The (South) American Dream:
Mobility and Economic Outcomes of First- and Second-Generation Immigrants in 19th-Century Argentina

Santiago Pérez

Santiago Pérez is Assistant Professor, Department of Economics, University of California, Davis, 1118 Social Sciences and Humanities, Davis, CA 95616. E-mail: seperez@ucdavis.edu.

I have benefited from various conversations with Ran Abramitzky, Carolina Biernat, Arun Chandrasekhar, Melanie Morten and Gavin Wright, as well as several participants in the Stanford Economic History seminar. I am also grateful to Enrique Pérez, Mercedes Pérez, Fernanda Righi, Mariana Righi, Lisandro Vaccaro and María Fabiana Vaccaro for helping me with data collection and with Fiona Stewart for proofreading. This paper benefited from funding from the Stanford University Economics Department; the Economic History Association through the Exploratory Data and Travel Grant and the Dissertation Fellowship; the Graduate Student Fellowships in International Development through a grant from the Stanford Center for International Development; and the Leonard W. Ely and Shirley R. Ely Graduate Student Fund Fellowship.

The (South) American Dream: Mobility and Economic Outcomes of First- and Second-Generation Immigrants in 19th-Century Argentina

I study the mobility and economic outcomes of European immigrants and their children in 19th-century Argentina, the second largest destination country during the Age of Mass Migration. I use new data linking males across censuses and passenger lists of arrivals to Buenos Aires. First-generation immigrants experienced faster occupational upgrading than natives. Occupational mobility was substantial relative to Europe; immigrants holding unskilled occupations upon arrival experienced high rates of occupational upgrading. Second-generation immigrants outperformed the sons of natives in terms of literacy, occupational status and access to property, and experienced higher rates of intergenerational mobility out of unskilled occupations.

During the Age of Mass Migration (1850-1913), 55 million Europeans left their countries of origin and moved to the New World. After the US, Argentina was the second largest destination country in the period, receiving 6.2 million immigrants. By 1914, 30% of Argentina's population was foreign-born.

The conventional view on this migration episode is that Argentina constituted a "land of opportunity," offering European immigrants a good chance to experience upward economic mobility.[1] Although this view has been pervasive in the historical literature,[2] there is little quantitative evidence to support it. Moreover, the quantitative evidence that does exist is based on the published tabulations of the census. While these tabulations provide useful information on the economic performance of immigrants at one point in time, they offer little insight into how immigrants progressed as they spent time in the country. In addition, the published census data contain no information on parental place of birth, which prevents a systematic study of second-generation immigrants' economic performance.

I study the mobility and economic outcomes of European immigrants and their children in

[1]For instance, Szuchman (1981) argues that "Argentines never rejected the belief that their society represented an open system of economic opportunities proven by the upwardly mobile population."

[2]See for example Baily (1983), Conde (1979), Diaz-Alejandro (1970), Klein (1983), among others.

19th-century Argentina. To do so, I use newly collected data linking males across the 1869 and 1895 national censuses of population and passenger ship lists of immigrant arrivals to the city of Buenos Aires. These data enable me to follow a large group of immigrants and their children and to track their progress while in the country. To the best of my knowledge, this paper is the first to use longitudinal data following individuals over time and across places to provide evidence on the economic performance of immigrants in late 19th-century Argentina.

The first part of my analysis looks at the occupational mobility of first-generation immigrants. I ask whether immigrants started in lower paying occupations than natives but converged to them as they spent time in the country. I find that, upon arrival, European immigrants held on average slightly lower paying occupations than natives. Yet, consistent with assimilation into the labor market of Argentina, my findings suggest that immigrants from most of the major sending countries outpaced natives in terms of occupational upgrading. These results contrast with recent evidence on the Age of Mass Migration in the US (Abramitzky *et al.* 2014), where immigrants appear to have experienced similar rates of occupational upgrading as natives.

Once I have characterized the occupational progress of immigrants *after* their arrival to Argentina, I study the extent to which immigrants who stayed in Argentina experienced progress relative to their pre-migration occupations. I find that immigrants were very likely to upgrade their occupations: About 75% of those who declared an unskilled occupation upon arrival experienced occupational upgrading in less than 15 years. Moreover, relatively skilled immigrants experienced little occupational downgrading. Comparing my results to evidence from similarly constructed data for the US (Ferrie 1997), my findings suggest that European immigrants in Argentina were better able to exploit their pre-migration human capital than those migrating to the US.

My data only allow me to characterize the occupational trajectories of immigrants who migrated permanently to Argentina. These immigrants are of special interest from a historical point of view, since they participated in the labor market of Argentina for many years and were also likely to raise children in the country. However, in interpreting the results described above, it is useful to bear in mind that the experience of the average immigrant might have been differed from the experience of those who settled permanently.

