

**Socioeconomic Attainment in the Ellis Island Era**

Michael J. White  
Erica Jade Mullen

Population Studies and Training Center  
Department of Sociology  
Brown University

Paper Presented to the  
Population Association of America  
San Francisco  
May 2012

We acknowledge support of the National Institutes of Health training award 5T32HD007338 to the Population Studies and Training Center, Brown University.

## **Socioeconomic Attainment in the Ellis Island Era**

### Abstract

This project addresses a gap in the assimilation literature. Contemporary immigrant assimilation theory compares today's immigrants to Southern and Eastern European immigrants from the great wave of the late nineteenth and early twentieth centuries (the Ellis Island era), yet the latter group's socioeconomic assimilation has not been tested empirically with longitudinal data. Using several decades of IPUMS census data, we utilize both double cohort methodology and OLS microdata regression to test the "Ellis Island myth" that those who arrived during the Ellis Island era managed rather quickly to climb the socioeconomic ladder. Our results show that while the first generation (the foreign born) exhibit decidedly inferior labor market outcomes, socioeconomic attainment (measured as SEI points) increases quickly with duration in the US. Persons of the second generation and those of mixed parentage show much less penalty than immigrants. We uncover differences in outcome by European region that do not disappear over the decades we examine.

## INTRODUCTION

### *The Ellis Island Mythology*

The discourse on American immigration is dominated by the “Ellis Island” narrative. A widespread perception prevalent in academic and nonacademic conversations about America’s history of immigration is that Europe’s tired and poor (most of whom were from Southern and Eastern Europe) arrived with little more than change in their pockets, but they managed with hard work and determination to achieve the American dream. That powerful image did more than crystallize the identity of the United States as a country of immigration. The “Ellis Island Myth”—the up from the bootstraps narrative—continues to frame scholarly and public discourse regarding the experience of contemporary immigrants. Although the current, early 21<sup>st</sup> Century period is described as one of record-level immigration for the United States, the demographic scale of immigration during the Ellis Island era was arguably larger (White and Glick, 2009).

To be sure, the United States received and assimilated these late 19<sup>th</sup> Century and early 20<sup>th</sup> Century immigrants. The broad historical sweep of evidence is clear: the waves of European immigrants did make their way successfully into American society, even despite concerns of inassimilability that grew in the wake of these boatloads of new arrivals. At the same time, our evidence on this point is aggregated and unrefined. Exactly how rapid was the rise? How did progress manifest itself across generations? How did different national origin groups fare?

Embedded in the American narrative about immigration are notions of answers to these questions. These, too, make up part of this Ellis Island myth and our current thinking in scholarly, lay, and policy circles. In this paper we aim to shed a brighter light on these questions and provide deeper empirical evidence on them. Thus, we help see the myth in greater refinement. In particular we aim to measure how much difference there was across time and space (origin, nationality) for these newcomers to America’s shores.

Consider the view expressed by Cecil Woodham-Smith in a classic historical work on the great Irish famine migration:

Very few of the poor Irish who fled from Ireland in the famine emigration were destined to achieve prosperity and success themselves.

...It was not until the second or third generation that...the children and grandchildren of the poor famine immigrants became successful...”

C. Woodham-Smith *The Great Hunger* (1962)

At base Woodham-Smith and others argue that it took multiple—many—generations for Ellis Island immigrants to enter the mainstream. Even mid-20<sup>th</sup> Century academic writing questioned the rate of assimilation of the Ellis Island wave. Writing with an eye toward policy, Hugh Carter in a 1949 volume of the *Annals* devoted to immigration offered:

Another important and frequently discussed factor in immigration policy is the assimilation of the foreign born...One hears less today [1949] of the old and naive theory of the “melting pot” according to which assimilation is both complete and rapid.

H. Carter (1949), p. 2-3

Here we can test these assertions directly. Much more recent writing on immigration often hearkens back to the presumed experience of the Ellis Island era. Perlmann, in his informatively titled *Italians Then, Mexicans Now*, makes the explicit comparison between major groups of the Ellis Island era and the contemporary wave, respectively:

Because the Mexican second generation is faring less well in relative terms than its SCEN counterpart...economic assimilation may take more time, four or five generations rather than two or three.

J. Perlmann, *Italians Then, Mexicans Now* (2007), p. 124

Clear then, the Ellis Island era did much to set the frame for thinking about immigrant assimilation.

In a similar vein much evidence has accumulated about national origin differences. We know, for instance, that the Irish started out in New York's most menial urban jobs but later moved into public employment, which afforded higher social status, and we know that the Jewish concentrated in New York's commerce and clothing industries, which helped buttress their social mobility in the US (Waldinger, 1995). But no study has examined 19<sup>th</sup> and 20<sup>th</sup> Century immigrants' socioeconomic progression across time on a national scale. Furthermore, Chicago School sociologists pointed to the colonies of settlement and suggested the associated differences in socioeconomic attainment that accompanied ethnicity and nationality. Subsequent scholars gave this more concrete and fixed form when they began to distinguish between "new" and "old" immigrant groups, further categorizing individuals by country of origin.

### *Contribution*

This project is an effort to test the assimilation paths of Ellis Island era immigrants, whether European immigrants from peasant backgrounds, those escaping famines, those arriving with little more than coinage, were able to "make it" in America. This is a necessary task if we are to shape our contemporary immigration understanding accurately on historical immigration patterns. We examine the "Ellis Island myth" empirically in a way that accounts for the truly temporal process of assimilation. We focus on socioeconomic outcomes and test the Ellis Island era immigrants' attainment using decennial census data available through the Minnesota Population Center's Integrated Public Use Microdata Series (IPUMS). We take insights from two parallel paths of investigation. On one path, we make use of cross-sectional microdata regressions that test for generation and national origin differences in socioeconomic achievement. In the other path, we gain further insight from double-cohort models that reveal the progress from decade to decade for persons of the same age who arrived at the same time. Our work will be able to explicitly test for the presence of differences by origin, their magnitude, and persistence over time. And not inconsequentially we will be able to do this with individual level models based on microdata with several key characteristics controlled. Our combination of double cohort methodology and microdata regression analysis provides us with both graphical depictions of immigrant socioeconomic attainment over time and hypothesis testing of the differential effects of duration in the US, generation status, and European origin on socioeconomic attainment.

## BACKGROUND

### *Historical Policy Context*

Both public policy and zeitgeist in the early 20<sup>th</sup> Century contradict the Ellis Island myth. The 1911 Dillingham Commission, a federal immigration report, concluded that immigrants from Southern and Eastern Europe were deleterious to American society and advised that a reading and writing test be implemented to filter out such undesirable immigrants. This report provided the policy infrastructure for the 1924 National Origins Act, which restricted the number of immigrants from Southern and Eastern Europe who were allowed to enter the US. Thus, there appears to be a disconnect between the government's response to Ellis Island era immigrants at the time and 21<sup>st</sup> Century perceptions of those immigrants. We are not suggesting that Ellis Island era immigrants did not actually climb the socioeconomic ladder; rather, we argue that more rigorous empirical testing of their success beyond analyses of single census years is needed before we can truly treat this era as a benchmark.

### *Ellis Island Era as the Benchmark*

Contemporary assimilation theory—both the new assimilation and segmented assimilation orientations—holds Ellis Island era immigrants as the reference group to which to compare current immigrants' assimilation patterns. The seemingly quick socioeconomic success of Southern and Eastern European immigrants—occurring in two or three generations—is compared to the success of certain groups today (e.g., Chinese and East Indian) and the struggles of other groups (e.g., Mexicans and Laotians). Some students of immigration remain optimistic that contemporary immigrant groups of various ethnic and national backgrounds will experience socioeconomic gains resembling those made by their Southern and Eastern European immigrants of yesterday (*see* Alba and Nee, 2005), while others foresee diverging paths into upward or downward mobility rather than the monolithic improvement witnessed by Southern and Eastern European immigrants of America's past (*see* Portes and Rumbaut, 2006). The problem is, we have yet to thoroughly test empirically just how quickly European immigrants progressed socioeconomically. How quickly did they catch up to the native white majority, and how did the patterns for “old” immigrant groups (from Northern and Western Europe) compare to those for “new” immigrant groups (from Southern and Eastern Europe)?

