

Brains versus Brawn: The Persistence and Convergence of Premigration Socioeconomic Status during the Age of Mass Migration

By

Peter Catron

Department of Sociology
University of Washington
catron@uw.edu

2019 PAA Submission

Extended Abstract

Short Summary: Immigrant assimilation studies often point to the success of immigrants who entered in the first half of the twentieth century to draw inferences about whether today's immigrants will follow a similar trajectory. However, there is disagreement over how long premigration differences take to disappear. Whereas some researchers suggest that convergence of premigration differences take up to 150 years (Borjas 1994), others argue that convergence happens over a few generations (Perlmann 2006; Water 1990; Alba and Nee 2003). This research, however, has been unable to link immigrants to their descendants and instead rely on pseudo-cohort analyses. Similarly, they have been unable to track immigrants from their sending country to the US. To adjudicate the speed of premigration convergence, therefore, I build a new panel dataset that links immigrants from their passenger record to their census record and then follows their children across time. The results suggest that while premigration differences matter when predicting the economic outcomes of the first generation, their effects are severely attenuated by the second and third generations. The data that links parents and children points to faster convergence than predicted by those arguing in favor of long-term differences.

Brains versus Brawn: The Persistence and Convergence of Premigration Socioeconomic Status during the Age of Mass Migration

This article seeks to understand how long premigration socioeconomic status differences persist across generations. It focuses on four immigrant groups – Italian, Jewish, German, and Russian immigrants – who entered in the first half of the twentieth century. A long-standing debate is concerned over how long initial differences in socioeconomic status persisted for these immigrant groups. On the one hand, researchers note that differences persist for over 150 years (Borjas 1994), but on the other hand, others argue that differences disappeared after just three generations (Alba and Nee 2003; Waters and Lieberson 1989). This previous research, however, has been unable to directly link immigrants to their descendants and instead rely on pseudo-cohort analyses. I enter this debate by investigating how long differences in premigration socioeconomic status persisted across generations by developing a new panel dataset that follows immigrants from their sending country to the US and then followed their US-born children over time.

This research is of particular interest to scholars of migration and stratification. Immigrants who come from higher status backgrounds are expected to perform better in the labor market than groups with lower status backgrounds. This belief leads to calls for merit-based immigration policies that favor high skilled immigrants that has been a popular sounding board of the Trump administration. The results from this article help us understand whether migration restrictions or selection policies are necessary to ensure strong migrants' performance in the labor market in a period of open borders. In addition, this article is of direct relevance to today's immigrants who enter with a wider range of skill than in the past. While data that track today's immigrants across generations are limited, slow or rapid convergence of premigration status in the past may point to how long we may expect differences to remain today.

Data and Matching

The data used in this article come from immigrant passenger records and the full-count 1910 and 1940 US censuses. Passenger records keep information about individuals coming to the US including their name, age, sex, literacy, occupation in the sending country, place of last residence, destination, country of origin, and travel compartment. I include all immigrants who arrived to the US from Germany, Italy, and Russia between 1880 and 1899 and match them forward to their 1910 census record. Then, I take their US-born children in 1910 and match them forward their 1940 record. These linked datasets follow immigrants from their sending country to settlement in the US and then follows their children to adulthood.

The linking procedure follows standard algorithms that match individuals using their name, age, and place of birth (Abramitzky, Boustan, and Eriksson 2012). This technique links individuals from their passenger record to their 1910 census record by first standardizing first and last names by correcting for nicknames (e.g., "Pete" v. "Peter") and then uses a New York State Identification and Intelligence System (NYSIIS) phonetic coding system to account for alternate and misspelling of names. Observations from the passenger records are matched forward to the 1910 census by first looking at exact matches based on the above criteria. If there is one unique match, the procedure stops and the individual is considered matched. If there is not a match, I try matching within a 1-year age band (older and younger) and then within a 2-year age band; if there is one unique match, the individual is included in the first generation sample. However, if there are multiple matches, or there is no match, the observation is discarded as

unmatched. After the first generation is matched, I identify their US-born sons, ages 0-18, and match them forward to the 1940 census using the same matching technique.¹

Methods

To understand how premigration status influences later success, I first regress occupation income in 1910 on a set of control variables including the immigrant's premigration occupation and ethnicity. The occupation income score (OCCSCORE) is calculated by IPUMS and reflects the median income of each occupation observed in the 1950 census in hundreds of dollars. For ease of interpretation, I convert this measure into 2010 dollars. I estimate outcomes of permanent immigrants to determine whether their premigration status predicts their occupational success using the first linked sample. This includes all immigrants who successfully match from the passenger record to the 1910 census and fit the data constraints defined above. Using the first linked dataset of first generation immigrants, I fit the following regression specification:

$$OCCSCORE_{it} = \alpha + \beta_1 HISCLASS_i + \beta_1 ETHNICITY_i + \gamma X_i + e \quad (1)$$

where $OCCSCORE_{it}$ is the occupation income score for person i measured in 1910. $HISCLASS$ is the premigration HISCLASS for the first generation immigrant, $ETHNICITY$ is the ethnicity of the immigrant, and γX_i is a vector of control variables that come primarily from the passenger record: age at arrival, cabin travelled on, and a year of arrival fixed effect. In addition, I add control variables from the immigrant's census record in 1910 including marital status, literacy, and a state fixed effect. These results allow us understand the social destinations of immigrants given their social origins in their sending country.