The second part of my analysis focuses on the children of European immigrants: the "second generation." I find that the sons of European immigrants experienced substantially better eco-

nomic outcomes than the sons of natives: They were more likely to be literate, held higher paying occupations and were more likely to own property as adults. The relative advantage of the second generation was not confined to any single sending country. Rather, the sons of immigrants from every major sending country outperformed the sons of natives. In explaining these results, I provide suggestive evidence that the higher "ethnic capital" (Borjas 1992) of the children of European immigrants might have contributed to their advantages in adulthood.

Finally, I contrast the intergenerational mobility of second-generation immigrants to that of the children of natives. I find a relatively similar persistence of occupational status among second-generation immigrants compared to children of natives. On the one hand, the children of unskilled European immigrants were more likely to exit those occupations than the children of unskilled natives. However, the sons of white-collar immigrants were more likely to work in white-collar occupations in adulthood than the sons of white-collar natives. As a result, occupational persistence was on average similar across the two groups.

A number of features make Argentina in the Age of Mass Migration an interesting case study of the economic performance of international immigrants. First, the magnitude of the migration flow relative to the native population was substantial by both historical and contemporary standards. Second, in contrast to immigrants entering the US during this time period and to many immigrants today, immigrants entering Argentina exhibited higher levels of human capital than natives and came from countries that resembled Argentina in terms of average living standards.[3] Hence, this migration episode can shed light on what immigrant assimilation might look like in a setting where immigrants had high human capital and were a fairly numerous group relative to natives. Finally, the opportunity to construct longitudinal data that follow a large number of immigrants and their children enables me to deal with some of the methodological challenges faced by researchers studying immigrant assimilation (Borjas 1985, Abramitzky *et al.* 2014).

[3]In 1869, the literacy rate among males over 18 years old was 26% for natives and 61% for immigrants. I estimated that the typical sending country had a per capita GDP that ranged from 80% to 120% that of Argentina in the 1875-1890 period and about 60% in the 1890-1914 period. In the US, this figure ranged from 40% to 60% in 1875-1890 and 40% to 50% in 1890-1914. Own elaboration based on Dirección General de Inmigración (1925), Ferenczi (1929), Maddison (2007).

Historical context and related literature: Argentina in the Age of Mass Migration

The 1853 Constitution made it a national priority to attract European immigrants to help populate the vast and sparsely populated Argentine territory. In the mind of the elites governing the country, immigrants from Europe – especially those from the north of the continent – were needed to provide a "civilizing influence" that would enable Argentina to grow as a prosperous and free nation (Alberdi 1852).[4]

From 1857 to 1930, Argentina received 6.2 million immigrants from Europe, becoming the second largest receiving country in the period (after the US) and the largest in per-capita terms (Germani 1966). Figure 1 shows the number of yearly arrivals of overseas immigrants to Argentina. Until 1862, the number of yearly arrivals was below 10,000, but started to increase rapidly thereafter. This increase coincided with the unification of the different provincial governments into a single national authority following the Battle of Pavón in 1861. By 1914, the year of the third national census, Argentina's population had grown from less than two million in 1869 to more than eight million, of which 30% were foreign born.

Despite the Argentine elites' desire to attract immigrants from the north of Europe, nearly half of the immigrants were of Italian origin. In 1895, Italian immigrants accounted for 11% of the population of Argentina. Spain was the second most numerous sending country, representing around a third of the total immigration. Although immigrants from France were not as numerous overall, they accounted for a relatively large fraction of the early-arriving immigrants that are the main focus of this paper.

Conventional accounts of the period describe Argentina as a country where "hard working" immigrants had an easy path to upward economic mobility (Alsina 1898). Although this view is also popular among early scholars (Diaz-Alejandro 1970, Conde 1979, Baily 1983), there is little quantitative evidence supporting it. In a series of widely debated studies, Germani (1966) uses the published census tabulations to study the extent of occupational mobility in 19th-century Argentina. His study finds that immigrants moved up the occupational ladder at a faster pace

[4]See Devoto & Benencia (2003) for an overview of the history of immigration to Argentina. See Taylor (1994) for a comparison between mass migration to Argentina and Australia.

than natives. However, a concern with inferring mobility from aggregate data is that the pool of immigrants changes from census to census, either because new immigrants arrive to the country or because some return to their countries of origin. Hence, it is not possible to disentangle changes in the social standing of immigrants from changes in the composition of the immigrant pool. In the case of Argentina, an added difficulty is the lack of information on year of arrival to the country in the 19th-century censuses.