### *Previous Studies*

Studies of socioeconomic gains made by Ellis Island era immigrants generally compare economic indicators of ethnic groups in the 1910 or 1920 census to those same ethnic groups in recent censuses, or they look at economic change across time for various birth cohorts in a single census. Some find the attenuation of differences between old and new immigrants, while others find continuing differences, or a less rapid assimilation process for new ethnic groups. While some studies find that the second generation of new ethnic groups are less represented in skilled occupations compared to those of old ethnic groups (Lieberson, 1980), others conclude that “the once major differences among specific white groups [in terms of occupational distribution and income] are largely gone” (Lieberson and Waters, 1988: 155).

Comparing SEI differences for twelve ethnic groups in 1910 and 1980, White and Sassler (1995) find that by 1980 ten of the twelve groups displayed uniform mean SEI scores, and whereas group dispersion from the mean SEI was large in 1910, dispersion was minimal in 1980. Additionally, a group's mean 1910 SEI score did not predict its 1980 SEI score. These findings suggest that duration in the host society ameliorates socioeconomic differences observed at arrival. Perlmann (2005) tracks occupational category and wage changes for the new ethnic group second generation of different age cohorts across the 1910, 1920, and 1940 censuses and documents real income gain over time, except during the Great Depression decade.

An abundance of manufacturing jobs is often invoked to explain new immigrants' socioeconomic attainment. While second generation Italians and Poles surpassed their parents in terms of occupation type and earnings after World War II, Katz, Stern, and Fader (2007) explain that this had little to do with the presence of manufacturing jobs, but rather to labor laws and unionization, which mandated decent wages and improved job stability. Waldinger (2007) corroborates Katz et al. by finding that only the Polish were overrepresented in manufacturing jobs, not the Italians, and second and 2.5 generation Polish were both overrepresented in manufacturing, suggesting that they remained in distinct occupation types across the generations, rather than branching out to other occupations. He also notes that manufacturing jobs did not enhance earnings or prestige for immigrants, compared to the earnings and prestige of the native majority. Also, Italians, who were not concentrated in manufacturing, earned higher wages and higher SEI than Poles. Daughters of old immigrant groups married later, thereby contributing to the family economy longer, and they worked in higher-status occupations that paid more than those in new ethnic groups (Sassler and White, 1997). Mellot and Sassler (2007) use the 1920 census to show that jobs status (Duncan's SEI) increases with generational status for old ethnic groups, but not for new ethnic groups.

In terms of school participation, English-speaking immigrants from old ethnic groups and their children experienced a positive, assimilation path, whereas new immigrants and Germans (who were not English-speaking) were less likely to achieve parity with the native stock in terms of school participation (Sassler, 2006). Residential segregation is another outcome of interest to assimilation researchers. Old ethnic groups, particularly the English and Irish, were less residentially segregated in 1910 than new ethnic groups, although the degree of segregation varies by city. New ethnic groups were among the most segregated groups in the 20<sup>th</sup> Century US (White and Sassler, 1995; White, Dymowski, and Wang, 1994).

None of these studies utilizes the double cohort method or attempts to model socioeconomic assimilation longitudinally by using several censuses to approximate a longitudinal panel, which more adequately tests assimilation. The methods we employ here will allow us to track the SEI progress for specific age and arrival cohorts across their labor force years, drawing a better picture of just how linear the process of socioeconomic assimilation was at the turn of the century. Using SEI as the outcome of interest also offers a better picture of assimilation because it is a proxy for social status. If immigrants from specific age and arrival cohorts increase their SEI over their labor force years, we can say that they were able to enhance their social status with time in the US. Analyses comparing occupational distributions of new and old ethnic groups and the native stock are not as informative of social status change for specific cohorts of immigrants.

## *Theory and Hypotheses*

We now turn to an explication of the theory invoked to explain immigrants' expected paths of incorporation into their host society.

The Chicago School social theorists Park (1926), Burgess (1967), Gordon (1964), and McKenzie (1984) developed a paradigm of immigrant integration into the host society that emphasizes the gradual adoption over time of the host society's culture and institutional participation that would lead to improvement in immigrants' and their progeny's socioeconomic attainment. In doing so, immigrants and their children become less distinct from the host society's majority group. Gans (1973, 1992) popularized the "straight-line" and "bumpy line" assimilation paths, both of which focus on intergenerational adjustments to the host society, with the latter accounting for the nonlinear yet upward nature of the assimilation process. We can generally picture the classical assimilation model as a ladder, where recent immigrants stand on the bottom rung and work in the lowest status occupations, but are elevated by the arrival of subsequent immigrants who take the lowest status jobs; the longer an immigrant's duration in the host society, the higher he moves up the ladder. This upward mobility is (presumably) transmitted to the immigrant's children, such that after several generations, differences between ethnic groups and the native stock disappear. Duration is a proxy for experience gained, skills acquired, and social networks formed in the destination society. From a more statistical vantage point, assimilation can be seen simply as a decline in the predictive power of nativity and generation status in determining socioeconomic attainment.

We therefore expect that immigrants who arrived earlier and their children will have higher SEI gain because they've had more time to acquire the occupational skills, English language, and social networks that help boost socioeconomic status, compared to immigrant cohorts who arrived later. The bulk of conventional assimilation literature would lead one to expect that immigrants from "old" ethnic groups in Northern and Western Europe and their children will display higher SEI gain because these groups have been in the US for longer, and immigrants from Southern and Eastern Europe and their children will have lower SEI gain because they've been in the US for a shorter period.<sup>1</sup> Both the first and second generations should have lower SEI than the native stock of native parentage. For ease of exposition we will term such individuals in the third and subsequent generations the "established" population.

## DATA AND METHODS

This project proceeds in three stages. First we analyze double cohort SEI progression in graphical form. Then we run OLS regression models on the aggregated double cohort data to summarize the information contained in these graphs. Third and finally, we turn to OLS cross-sectional regression models to discern differences between first, second, and 2.5 generation working age adults.

### *The IPUMS Data: Years and Variables Included*

We analyze data from the 1900, 1910, 1920, and 1930. These censuses capture the population of European immigrants who arrived in the United States between 1880 and 1920—

---

<sup>1</sup> We of course need to keep selection in mind; individuals may have different characteristics before emigration that influence assimilation once in the destination.

the largest wave of immigration in America's history. These four censuses also provide information on immigrants' year of arrival in the US, allowing us to track the SEI change of immigrants from different arrival cohorts using the double cohort model.

In our cross-sectional microdata regressions, we model the first, second, and 2.5 generations separately. The first generation consists of foreign born individuals. The second generation is US born individuals whose parents are both foreign born. The 2.5 generation is also US born individuals, but only one parent is foreign born—either the mother or the father. The third generation is the reference category, and it consists of individuals born in the US to US born parents. The third generation is also referred to as the “native stock”. For the first and second generations, we distinguish European region of origin. These censuses do not provide information on the native stock's ancestry.

European origin is divided into two groups to avoid sparseness in regression models: Northern and Western European (the “old” ethnic groups), and Southern and Eastern European (the “new” ethnic groups). All individuals from Northern and Western Europe are combined because these nationalities have the longest duration in the US. Southern and Eastern Europeans arrived later—mostly between 1880 and 1920—to a different context of reception. Southern and Eastern European nationalities were considered racially inferior to Northern and Western Europeans, and many had few skills upon arrival to the US. The exception is the Jewish population, which was highly skilled at the time as merchants and artisans. We test for differences between the Jewish population and the rest of the Southern and Eastern Europeans (mostly Italians and Poles) and note differences.

Because of changing national borders, we rely as much as possible on mother tongue to distinguish the two regional groups. For example, we are careful to make sure that certain “new” ethnic groups (the Polish and Jewish) that at times lived in Germany (part of Western Europe) are not classified as Northern or Western Europe. Since the mother tongue variable is not available in the 1900 census, this means that in that census, the Northern and Western European group will not include Germans (since we code as German only those born in Germany whose mother tongue is German), while the 1910, 1920, and 1930 censuses, will include Germans. Similarly, the Jewish and Polish-speaking are not identified in the 1900 census, but they are identifiable in the other three censuses. The 1900 census does specify individuals born in Poland (but they may or may not speak Polish). We include only these European groups of interest in the model and lump together all other nationalities and ethnic groups (Latin Americans, Africans, Asians, and others). These groups represented a small proportion of all immigrants in the US in the late nineteenth and early twentieth centuries. For our microdata regressions, we model interactions between generation and European region of origin to test for differences in the effect of European region by generation status.

