Sklar (1997) examine the economic attainment of modern Canadian

migration, this paper will examine the economic attainment of American migration during the Age of Mass Migration.

By connecting cultural and economic assimilation for immigrants to the United States during the Age of Mass Migration, this paper will contribute to the existing literature on the subject. To that end, Section II will discuss the empirical methods used in this paper, Section III will discuss the data source, Section IV will present the main findings and examine the results for robustness, and Section V will provide conclusions and discussions for future research.

II. Empirical Methods

In order to determine the relationship between economic and cultural assimilation, I will be using the 1880-1900 IPUMS Linked Representative Samples, created by the Minnesota Population Center (2010). These 1% samples (meaning that they are 1% of all linked census entries between the two years) include characteristics such as name, birthplace, relation to household head (i.e. spouse, child, etc.), and occupational rank. IPUMS provides linked representatives for a variety of time intervals, however, all samples are based on the 1880 complete count dataset. I chose the 1880-1900 sample because I believe it is a large enough time frame to "catch" the parents having children while ensuring they are not yet retired (so they still have occupational scores). IPUMS also provides separate datasets for women, men, and married couples. For the purposes of this paper, I will be using married couples as it will allow me to see if there is any relationship between economic assimilation and having an American spouse, suggesting that there was actually less cultural assimilation in the immigrant parent.

The main variables I am interested in are name, occupational score, relationship to household head, birthplace, and age. For the main independent variable, I need the name field to

calculate the foreignness index of the names of the immigrant parent and their children. In order to calculate the foreignness index of a name, I will take the distribution of names in 1880 (as these are the names the parents have been exposed to) and calculate the proportion of the migrant population versus the total population with that name. This index is the same one used in Abramitzky (forthcoming). It is written as:

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

I will then calculate the change in cultural assimilation of the parent using two benchmarks. First, I will measure the change in the foreignness index between the parent's name and their last child's name. I will also measure the change in the foreignness index between the first child's name and the last child's name. If the parent only has one child between 1880 and 1900, then I will remove that individual from my regression. This is to check the robustness of my results. For the dependent variable, economic assimilation, I will use the occupational score provided in the dataset. The occupational scores provided in this dataset reflect the median nominal income of the occupation in 1950. As such, any change in occupational score can be converted to an increase in income (and therefore, improved labor market outcomes). Given that change in occupational score is my dependent variable, it is very important that I have values for both 1880 and 1900 and as such, will be discarding any observations with incomplete data.

The other variables of interest will allow controlling for foreseeable sources of variance, thereby helping me isolate the true relationship between economic and cultural assimilation. The relationship to household head variable will ensure that I am accurately capturing the parent, the spouse, and the children (rather than mistaking a cousin or a grandchild for a child). I will use the birthplace to determine which individuals are immigrants and build interaction terms and

indicator variables. I believe that there may be a muted effect for immigrants arriving from English-speaking countries and as such, believe that a regression without interaction and indicator variables for this trait would underestimate the impact of cultural assimilation on economical assimilation. I also believe the child of a native and an immigrant may have a more culturally assimilated name than the child of two immigrants. As such, I will also control for the birthplace-status of the immigrant's spouse. Finally, the age variable will be used to ensure that first, the children used to measure the change in foreignness index were born in 1880-1990, our time range of interest. Although I would have liked to control for age arrived in the United States, as younger immigrants may assimilate faster, I did not have the necessary data.

The strengths and limitations in my dataset are both due to its representativeness. More specifically, because these entries were taken from complete census counts, I have access to a huge swath of variables for a representative slice of individuals. However, the dataset is only representative by demographics, not necessarily by name representation. Considering that this dataset only includes 1% of all total entries, there is a high probability that not all names are present and as such, will yield 0 foreignness indices when the indices are actually slightly higher. Another limitation of this dataset is its incompleteness. There are a large number of entries where there are either entries for 1880 or 1900 but not both. Because I seek to measure change in this time interval, I cannot use those observations. The birthplace values are also sometimes empty, meaning that I have no way to detect if they are an immigrant or a native. Some names are either empty or misspelled, leading to additional challenges when calculating the foreignness index. Finally, because the occupational rank at the time a child was born is likely not recorded (considering that 1880 and 1900 are only two years out of a twenty-year span), it is impossible to

track cultural assimilation trajectory within the two years, merely the absolute change across the period.