Once it is established that premigration occupation has an impact on later success for the first generation, I then follow their children and calculate the intergenerational elasticities to directly measure occupational income persistent. Following previous research on intergenerational persistence, I fit the specification:

$$y_{i,g} = \alpha_0 + \alpha_1 y_{i,g-1} + e_{i,g} \quad (2)$$

where $y_{i,g}$ is the log occupation income score for individual i from generation g . The coefficient for β_1 measures the intergenerational elasticity (IGE), that measures the amount of occupational income that a father passes to their son. That is, β_1 is a persistence parameter where higher estimates mean a tighter link between father and son's occupational outcomes that therefore less mobility. I calculate the IGE of individuals at different premigration statuses and control for a quartic of son's age, a quartic of father's age, and son's age interacted with father's occupation to minimize life-cycle bias following previous research

Results

Below are the first generation results using the first link. As shown, premigration occupation has a significant impact on later first generation success.

¹ Reasons for not matching include having a common name, return migration, and name changes. A challenge of matching passenger records to census records is that name changes could have occurred at Ellis Island. Some researchers who have matched passenger records to censuses have allowed for name Americanization to account for this issue (Spitzer; Ward). While match rates are slightly higher for these algorithms, dramatically morphing names like these methods do could lead to potential false matches. Similarly, there is little evidence that immigration agents changed names at Ellis Island as family lore often suggests, but rather name changes happened later after migration (see Cannato 2009). I therefore do not use these algorithms for these data.

Table 1: OLS estimates predicting first generation 1910 occupational income (in \$2010)

	OCCSCORE			LN(OCCSCORE)
	Model (1)	Model (2)	Model (3)	Model (4)
<i>Passenger List Variables</i>				
Premigration Occupation (white collar ref.)				
Skilled	-1032.411*** (95.696)	-784.979*** (94.225)	-815.401*** (94.118)	-.029*** (.004)
Farmer	-2378.774*** (92.439)	-1623.027*** (91.786)	-1619.344*** (91.739)	-.069*** (.004)
Semi-skilled	-1465.758*** (113.523)	1113.447*** (111.482)	-1152.211*** (111.448)	-.045*** (.005)
Unskilled	-1807.187*** (87.717)	-1412.722*** (86.937)	-1361.526*** (86.885)	-.056*** (.004)
Ethnicity (Italian ref.)				
Jewish	635.237*** (60.875)	1944.490*** (74.760)	1232.215*** (77.435)	.046*** (.004)
German (non-Jewish)	234.201*** (53.999)	1690.258*** (70.364)	930.017*** (73.891)	.035*** (.003)
Russian (non-Jewish)	-540.398*** (99.387)	923.868*** (103.630)	548.47*** (103.625)	.022*** (.005)
Arrival age (15 to 20 ref.)				
21 to 25		-539.674*** (55.059)	-548.155*** (55.062)	-.025*** (.003)
26 to 30		-1083.998*** (59.902)	-1059.501*** (59.995)	-.048*** (.003)
Over 30		-2030.703** (61.801)	-1922.626*** (61.829)	-.094*** (.003)
Class traveled on (first/second cabin ref.)				
Steerage		-252.947** (86.474)	-255.201** (86.413)	-.008* (.004)
Stowaway/missing		-237.371 (191.616)	-216.715 (191.807)	-.004 (.009)
<i>1910 Census Variables</i>				
Married			1706.456*** (58.157)	.086*** (.003)
Literate			2590.962*** (56.745)	.107*** (.003)
<i>Fixed Effects</i>				
Year of arrival fixed effect (passenger)	N	Y	Y	Y
State fixed effect (1910)	N	Y	Y	Y
Constant	22474.700*** (92.317)	22282.030*** (676.864)	19678.990*** (662.949)	9.831*** (.031)
Observations	154,600	154,098	151,864	151,411
R-squared	.008	.075	.089	.100

+ .05 < p < .1, * p < .05, ** p < .01, *** p < .001 (two-tailed)

Note: Data come from the matched passenger record to 1910 census record dataset.

This second table provides the intergenerational elasticities (IGE) for second generation immigrants. As shown, the IGEs are low for all groups pointing to the idea that the effects of premigration status after arrival do not last long.

Table 2: Intergenerational elasticities of various immigrant groups

Sample	IGE	IGE (with state fixed effect)	Observations	Clusters
Pooled	.239*** (.012)	.174*** (.011)	52467	34665
Italians Only	.069** (.026)	.071** (.026)	7218	5366
Jews Only	.294*** (.022)	.215*** (.021)	14651	9673
Germans Only	.277*** (.016)	.205*** (.016)	28310	18103
Russians Only	.286*** (.055)	.143* (.056)	2288	1539

+.05<p<.1, *p<.05, **p<.01, ***p<.001 (two-tailed)

Note: Data come from the matched second generation sample 1910-1940. Father's occupational income is measured in 1910 and son's occupational income is measured in 1940.

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(5): 1832-1856.
- Alba, Richard and Victor Nee. 2003. *Remaking the Mainstream: Assimilation and Contemporary Immigration*. Cambridge: Harvard University Press.
- Borjas, George. 1994. "Long-Run Convergence of Ethnic Skill Differentials: The Children and Grandchildren of the Great Migration." *Industrial and Labor Relations Review* 47(4): 553-573.
- Liebertson, Stanley and Mary Waters. 1980. *From Many Strands: Ethnic and Racial Groups in Contemporary America*. Census Monograph.