We assign ancestry on the basis of father's place of birth when the father and the mother are from different places. For example, if the father is from Poland and the mother is from France, the individual is coded as Southern or Eastern European, not Northern or Western European. We do this because we anticipate that father's background and experience in the labor market will impact male children's socioeconomic attainment more than will mother's, given the greater labor force participation of men in the late nineteenth and early twentieth centuries.<sup>2</sup>

In the microdata regressions, duration in the US is a continuous variable, and we include a quadratic duration term to account for the curvilinear relationship between time spent in the US and socioeconomic attainment. Arrival cohort, used in the double cohort graphs, is derived from

---

<sup>2</sup> In doing so we recognize the need for empirical analysis on the implications of alternative classification schemes.



year of arrival in the US, and it is divided into three categories: pre-peak (arrived before 1880), early peak (arrived during 1880s and 1890s), and late peak (arrived during 1900s and 1910s). The “peaks” refer to the magnitude of immigration witnessed at the turn of the 20<sup>th</sup> Century. The last two decades of the nineteenth and the first two decades of the twentieth centuries were when a huge wave of Southern and Eastern European immigrants arrived in the US. Distinguishing the 1880 – 1899 score from the 1900 – 1919 score allows us to test for period differences within the heightened immigration decades. Pre-peak is the reference category in the aggregated cohort regression models. Differences between these three arrival cohorts in SEI progression also may represent unobserved characteristics of immigrants who arrived in the country at different times that may be related to the historical period, e.g., the economic situation in the place of origin at the time of emigration. Half a percent of the immigrants in the sample are missing their arrival year. Most of them are from Japan (59%), China (22%), or Korea (4%). We replace these missing values with the mean year of arrival for their national origin group and their 5-year age group.

Age is a continuous variable in the microdata regressions. We include also a quadratic age term to account for the curvilinear relationship between age and socioeconomic attainment anticipated by the standard labor market model. In the double cohort graphs, age is categorized into ten-year age groups to reduce sparseness. In the aggregated cohort regressions where we regress each cohort’s SEI on its SEI from the previous decade, we control for age at the end of the decade.

### *SES Outcome Measure – Duncan’s SEI versus Other Measures*

We use Duncan’s Socioeconomic Index (SEI), as provided in the IPUMS, in all our analyses. Duncan’s SEI is a measure of occupational status based on the income level and education level associated with each occupation in 1950. Prestige is built into this index. Respondents rate occupations according to their perceived social status, and then those prestige ratings are regressed onto occupational income and education. That model is then used to create scores for each occupation. An advantage of the SEI is that the scores are equivalently calibrated across all censuses. Occupational income score is another index provided by IPUMS. This score assigns each occupation a value representing the median total income (in hundreds of 1950 dollars) of all persons with that occupation in 1950; the scale’s values may change slightly across censuses (Minnesota Population Center). The elapsed time between 1950 and the censuses we use and the lack of individual income data give us pause in using this scale alone. The SEI and occupational income score are highly correlated (.85). While Treiman (1977) argues that socioeconomic scores, rather than prestige scores, may be a better indication of whether occupations are resources that transmit advantage across generations, prestige is precisely what we aim to measure when assessing assimilation, which is why we choose to model SEI. Both SEI and occupational income score categorize occupations, which captures skill, another measure that indicates assimilation.<sup>3</sup>

---

<sup>3</sup> The Nam-Powers-Boyd index is another measure of occupational status based on the median earnings and median educational attainment associated with each occupation. It assigns equal weights to education and earnings (Minnesota Population Center; Nam, 2000; Hauser and Warren, 1997). Its correlation with SEI is much lower (-.42). Hauser and Warren (1997) note that because the scores are based on the average percentile of an occupation’s workers in the cumulative distribution of workers, after occupations are ranked by median education and median income, they will vary with time as the relative standing of occupations changes, thus complicating temporal analyses of occupational change such as ours.

### *Cohort Structure: Age and Arrival Cohort across Censuses*

The double cohort method is a particularly insightful tool for examining assimilation (*see* Myers and Lee, 1996; Myers and Cranford, 1998). Categorizing and nesting immigrants jointly into their *age cohort* and *arrival cohort* and tracing each “double cohort’s” SEI score across censuses allows us to track immigrants’ socioeconomic attainment over time, discern separately age and arrival cohort patterns, and test whether the combination of specific age and arrival cohorts produces distinct socioeconomic outcomes. This provides us with a pseudo-longitudinal data file. Due to sample size considerations, we use broader age and arrival time groupings than some other double-cohort studies, but the concept is the same. Change in SEI for each double cohort is graphed on a line chart. We track double cohorts across the four censuses for the two European regions of interest here (Northern and Western Europeans, and Southern and Eastern Europeans) as well as for the Jewish since they started off in the US with more skills than other Southern and Eastern Europeans.

It is important to note that the people we track across the censuses are different people, not the same people as in a real panel study. The population changes from census to census, so the people we are comparing across censuses are slightly different; in fact the IPUMS data are samples of these individuals, so sampling error and differential cohort gain/loss influence our values. Nonetheless these values are conceptualized in terms of true underlying cohorts.

In addition to graphs showing SEI progression across the three decades, we run OLS regression models for the same aggregated double cohort data, regressing each cohort’s SEI onto its SEI for the previous census (its “lagged” SEI). This lagged SEI approach tells us the ten-year average change in SEI. We include other regressors in the model—age, European region of origin, and arrival cohort—to see how the SEI slope changes with these covariates. The regressions are also run separately for each census year to parse possible period effects. These regressions are restricted to cohorts ages 30 to 64, who were ages 20 to 54 in the prior decade. We apply a weight to the regressions to control for the different sizes of the double cohorts. For example, we have 3,368 men from Southern or Eastern Europe who were age 30 to 35 in the 1910 census and who arrived in the 1900s or 1910s, while we have 1,054 men of the same census year, arrival cohort, and age but who are from Northern or Western Europe.

### *Cross-sectional Microdata Regression Models*

In addition to the aggregated double cohort graphs and regressions, we run regressions on the individual level microdata. By retaining more observations these regressions allows us to formally tease out the differential effects of other characteristics and SEI attainment, for example, duration and metropolitan status, even though we sacrifice the more longitudinal inferences of the double cohort approach.

We elect to restrict the cross-sectional microdata models to the metropolitan population. This allows us to examine the comparative performance of immigrants and natives, without the confounding effect of metro-nonmetro (urban-rural) labor market differentials and the heavy weight that the SEI score for farmer (and other rural-based occupations) would place on these results. We note that 62% of our immigrant and second-generation populations reside in metropolitan areas. We also, at least initially, examine regression models restricted to the white population. We do so for two related reasons. First much previous research on immigrant

assimilation, e.g. Lieberson and Waters (1988) tends to focus on white immigrants and ethnics. Second, reflecting the reasoning made by others in this earlier ethnic mobility, the sharply discriminatory environment -- and labor market -- faced by African American in this era would serve to obscure the generational comparisons we wish to make. Despite the presence of discrimination within the white population in this era (likely manifest in our statistical results), new immigrants from Europe likely benefited from a socio-cultural regime which favored their complexion over that of US born natives who traced parentage to recent slavery.

[We do examine some models that include all spatial territory, and we are investigating parallel models that examine the fortunes of African American within the same IPUMS framework.]

We run our models first pooled across all decades, and then we estimate separate regressions by decade, allowing us to parse possible period effects (e.g., economic depressions) on socioeconomic status. Only labor force age men (ages 15 to 64) are included in the regressions and are not weighted due to the near-fixed weight IPUMS design. Our models initially include only basic covariate of age (and its square), generation, and US duration (for the foreign-born). We then add region-generation dummy variable interaction terms) to test for variation in the assimilation process among those recent arrivals.