Using the aforementioned variables, I will perform three types of regressions: a level analysis, a general differences regressions, and a country-specific differences regression. Given the difficulties of conducting censuses in 1880 and 1900, the dataset is incomplete and contains many missing values. Compounded with very stringent requirements for usage, the number of observations eligible for my regressions shrinks considerably. To combat this problem, I will conduct a level analysis that regresses the occupational rank of a parent on the foreignness index of their last child. I will conduct this regression using both the occupation rank of the parent in 1880 and 1900. The regression will take the form:

$$Occupational\ Rank_{i,t} = \alpha + \beta \Delta Foreignness\ Index_{last\ child,i} + \chi' \gamma_i + \varepsilon_i$$

Where i represents each parent, t is the year the census was conducted (i.e. either 1880 or 1900), α is the intercept, β is the coefficient of the change of foreignness index, $\chi' \gamma$ is a vector of interaction and indicator variables (such as whether they immigrated from an English-speaking country), and ε is the error term. The strength of this analysis is that it will allow a larger dataset. The limitation is that it does not measure the change in assimilation, simply assimilation levels. My subsequent analysis will address this limitation.

My general differences regression will calculate the magnitude of the relationship between change in occupational rank and change in foreignness index. I will use the general form:

$$\Delta\ Occupational\ Rank_i = \alpha + \beta \Delta Foreignness\ Index_i + \chi' \gamma_i + \varepsilon_i$$

Where i represents each parent, α is the intercept, β is the coefficient of the change of foreignness index, γ is a vector of interaction and indicator variables (such as the nationality of their spouse), and ε is the error term. The strength of my empirical strategy is that it will allow me to quantify both changes in economic and cultural assimilation, the major concern of the first regression. The limitation of my empirical strategy is that it does not incorporate weighting, meaning that some individuals could be overrepresented in the share of the total population with their demographics. In an ideal world, I would incorporate weighting into my model to account for this potential bias.

The country-specific differences regression will calculate the magnitude of the relationship between the change in occupational rank and the change in foreignness index for immigrants from the most common origin countries, Germany, England, and Canada. I will use the general form:

$$\Delta Occupational\ Rank_{c,i} = \alpha + \beta \Delta Foreignness\ Index_{c,i} + \varepsilon_i$$

Where c represents each country, α is the intercept, β is the coefficient of the change of foreignness index, and ε is the error term. The strength of this analysis is that it will allow me to narrow my focus to specific countries. This is particularly useful when we want to observe differences between countries. The main limitation of this analysis is that there is a potential for bias if the average values are not representative of the entire dataset (e.g. if there are outliers, the dataset skews in either direction, etc.).

III. Data

The cleaned dataset used for this paper consisted of 1101 immigrants. Among these immigrants, 2 of them were women and 1099 of them were men. The immigrants represented 22

countries with Germany, England, and Canada being the most common; there were 500 immigrants born in Germany, 154 immigrants born in England, and 119 immigrants born in Canada. In 1880, the mean age of the immigrants was 34.26 years. Given the median age was 34 years, the dataset is roughly symmetrical in relation to age. The average age of migration was 17.88 years, Considering that the median age of migration was 19 years, we can conclude that the dataset is skewed to the left. In 1880, the average number of years an immigrant had spent in the United States was 16.38 years. The dataset is skewed left considering the median value is 13 years. The mean occupational score in 1880 was 22.37 and the median occupational score was 23. This suggests that the dataset is roughly symmetrical. The dataset is slightly skewed right for the occupational score in 1900. It is interesting to note that although the mean occupation score in 1900 is higher than the occupational score of 1880 (23.25 vs. 22.37), the median occupational score in 1900 is lower than the median occupation score of 1880 (20 vs. 23). This suggests that either part of the population shifted to lower-paying jobs between 1880 and 1900 or the median income for an average income occupation decreased between 1880 and 1900.