## EMPIRICAL RESULTS

### *Double Cohort Results*

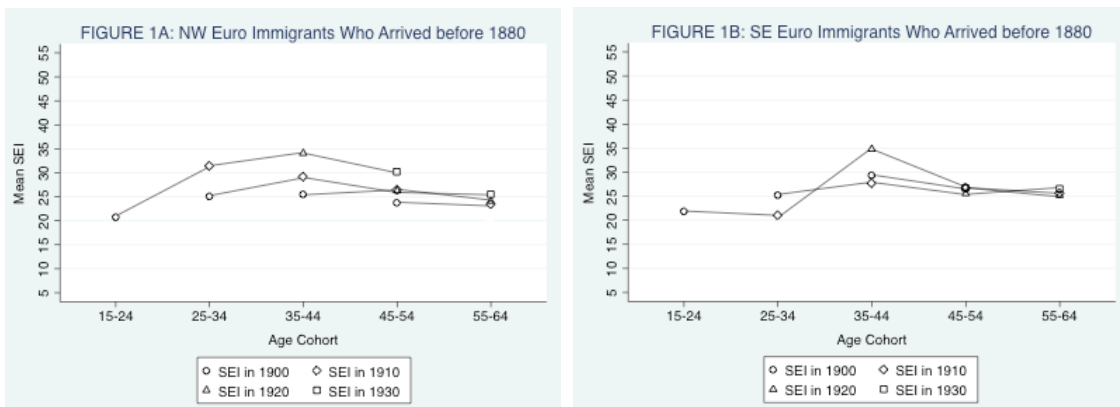
The variables used in the double cohort analysis are summarized in Table 1. As expected, we have more cohorts of immigrants who arrived in the early and late peak years than in the pre- and post-peak years. Just over thirty percent of the cohorts are represented by immigrants from either Northern or Western Europe (32%) or Southern or Eastern Europe (34.2%). Fourteen percent of the cohorts are Jewish and 23.4 percent are made up of either Italians or Poles. A quarter of the other white immigrants in the analysis are from all other world regions. The mean age across the cohorts is the late forties. The mean SEI score for all the cohorts is 27 (with a minimum of 12 and a maximum of 59), and the lagged SEI is slightly less at 24 (with a minimum of zero and a maximum of 58).

	Standard Deviation	% or mean§
Arrival cohort		
Pre-peak (before 1880)*	.42	24.5
Early peak (1880s or 1890s)	.48	34.2
Late peak (1900s or 1910s)	.48	34.2
Region/ethnicity of origin		
Northern or Western European*	.47	32.0
Southern or Eastern European	.48	34.2
Jewish	.35	14.3
Italian or Polish	.42	23.4
Other white immigrants	.43	25.0
Age group (5-year intervals)	9.8	45-49
Duncan's SEI	8.4	27
Lagged Duncan's SEI	9.4	24

Source: Integrated Public-Use Microdata Series 1% Sample of 1900, 1910, 1920, and 1930 Censuses

Note: \*Reference category in multivariate regressions. §Unweighted. Lagged Duncan's SEI is used in the double cohort regression analysis only. Only white immigrants are included in cohorts. N cohorts = 243.

*Cohort trajectories.* Figures 1 through 3 graphically depict the double cohort method using the censuses from 1900 through 1930. Figure 1A displays the SEI trajectories for Northern and Western European immigrants who arrived before the 1880s, by ten-year age cohort. By following the lines between points on the graph, we can see the SEI of those, for example, ages 15 to 24 in 1900, their SEI ten years later in 1910 when they were ages 25 to 34, and so forth. Figure 1B shows the Southern and European immigrants who arrived during the pre-peak years. Figure 2 shows the same SEI progression for the two European groups who arrived during the early peak years, and breaks the Southern and Eastern European group down into Jewish on the one hand and Italian or Polish on the other to see whether the two groups experienced different SEI progression (we show only the early and late peak cohorts due to small N sizes for the pre-peak cohort). Figure 3 shows the four groups who arrived during the late peak years. We do not graph the post-peak arrival cohort due to small N sizes.

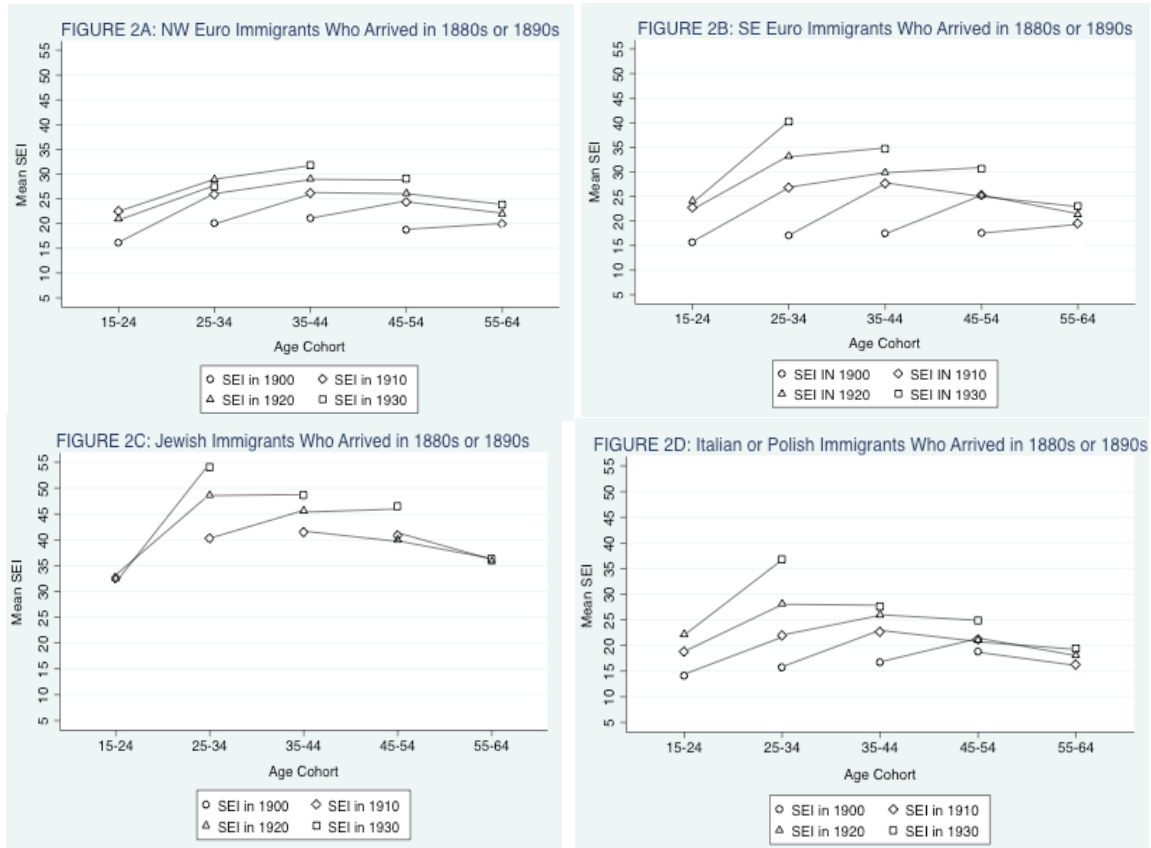


Source: Integrated Public-Use Microdata Series 1% Sample of 1900, 1910, 1920, and 1930 Censuses

Across all arrival cohorts, the graphs depict the expected curvilinear trend of SEI increasing then leveling off with age. Figure 1A shows that SEI increases for the Northern and Western European group (hereafter, NW Euro), and then either decreases or is stagnant when the cohort reaches the oldest ages. SEI progression for Southern and Eastern Europeans (hereafter, SE Euro) is the same as for the NW Euro, except for the 15 to 24 age cohort starting off in 1900, which is doing worse than its NW Euro counterpart by age 25 to 34, then experiences a sharp increase in SEI, catching up to the NW Euro group (although only 14 individuals make up the mean 1920 SEI point for that age cohort). Otherwise, there is no large difference in these two European groups' SEI progression across the four censuses.