IV. Results

Our analysis does not show a statistically significant relationship between cultural and economic assimilation. These results can be viewed in Table 2 and Table 3 of the appendix. However, there appears to be a statistically significant relationship between an immigrant parent's last child's name foreignness index and their occupational score in 1880. I found these results during my levels analysis. The results can be viewed in Table 1. Below, I discuss my results in greater detail and provide possible interpretations and reasoning.

My first type of analysis, my levels analysis, sought to answer the question: Are immigrants who give their kids more foreign names more or less professionally successful in 1880? In 1900? Controlling for the immigration status of their spouses, their native languages, and their age when they migrated, I found a very small positive association between the last child's foreignness index and an immigrant's 1880 occupational score that was statistically significant at the 90% confidence level. My results suggest that for every 1% increase in the probability that an immigrant would have that name, we expect an increase of 0.00037 in the immigrant parent's occupational score, or $3.70 (1950 US dollars). This association may occur because immigrants in lower-paying jobs may experience more discrimination than those in higher-paying jobs because there are fewer concrete differentiating factors, such as education. As such, immigrants in lower-paying jobs may value the "protection" of an American-sounding name more than those in higher-paying jobs.

The native spouse indicator variable and interaction terms were also statistically significant in both 1880 and 1900, though in opposite directions. The native spouse indicator variable was 3.4 in 1880 and 2.979 in 1900. Both of these results were statistically significant at the 99% confidence level. This suggests that if an immigrant were to have a native spouse in 1880, we would expect the median income of his occupation to be $3,400 higher than if he had an immigrant spouse. Likewise, we would expect an immigrant with a native spouse to be in an occupation with a median income $2,979 higher than an immigrant with an immigrant spouse, holding all else equal. The negative direction of the interaction terms in both years (-0.042 in 1880 and -0.050 in 1900) suggest that for immigrants with native spouses, there is a negative association that is statistically significant at the 95% level between their last child's foreignness

index and their occupational score (i.e. the more foreign the child's name is, the less money we expect them to make). There is a myriad of reasons behind these relationships; one possibility is related to increasing concrete differentiating factors in higher-paying jobs.

In my general differences regressions, where I regressed change in occupational score for each immigrant on change in foreignness, I did not find a statistically significant relationship between the two types of assimilation. This result was in opposition to my hypothesis, as I previously an increase in cultural assimilation to correlate with an increase in economic assimilation. In order to test the robustness of my results, I used both the change between the parent's and last child's names and the change between the first and last child's name, ignoring families who only had one child. In both regressions, the relationship was not statistically significant. I also included controlled for whether the spouse was native, whether the immigrant spoke English as a relationship, and when the immigrant arrived in the United States. With or without these controls, my results remained not statistically significant. As such, I cannot conclude that there is a statistically significant relationship between economic and cultural assimilation.

There are a few possibilities as to why my results were not statistically significant. First, names may not be a good way to measure cultural assimilation. Rather than reflecting the culture that surrounds them, names are often used to commemorate loved ones, emphasize values important to a family, etc. Second, because I did have access to individual wage data, my regression may have missed wage growth (or decrease) with an occupation. For example, if an immigrant was a carpenter in both 1880 and 1900, but had risen to the top of his craft in 1900 (and subsequently, earned more), my regression would not have noted that change.

Given there were several countries that represented a large portion of my dataset (e.g. over 45% of the immigrants in my dataset were born in Germany), I decided to run country-specific regressions for the three most common origin countries: Germany, England, and Canada. I expected to see a greater association between economic and cultural assimilation as large populations allow for the development of enclaves, where cultural assimilation is less rapid than elsewhere. Assuming that economic assimilation proceeded at an equal pace to immigrants with other origin countries, each incremental increase in cultural assimilation would have lead to a larger increase in economic assimilation. However, I did not find this result; there was no statistically significant relationship between change in occupational rank and foreignness index for Germany, England, and Canada.

V. Conclusion

In this paper, I did not find a statistically significant relationship between economic assimilation (measured by a change in occupational score) and cultural assimilation (measured by a change in foreignness index). However, based on the result of the levels analysis, there is a weak positive association between the foreignness of names and higher-paying occupations. Moreover, there is a very strong positive association between having a native spouse and higher-paying occupations. Finally, having a native spouse appears to negatively impact the weak positive association between the foreignness of names and higher-paying occupations.