Turning next to the early peak arrival cohort in Figure 2, we see upward SEI progression for both the NW and SE Euro groups. Every age cohort in the SE Euro group starts off with slightly lower SEI than its corresponding age cohort in the NW Euro group, except for the youngest age cohort, but they all catch up to the NW Euro cohorts. The SE Euro 15 to 24 age cohort that starts off in 1910 actually surpasses its corresponding NW Euro cohort. Figures 2C and 2D split the SE Euro groups into the Jewish and the Italian and Polish, to see whether the Jewish are driving the overall SE Euro trends. As expected, we see much higher SEI scores for the Jewish. Like the NW Euro group, they experienced upward SEI progression attenuating with age. The Jewish SEI scores were likely driving what we saw in the SE Euro graph because the Italian and Polish SEI scores are about 5 to 8 points lower across the years than those seen in the

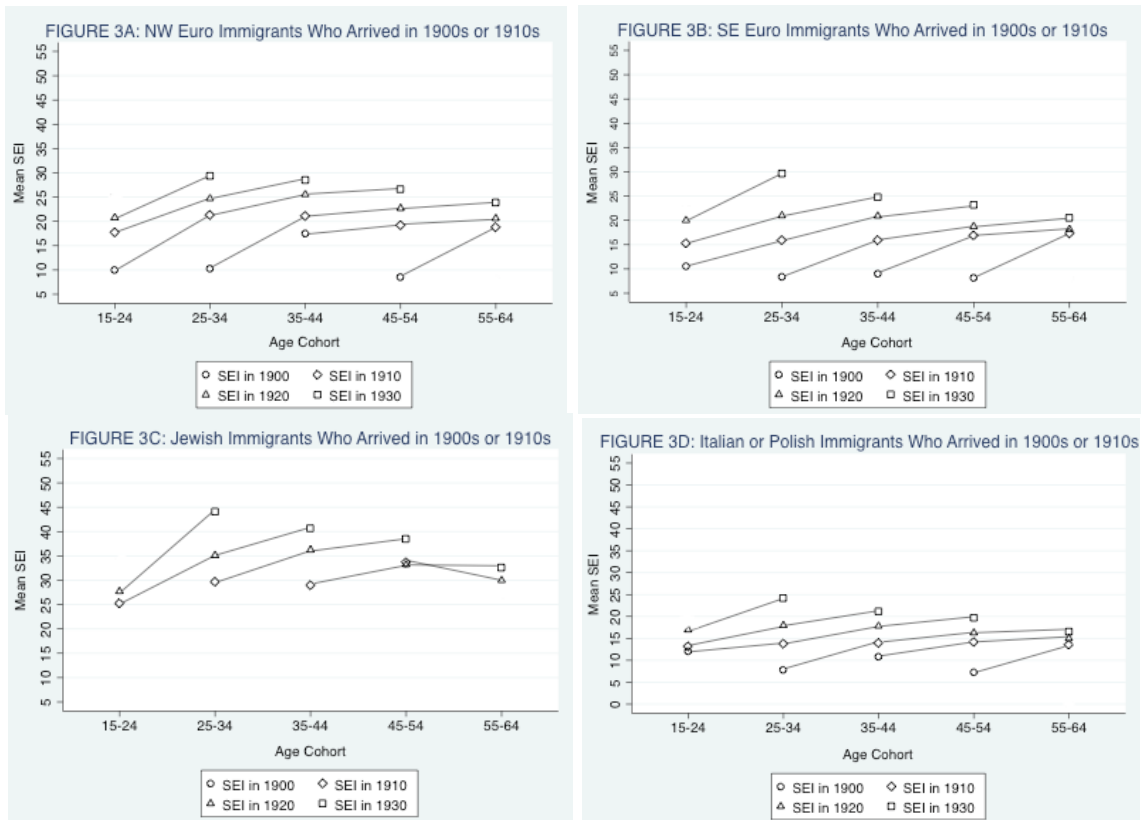
combined SE Euro graph. The younger Italian/Polish age cohorts do not experience as much increase in SEI across the decades as in the aggregated SE Euro graph. Compared to the NW Euro group, the Italian and Polish sit lower on the X-axis with one major exception—the 15 to 24 age cohort starting off in 1920, which surpassed the corresponding NW Euro cohort.



Source: Integrated Public-Use Microdata Series 1% Sample of 1900, 1910, 1920, and 1930 Censuses

Both the NW and SE Euro groups who arrived in the late peak years are noticeably lower on the X-axis than earlier arrival cohorts (Figures 3A and 3B). While they experience upward SEI progression like the two earlier arrival cohorts, they start off at much lower levels. The NW Euro almost reach the highest SEI of the pre-peak and early peak arrival cohorts, but the SE Euro are well below, and they're not making up the difference as they age like their NW Euro counterparts. For both European groups, with each arrival cohort, the SEI reached by the 45 to 54 or 55 to 64 age cohort is lower than that reached in the preceding arrival cohort. The two SE Euro 15 to 24 age cohorts starting in 1900 and 1910, and the SE Euro 25 to 34 age cohort starting in 1900, each experience less SEI progression than the corresponding NW Euro cohort.

We see an overall lower placement on the X-axis for the Jewish when looking at the late peak arrival cohort (Figure 3C), as for the NW Euro group, and the Jewish do not catch up to the SEI achieved in the early peak cohort. Still, the Jewish sit higher on the X-axis than the Italian/Polish (Figure 3D). For neither arrival cohort do the Italian/Polish ever reach the SEI achieved by the Jewish. The two groups never even overlap in SEI scores—the Jewish are always higher.



Source: Integrated Public-Use Microdata Series 1% Sample of 1900, 1910, 1920, and 1930 Censuses

Overall, the graphs depict, within each arrival cohort and for both regions of Europe, a uniform starting point for immigrants of all ages, with the youngest immigrants making the most SEI progress as duration in the US increases. Duration thus appears to have a bigger impact than age, supporting classical assimilation theory’s prediction that SEI improves with time in the host society. Arrival cohort also differentially impacts SEI progression. We observe that—when looking at the same four census years—subsequent arrival cohorts do worse. We therefore expect a negative beta for the early peak and late peak arrival cohorts compared to the pre-peak arrival cohorts in the aggregated regression models discussed next. The NW Euro group does not appear to derive an advantage in terms of SEI progression from the fact that this group is more established in the US, counter to our expectation based on classical assimilation theory.

*Cohort regressions.* Turning now to the aggregated cohort regressions, the results corroborate the patterns seen in Figures 1 through 3. First looking at the effect of a cohort’s lagged mean SEI on its SEI in 1910, 1920, and 1930 (the lagged SEI is from 1900, 1910, or 1920), we see in the pooled results a positive and significant association. That is, for any cohort, each single SEI point higher in the previous decade translates into just 0.6 of a point higher SEI, holding the other variables constant.<sup>4</sup> Most likely regression to the mean effects are operating

<sup>4</sup> In the double cohort graphs and the microdata regressions below, NW Euro includes all countries in Northern and Western Europe. In the aggregated cohort regressions, NW Euro is the English, Irish, and German only, which are the largest groups who emigrated to the US. We plan to update the results using the more inclusive version of NW Euro that includes all countries in NW Europe in the aggregated cohort regressions. The findings should remain the same.

here, and hence, the lagged SEI coefficient takes on a value less than unity. This result has the further complication that we are pooling across pairs of years with different age-arrival cohorts and numbers of cohorts contributed per decade. Subsequent columns show the regression results separately for each concluding census year, e.g., Column 4 presents coefficients for the regression of 1920 cohort-specific SEI on its corresponding value in 1910 and the other cohort covariates. While these results provide the more specific manifestation of immigrant socioeconomic advancement, the results are not uniform across the decades. While the SEI change between 1910 and 1920, and between 1920 and 1930, reflect the positive and significant pooled results, the coefficient on lagged estimated SEI for the first decade is even slightly negative and non-significant. This indicates that there was no significant predictive power from mean SEI score in 1900 on its 1910 value when all other variables controlled. Since this 1900-1910 model still predicts well overall ( $R^2=.94$ ), we may be seeing other factors competing away explanatory power. We checked further on this. All three split decadal regression models show a significant positive association between the lagged SEI score and the cohort SEI value as of the concluding decade, but the coefficient value, and hence strength of the association, is weakest for 1900-1910.<sup>5</sup> This suggests either changes in structural (attainment) conditions in the first decade of the century, or perhaps less consistency or precision in assigning SEI scores in 1900.

We have modeled age through a quadratic effect. In many models, age coefficients display the expected curvilinear relationship between age and socioeconomic attainment, where SEI increases with age at a decreasing rate. It is likely that our regression models with these cohort pairs and a quadratic in age push the limits of our sample size [ $42 \leq N \leq 92$ ], which sometimes leads to anomalous results, such as the negative coefficient on the first order effect of age in the 1930 outcome regressions.

Very strong ethnicity effects are discerned. Jewish immigrants outpace their NW Euro counterparts: cohorts of Jewish immigrants are predicted to score 7 points higher SEI than NW Euro cohorts by 1920, holding age and arrival cohort constant. The Italian and Polish origin groups, on the other hand, exhibit a deficit with respect to NW Euro counterparts in every decade. For example, Italian or Polish immigrant cohorts are predicted to score 3.7 points lower SEI than those from NW Euro cohorts by 1920, net of age and arrival cohort. Other white immigrants were also falling short of the NW European groups, but with a smaller gap than Italian and Polish immigrants. Thus, the NW European origin groups exhibit an advantage over some groups (the Italian and Polish), but not others (the Jewish). This may be linked directly to human capital upon arrival within the origin groups, with NW Europeans having an advantage over some (perhaps the Italian and Polish) and a disadvantage with respect to others (perhaps Jews).

Consistent with the double cohort graphs, we see in this multivariate framework that cohorts arriving *later* in the era (“late peak”) exhibit generally lower socioeconomic outcomes than the reference group of pre-1880 arrivals and often the early peak (1880s and 1990s) arrivals, once age and region/ethnicity are controlled. (Data depth is not sufficient to estimate effects for the post-peak arrival cohort. Further, immigration was abating by 1930 and restrictions had already begun to be put in place.) The pooled results (col 1) show that on average, immigrants who arrived most recently gained less than their earlier vintage peers in SEI. Such findings as these may reflect better skill endowments among the earliest arrivals, i.e., “pre-peak” cohorts in our analysis. There may be no specific disadvantage to being the newest group in the host society; rather, earlier immigrants may have been positively selected at their origin.

---

<sup>5</sup> The regression model of SEI 1920 on only SEI 1910 provides a coefficient of 0.72 with an  $R^2$  of 0.12.

TABLE 2: Aggregated Cohort-Level OLS Regression Results				
	Pooled (all 4 censuses)	1910 SEI	1920 SEI	1930 SEI
Lagged SEI (previous decade)	0.607*** (.033)	-0.061 (.076)	0.443*** (.103)	0.719*** (.114)
Age group at end of decade	0.106 (.127)	0.971*** (.169)	0.207 (.188)	-0.204 (.225)
Age group squared	-0.004*** (.001)	-0.013*** (.002)	-0.005** (.002)	-0.000 (.003)
Region/ethnicity of origin‡				
Jewish	4.919*** (.670)	<i>N.A.</i>	7.421*** (1.279)	4.406*** (1.400)
Italian or Polish	-2.614*** (.402)	-4.846*** (.515)	-3.729*** (.782)	-1.993** (.959)
Other white immigrants	-0.933*** (.307)	-1.603*** (.291)	-1.292** (.513)	-0.385 (.612)
Arrival cohort‡				
Early peak (1880s or 1890s)	0.077 (.455)	-3.401*** (.479)	-0.396 (.700)	-0.737 (1.204)
Late peak (1900s or 1910s)	-0.690 (.522)	<i>N.A.</i>	-3.039** (1.241)	-0.935 (1.494)
Constant	13.752*** (2.584)	14.239*** (2.999)	18.359*** (3.808)	19.968*** (4.275)
R <sup>2</sup>	.911	.921	.934	.942
N cohorts	243	42	95	85

Source: Integrated Public-Use Microdata Series 1% Sample of 1900, 1910, 1920, and 1930 Censuses

Note: *i* indexes the cohort. Standard errors in parentheses. N.A. means not available in that census year. The Jewish cannot be identified in the 1900 census. The late peak arrival cohort in the 1910 model would include only individuals who arrived in 1900, so we exclude them from the 1910 model. Age groups are measured in 5-year intervals here. Results reported are weighted using the size of the cohort in the current year. As an alternative, we weight the results using the size of the cohort in the lagged year. The results are consistent.

‡Reference categories are pre-peak for arrival cohort and NW Euro for region/ethnicity of origin.

\*\*\*  $p \leq .01$  \*\*  $p \leq .05$  \*  $p \leq .10$

### Cross-Sectional Regression Results

Model 1 of Table 3 predicts socioeconomic outcomes for simple covariates of age, generation, and duration. These models are pooled 1900-1930 and include white males in the labor market years. We find that age strongly predicts labor market achievement as indexed by SEI score. This is in keeping with the bulk of prior labor market studies (irrespective of era and outcome measure), and so lends further confidence to using the SEI index to track labor market outcomes in this Ellis Island Era. All coefficients in this pooled census sample are highly significant with  $p$ -values under 0.001.

We find that SEI rises sharply with age but at a decreasing rate, also in keeping with most labor market studies predicting socioeconomic outcomes. Our model [ $1.227A - 0.014A^2$ ] indicates that SEI would be predicted to peak at age 42.6. At age 30 the slope of SEI is rising (at a decreasing rate) at about 0.8 points per year.

Model 1 predicts that immigrants (first generation white men) would score over 16 SEI points below established (3<sup>rd</sup> generation or later) white men of the same age in metropolitan labor markets in this era. This is quite a sharp differential, on the order of a standard deviation, in outcomes. Not nearly so large a differential—a generational or immigrant penalty—is visible in the results for the second generation. Native men born of two immigrant parents face a 2.2 unit deficit in SEI from age-mates in the established population. For men who have one



immigrant and one US-born parent (classified here as the 2.5 generation) there is no penalty at all; in fact there is a statistically significant *improvement* in their predicted SEI score over established white males of the same age. Taken together, these first results strongly suggest that much of the immigrant deficit—whether driven by lack of skills, limited acculturation, or discrimination—is erased within a generation, at least among white males in the labor force ages living and working in America’s cities.

TABLE 3: Individual-Level Microdata OLS Regression Results – Pooled		
	Model 1	Model 2
Age	1.228*** (0.016)	1.221*** (0.016)
Age squared	-0.0144*** (0.000)	-0.0144*** (0.000)
Duration in US	0.380*** (0.015)	0.391*** (0.015)
Duration squared	-0.00222*** (0.000)	-0.00284*** (0.000)
Generation 1 (=1)	-16.42*** (0.156)	-13.59*** (0.193)
Generation 2 (=1)	-2.207*** (0.086)	-1.113*** (0.147)
Generation 2.5 (=1)	1.050*** (0.125)	1.002*** (0.217)
Gen1*NW Europe‡		-2.042*** (0.158)
Gen1*SE Europe‡		-4.324*** (0.151)
Gen2*NW Europe‡		-1.195*** (0.168)
Gen2*SE Europe		-2.345*** (0.217)
Gen2.5*NW Europe		0.0502 (0.254)
Gen2.5*SE Europe		0.156 (0.569)
Constant	13.03*** -0.282	13.20*** -0.284
R <sup>2</sup>	0.063	0.065
Observations	475,662	475,662

Source: Integrated Public-Use Microdata Series 1% Sample of 1900, 1910, 1920, and 1930 Censuses

Note: *i* indexes the individual. Standard errors in parentheses.

‡Reference category is generation 3+. Results unweighted.

\*\*\* p ≤ .01 \*\* p ≤ .05 \* p ≤ .10

In model 2 we keep measures of age, duration, and generation status, and we add several covariates that capture interactions between generation and region of origin. We limit these regional comparison to Europe (a) North and West (NW); and (b) South and East (SE), in keeping with previous classifications of immigrant origins. While some earlier writing on this period of immigration referred to NW as “old” and SE as “new”, we adhere more strictly to the regional aspect of the classification, since we are controlling for timing in that we include in our regression analysis a defined set of periods (censuses) and introduce controls for age and duration.

Model 2 indicates that first generation labor force age men from both European regional origins pay an additional penalty beyond that of (gen1) immigrant status itself. NW men are

predicted to have an SEI 2 points yet again lower, while the added deficit for SW men is larger at over 4 SEI points. This regional immigrant deficit is halved by the second generation. Although both coefficients remain statistically significantly different from zero, for the 2.5 generation (one parent born in the US; one born outside) the difference is not statistically detectable at conventional levels of significance.

[Note that our model, by design, does not include first-order dummy variable terms for region, since region of origin is meaningful only for immigrants and their immediate descendants. As a consequence the overall reference group in these regional-generation dummies is the established population and those immigrants from other origins. Immigrants from other than NW and SE are a small fraction of the population, so the reference group really reflects the established white male metropolitan labor force. Of course this also means that immigrants not of these regional origins would not have these additional increments/decrements to predicted achievement; thus immigrants from other regions would be expected to score 13.6 points lower on SEI versus the 13.6 - 2.0 for immigrants from NW Europe.]