At the beginning of this paper, we asked ourselves, “Why should we care about cultural assimilation?” And although there does not appear to be a relationship between economic and cultural assimilation, there are still reasons why cultural assimilation matters. Perhaps even if immigrants fully economically assimilate, a lack of cultural assimilation culturally leads to

ostracization by natives. This could lead to a decrease in the immigrant's quality of life even if labor market outcomes are identical. Another possibility is related to the propagation of "American" values; if policymakers value common belief in "American" ideals such as democracy and liberty and believe that immigrants do not hold those beliefs, then they might still place an emphasis on cultural assimilation.

References

Abramitzky, R., Boustan, L., Eriksson, K., & Hao, S. (2020). Discrimination and the Returns to Cultural Assimilation in the Age of Mass Migration. *AEA Papers and Proceedings, 110*, 340.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. (Forthcoming). "Do Immigrants Assimilate More Slowly Today than in the Past?" American Economic Review: Insights.

Collins, William J. and Ariell Zimran. 2019. "Immigrants' Changing Labor Market Assimilation in the United States during the Age of Mass Migration." NBER Working Paper #26414.

Goldstein, J., & Stecklov, G. (2016). From Patrick to John F: Ethnic Names and Occupational Success in the Last Era of Mass Migration. *American Sociological Review, 81*(1), 85-106.

Minnesota Population Center. (2010). 1880-1900 IPUMS Linked Representative Samples, Couples (final version). Retrieved from https://usa.ipums.org/usa/linked_data_samples.shtml

Reitz, J., & Sklar, G. (1997). Culture, Race, and the Economic Assimilation of Immigrants. *Sociological Forum, 12*(2), 233-277.

Appendix

Table 1: Levels Analysis Using 1880 and 1900 Occupational Data

	Dependent variable:	
	1880 Occupational Score	1900 Occupational Score
	(1)	(2)
Last Child's Foreignness Index	0.037*	0.022
	(0.021)	(0.023)
Native Spouse (indicator)	3.400***	2.979**
	(1.250)	(1.412)
Native English Speaker (indicator)	−0.674	−0.254
	(1.237)	(1.398)
Age Migrated (indicator)	0.031	0.012
	(0.044)	(0.050)
Native Spouse (interaction)	−0.042**	−0.050**
	(0.019)	(0.022)
Native English Speaker (interaction)	0.014	0.007
	(0.020)	(0.022)
Age Migrated (interaction)	−0.0004	−0.0003
	(0.001)	(0.001)
Constant	19.572***	22.128***
	(1.355)	(1.531)
Observations	982	982
F Statistic (df = 7; 974)	1.398	1.100
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2: Differences Regression

	Dependent variable:			
	Change in Occupational Score			
	(1)	(2)	(3)	(4)
Between Parent and Last Child	−0.008	0.006		
	(0.009)	(0.019)		
Between First and Last Child			−0.009	−0.020
			(0.009)	(0.026)
Constant	0.687**	3.091***	0.804**	3.858***
	(0.336)	(0.981)	(0.382)	(1.086)
Native Spouse Controls	No	Yes	No	Yes
Native English Speaker Controls	No	Yes	No	Yes
Age Migrated Controls	No	Yes	No	Yes
Observations	879	879	571	571
Adjusted R^2	−0.0001	0.005	−0.00001	0.014
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3: Country-Specific Differences Regression

	Dependent variable:		
	Change in Occupational Score		
	Germany	England	Canada
	(1)	(2)	(3)
Between Parent and Last Child	−0.002	−0.025	−0.043
	(0.027)	(0.063)	(0.049)
Constant	2.637*	3.959	3.766
	(1.570)	(2.444)	(2.381)
Native Spouse Controls	Yes	Yes	Yes
Age Migrated Controls	Yes	Yes	Yes
Observations	407	127	94
Adjusted R^2	−0.002	−0.019	−0.003
Note:			*p<0.1; **p<0.05; ***p<0.01