Table 4 presents companion results to those of Table 3, Model 2, but in which we estimate separate models for each census year 1900 through 1930. This allows us to examine the variation in predictive traits—duration, generation, national origin—through key early decades of the 20<sup>th</sup> Century. (These models also allow the effect of age on SEI to vary across time.) Results of this model exhibit some parallels with those of Table 3, but there are also a few aspects in which the pooled results are not simply replicated across the four decade-specific regressions.

The variable Age (including both linear and quadratic terms) does perform much as one would expect in a standard labor market attainment model. All four models exhibit broadly similar coefficients for age, and thus, SEI improvement would trace similar patterns with age across the four census decades. So, too, duration (including its square) traces similar patterns from 1900 to 1930. In all three of these census years, one predicts an appreciable socioeconomic return to each year of US residence for immigrants. This return, while remaining positive, declines modestly with further length of stay in the US. In 1930 however, the model deviates: both linear and quadratic terms are positive. While seemingly anomalous, the result is also intriguing. By 1930 the US had seen the imposition of new immigration policies with restriction through the 1924 Naturalization Act (and related policies), while also registering the start of the Great Depression. While both linear and quadratic coefficients for 1930 are positive, substantive examination indicates that the net effect is modest in size and traces a flatter trajectory. For instance in 1910, the combined duration effects predict an increment to SEI of 5.38 with 10 of years US experience and an increment of 12.0 SEI points with 30 years of experience (net of other controls in the 1920 model). For 1930, the corresponding SEI increments are 0.68 at 10 years and 4.0 at 30 years (with other traits controlled), a much less pronounced profile.

These split sample results recapitulate the earlier finding of a significant deficit in predicted SEI among immigrants themselves (Gen 1), and we now observe some variation across decades. We see that immigrants themselves (net of duration, which works to improve SEI), experience SEI deficits of about a dozen to nearly twenty points, depending of origin group and decade. For instance, immigrants from Northern and Western Europe experience an SEI deficit of nearly 18 points in 1900, with those from Southern and Eastern Europe slightly worse. The added deficit for those from Southern and Eastern Europe—a key comparison here—persists across the 1900-1930 decades, and notably, this regional differential (see the magnitude of the coefficients) actually increases from about 1.3 SEI points in 1910 to about 3.0 SEI points by

1930. Even though the overall first-generation offset seems to be less in the final census year, the gap by origin within Europe has widened a bit.

TABLE 4: Individual-Level Microdata OLS Regression Results – Split by Census Year				
	Model 1 – 1900	Model 2 – 1910	Model 3 – 1920	Model 4 – 1930
Age	1.079*** (0.040)	1.021*** (0.033)	1.200*** (0.030)	1.412*** (0.029)
Age squared	-0.013*** (0.001)	-0.012*** (0.000)	-0.015*** (0.000)	-0.017*** (0.000)
Duration in US	0.407*** (0.037)	0.607*** (0.027)	0.482*** (0.034)	0.036 (0.031)
Duration squared	-0.003*** (0.001)	-0.007*** (0.001)	-0.005*** (0.001)	0.003*** (0.001)
Generation 1 (=1)	-13.300*** (0.435)	-16.130*** (0.386)	-13.320*** (0.422)	-9.611*** (0.404)
Generation 2 (=1)	-3.183*** (0.296)	-2.912*** (0.485)	-0.641 (0.401)	0.184 (0.214)
Generation 2.5 (=1)	-0.652 (0.512)	-0.363 (0.681)	0.260 (0.556)	1.685*** (0.297)
Gen1*NW Europe‡	-4.318*** (0.297)	-1.734*** (0.347)	-2.313*** (0.333)	-1.032*** (0.319)
Gen1*SE Europe‡	-5.674*** (0.373)	-3.253*** (0.344)	-5.323*** (0.309)	-4.056*** (0.302)
Gen2*NW Europe‡	-1.727*** (0.355)	-0.111 (0.500)	-0.909** (0.423)	-0.226 (0.291)
Gen2*SE Europe	-0.907 (1.035)	0.296 (0.664)	-0.917* (0.495)	-4.937*** (0.306)
Gen2.5*NW Europe	-0.011 (0.643)	0.596 (0.720)	1.044* (0.596)	0.478 (0.399)
Gen2.5*SE Europe	6.632** (2.842)	2.859* (1.641)	2.259** (1.128)	-1.863** (0.777)
Constant	15.070*** (0.701)	17.040*** (0.571)	13.850*** (0.531)	9.871*** (0.521)
R <sup>2</sup>	0.067	0.088	0.062	0.051
Observations	67,748	104,244	133,011	170,659

Source: Integrated Public-Use Microdata Series 1% Sample of 1900, 1910, 1920, and 1930 Censuses

Note: *i* indexes the individual. Standard errors in parentheses.

‡Reference category is generation 3+. Results unweighted.

\*\*\*  $p \leq .01$  \*\*  $p \leq .05$  \*  $p \leq .10$

The case of the second generation (Gen2 and also Gen 2.5 here) is also informative. In the early two decades (1900 and 1910) we observe a net negative increment in most comparisons (by parentage and region), although we see, mirroring our pooled results, a much weaker effect (or even none) for those working age men with one US-born parent. This effect begins to break down, however. By 1920 and 1930 this second generation offset is more modest (even pressed to nonsignificance in the case of the reference population). The only exception seems to be for SE European individuals in the second generation in 1930 who exhibit a deficit in SEI score of nearly 5 points after other determinants have been controlled. Again, while the overall generational effect seems to be waning, by 1930 a widening of the regional gap is observed. This well may be due to increasing ethnic differentiation (outright discrimination or other factors) in the labor market, but certainly it aligns with the view that there developed a less welcome context of reception for “new” immigrants (and their immediate descendents) from Southern and Eastern Europe.

## CONCLUSION

Ellis Island, myth and reality, continues to hold significant sway over the narrative of the American immigrant experience. To be sure, those who came to America's shores in the late 19<sup>th</sup> Century and the early 20<sup>th</sup> Century found, on balance, success in their new land. Much has been written to document this assimilation. At the same time, questions arise about the uniformity and rapidity of that assimilation. This is all the more so, and more importantly so, as 21<sup>st</sup> Century writing about 20<sup>th</sup> Century immigration criticizes and questions the universality and “straight-line” nature of assimilation.

We exploited the richness of historical census data to begin to photograph statistically the Ellis Island era in greater resolution. Just as the greater accessibility of photography provided more accurate pictures of the new arrivals and their early US lives, so too these early decennial census data files provide a more accurate picture of the assimilation experience: what was common across groups and what was not. We approached this task with a two-pronged method. In one line of work, we aggregated the IPUMS decennial census microdata into double-cohorts, groupings of individuals by age and period of arrival that could be aligned across subsequent decades. These give trajectories analogous to true cohorts, although the data are not (necessarily or identifiably) for the same persons. In the second prong of the approach, we considered each census separately and cross-sectionally, exploiting the full sample. Here we conducted cross-sectional regressions on socioeconomic attainment (Duncan SEI index as the outcome) and determined these as a function of age, duration in the US, generation, and region of origin. In this alternative approach, we were particularly interested in examining the progression of origin groups across generations.

Our results buttress some conventional scholarly viewpoints on immigration. (One must acknowledge however that viewpoints are not all in consensus and surely have changed themselves over the decades!). On balance we find both continuity and progress from decade to decade. Our simple summary graphics point to SEI improvement for most groups across most decade. We also see a slowing of the rate of progress as time wears on, something that is not surprising for those familiar with labor market studies. At the same time we do detect region (ethnicity) of origin and vintage (arrival time) differences in our double cohort models. Cohorts we identify (from language and origin information, following prior scholarship) as Jews achieve higher scores—about 5 SEI points—than other groups. Italian and Poles—merged here as major representatives of the “new” immigration—do less well than the reference group of immigrants from Northern and Western Europe. Finally, although it pushes the limits of our data, there is some suggestion that immigrants arriving during the “late peak” years, i.e., 1900s and 1910s, fare somewhat less well.

The cross-sectional microdata regressions provide important alternative insights regarding the Ellis Island era and the assimilation experience within the early 20<sup>th</sup> Century, also improving our empirical precision for measuring the assimilation process. Our regression equations confirm the expected gains in socioeconomic attainment (as tapped by the SEI index) that come with labor market experience (age) and duration in the US. At the same time we see appreciable generational advance, arguably larger than that often anticipated by writers of the time or those looking back on the era. For urban working age (and white) males, the second generation offset was detectable but much less pronounced than that of the first. Time clearly bestows benefits: we observe appreciable gains over time among the immigrants in their own working years and then across the generations. We do also find some deviations from any all-

encompassing model of assimilation and inclusion. While our results are bound by the necessary limitation of working with historical census data, we appear also to be able to detect a widening regional (ethnic) gap by 1930. One possible explanation is that the concern seen in the national immigration discourse, particularly voiced with reference to new arrivals from Southern and Eastern Europe, is also translated into the context of reception in the labor market and an ensuing penalty in SEI scores we detect.

Taken together, our empirical analyses of these historical census data highlight immigrant and second generation success in the Ellis Island era. Gains were broad-based across origin groups and period of arrival. Most new arrivals (who were recorded in these census data) found their way in the American setting. Most striking in our findings, perhaps, is the weakening of socioeconomic attainment differentials already by the second generation. In many of our models (including some not presented here) second generation urban working-age men exhibit modest socioeconomic outcome differentials from the more established population.

Not all was equivalent, however. A uniformly operating American assimilation machine would have made no room for explanatory power due to ethnic origin or period of arrival, and we find evidence for both. Some of these differences are, no doubt, attributable to differences in skills brought with them by immigrants to Ellis Island and other US ports of entry. But we may also be observing in these results the rise in differential treatment visited upon new arrivals, a differential reception that would become manifest in growing doubt about assimilability and ensuing restrictions on immigration by the end of the Ellis Island period.

While our results remain only as refined as census data allow, they do point to a more rapid assimilation than many observers of the period, whether historians such as Woodham-Smith or policy-makers such as the Dillingham commission, allowed.

## METHODOLOGICAL APPENDIX

Our aim in this analysis is to tap the same pool of people across the censuses, yet, we acknowledge that when we combine several censuses into one large population, we end up with a meta-population, so to speak, because the dataset no longer represents a snapshot of the population from one year. The population changes depending on, for example, how people identify themselves in terms of ancestry or ethnicity, emigration from the US, and subgroup coverage.

We have some degree of omitted variable bias because the census does not include all relevant variables and because of the cross-sectional nature of the census.

The 1900, 1910, 1920, and 1930 censuses are unweighted “flat” samples, meaning each observation represents a fixed number of persons in the US population:

<http://usa.ipums.org/usa/intro.shtml#weights>). Our empirical analysis of the person weight variable (*perwt*) indicate that for 1900 and 1910 virtually all observations took values of 100 or 101, and for 1920 and 1930 weights took on a variety of fractional values, but all clustered tightly around 100. Weighted and unweighted regression were virtually identical.

The labor force universe in these censuses is individuals ages 16 or older. We have occupations recorded for younger individuals because the universe for the *occupation* question is those ages *10 or older* in the 1900 census, and *all* persons in the 1910 through 1930 censuses. (The IPUMS team assigns an SEI to individuals who have an occupation recorded.) Census enumerators were instructed to record an occupation for any person “gainfully employed”—a somewhat ambiguous term when applied to children. These individuals are likely working for their parents, not independently (Minnesota Population Center, 2011), so any observed SEI change for them would be partly a product of their working with their parents. We choose to include individuals ages 15 to 64 to be consistent with other literature covering the same era, and because individuals commonly began working in their early teens in the nineteenth and early twentieth centuries.

## REFERENCES

- Alba, Richard and Victor Nee. 2005. *Remaking the American Mainstream: Assimilation and Contemporary Immigration*. Cambridge: Harvard University Press.
- Burgess, Ernest W. 1967. "The Growth of the City." In *The City*, edited by R. E. Park, E. W. Burgess, and R. D. McKenzie. Chicago: University of Chicago Press.
- Carter, Hugh. 1949. "Reappraising Our Immigration Policy." *Annals of the Academy of Political and Social Science* 262.
- Foner, Nancy. 2000. *From Ellis Island to J.F.K.: New York's Two Great Waves of Immigration*. New Haven and New York: Yale University Press and Russell Sage Foundation.
- Gans, Herbert. "Second Generation Decline: Scenarios for the Economic and Ethnic Futures of Post-1965 American Immigrants." *Ethnic and Racial Studies* 15:173-192.
- 1973. "Introduction." In *Ethnic Identity and Assimilation: The Polish Community*, edited by N. Sandberg. New York: Praeger.
- Gordon, Milton M. 1964. *Assimilation in American Life*. New York: Oxford University Press.
- Hauser, Robert M. and John Robert Warren. 1997. "Socioeconomic Indexes for Occupations: A Review, Update, and Critique." *Sociological Methodology* 27:177-298.
- Katz, Michael B., Mark J. Stern, and Jamie J. Fader. 2007. "The Mexican Immigration Debate: The View from History." *Social Science History* 31(2):157-89.
- Lieberson, Stanley. 1980. *A Piece of the Pie: Blacks and White Immigrants Since 1880*. Berkeley: University of California Press.
- Lieberson, Stanley and Mary C. Waters. 1988. *From Many Strands: Ethnic and Racial Groups in Contemporary America*. New York: Russell Sage Foundation.
- McKenzie, R. D. 1984. "The Ecological Approach to the Study of the Human Community." In *The City*, edited by R. E. Park and E. W. Burgess. Chicago: University of Chicago Press.
- Mellott, Leanna M. and Sharon Sassler. 2007. "Growing Up with Single Mothers: Occupational Attainment of Daughters in the Early 20<sup>th</sup> Century." *Research in Social Stratification and Mobility* 25:73-88.
- Minnesota Population Center. 2011. "Chapter 4: Integrated Occupation and Industry Codes and Occupational Standing Variables in the IPUMS." Accessed 6-15-2011.
- Myers, Dowell and Cynthia J. Cranford. 1998. "Temporal Differentiation in the Occupational Mobility of Immigrant and Native-Born Latina Workers." *American Sociological Review* 63(Feb.):68-93.
- Myers, Dowell and Seong Woo Lee. 1996. "Immigration Cohorts and Residential Overcrowding in Southern California." *Demography* 33(1):51-65.
- Nam, Charles B. 2000. "Comparison of Three Occupational Scales." Working Paper, Florida State University.
- Park, Robert E. 1926. "The Concept of Position in Sociology." *Papers and Proceedings of the American Sociological Society* 20:1-14.
- Perlmann, Joel. 2005. *Italians Then, Mexicans Now: Immigrant Origins and Second Generation Progress, 1880 to 2000*. New York: Russell Sage Foundation.
- Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Shroeder, and Matthew Sobek. 2010. Integrated Public Use Microdata Series: Version 5.0. Minneapolis, MN: Minnesota Population Center.
- Sassler, Sharon L. 2006. "School Participation among Immigrant Youths: The Case of Segmented Assimilation in the Early 20<sup>th</sup> Century." *Sociology of Education* 79(Jan.):1-24.
- Sassler, Sharon and Michael J. White. 1997. "Ethnicity, Gender, and Social Mobility in 1910." *Social Science History* 21(3):321-57.
- Waldinger, Roger. 1995. "When the Melting Pot Boils Over: The Irish, Jews, Blacks, and Koreans of New York." In *The Bubbling Cauldron: The New Political Economy of Race and Ethnicity*, edited by Joe Feagin and Michael P. Smith. University of Minnesota Press.
- 2007. "Did Manufacturing Matter? The Experience of Yesterday's Second Generation: A Reassessment." *International Migration Review* 41(1):3-39.
- Woodham-Smith, Cecil. 1962. *The Great Hunger: Ireland 1845 – 1859*. Penguin Books.
- White, Michael J., Robert F. Dymowski, and Shilian Wang. 1994. "Ethnic Neighbors and Ethnic Myths: An Examination of Residential Segregation in 1910." In *After Ellis Island: Newcomers and Natives in the 1910 Census*, edited by S. C. Watkins. New York: Russell Sage Foundation.
- White, Michael J. and Jennifer Glick. 2009. *Achieving Anew: How New Immigrants Do in American Schools, Jobs, and Neighborhoods*. New York: Russell Sage Foundation.

White, Michael J. and Sharon Sassler. 1995. "Ethnic Definition, Social Mobility, and Residential Segregation in the United States." In *Population, Ethnicity, Nation-Building*, edited by Calvin Goldscheider. Boulder, CO: Westview.