A more recent study by Da Orden (2005) also offers some support for the optimistic view of immigrant social mobility. The author studies the occupational mobility of Spanish immigrants in Mar del Plata, a coastal city in the Province of Buenos Aires. By linking birth and marriage records of the children of immigrants, she is able to observe immigrants' occupation at two different points in time. She documents that, after an average of 26 years, 51% of the immigrants in her sample had moved up in the occupational ladder.

Other studies present a more negative outlook on the possibilities for social progress in 19th-century Argentina. Szuchman (1981) links census records from 1869 to 1895 for immigrants and natives residing in the city of Córdoba. He finds that upward mobility was rare among immigrants, although the author is only able to follow individuals who had stayed in the city of Córdoba until 1895. Sofer (1982) examines the occupational mobility of Eastern European Jewish immigrants in the city of Buenos Aires at the late 19th-century. The author links individuals from the 1895 census to the records of the Chevrah Kedyscha Ashkenazi, the main Jewish association in Buenos Aires. He documents little upward mobility, with most immigrants remaining trapped in unskilled jobs or even experiencing downward mobility.

Existing studies that use individual level data to assess the economic mobility of immigrants suffer from two main limitations.[5] First, these studies focus on either specific immigrant groups or on immigrants living in specific places within Argentina. Second, because economic and geographical mobility are probably associated, limiting the analysis to immigrants who did not change their place of residence is likely to underestimate the economic mobility experienced by the typical immigrant. Below, I show that failing to track internal migrants indeed results in lower estimated rates of occupational mobility among first-generation immigrants.

[5]Other studies on specific immigrant communities include Míguez's 1993 on the Province of Buenos Aires, Otero's 1994 study on French immigrants in the city of Tandil and Tolcachier's 1995 study on Israeli immigrants.

From a methodological point of view, this paper is closely related to Ferrie (1997) and Abramitzky *et al.* (2014). Ferrie (1997) links records of ship arrivals of immigrants to US censuses in the Antebellum period to look at the occupational mobility of immigrants relative to Europe. The author finds that immigrants, in particular those from Britain and Germany, experienced relatively high rates of upward occupational mobility during this time period. Abramitzky *et al.* (2014) use linked census data to study the labor market assimilation of immigrants in the US at the early 20th-century. The authors find that immigrants exhibited similar rates of occupational upgrading as natives.

This paper is also related to a growing body of literature in economic history that uses linked data to study historical migration episodes. In addition to the two articles described above, some other examples include Abramitzky *et al.* (2012; 2013), Boustan *et al.* (2012), Collins & Wanamaker (2014; 2015a), Long & Ferrie (2013), Long (2005), Kosack & Ward (2014) and Salisbury (2014). In contrast to this paper, the focus of this literature has been to study either internal migrations within the US or the UK or international migrations to the US.

Data

Linking the 1869 and 1895 censuses

I constructed a new sample following natives and immigrants across Argentina's national censuses of 1869 and 1895. To do so, I took advantage of the fact that both censuses' handwritten manuscripts are indexed and can be searched through the genealogy website FamilySearch.org.[6] The sample includes males – natives and immigrants – who were of working-age in both census years and males – sons of natives and native-born sons of immigrants – who were observed in their childhood household in 1869 and as adults in 1895.

To construct this sample, I identified two groups of individuals in the 1869 census full count: (1) males 18 to 35 years old, born in either Argentina or one of the six largest European sending countries (England, France, Germany, Italy, Spain and Switzerland), (2) males 0 to 17 years old, born in Argentina, with father present in the household and father born in Argentina or one of

[6]These are the only two national censuses of Argentina for which individual records with names are available. The next national census took place in 1914. Unfortunately, the individual records of this census were lost, so it is not possible to extend the sample ahead in time.

the European countries listed above. These six European countries were the only sending countries with more than 1,000 residents in the relevant age cohort in 1869 Argentina, accounting for more than 95% of all European immigrants at that time.[7] These two groups included a total of 448,201 individuals, of which 58,755 were born in one of the European sending countries included in the analysis and 22,932 were native-born sons of immigrants from these countries.

I then searched the 1895 census full count for a set of potential matches for each of these individuals. Based on the similarity in reported names and (estimated) years of birth, I calculated a linking score ranging from 0 to 1 for each pair of potential matches: Higher scores represented pairs of records that were more similar to each other. Full details on the procedure used to compute the linking scores are provided in the online appendix.

I used these linking scores to inform my decision rule on which records to incorporate into the analysis. To be considered a unique match for an individual in the 1869 census, a record in the 1895 census had to satisfy three conditions: (1) be the record with the highest linking score among all the potential matches for that individual, (2) have a linking score above a minimum threshold ($p_1 > \underline{p}$) and (3) have a linking score sufficiently higher than the second-best linking score ($\frac{p_1}{p_2} > l$).[8]

Because the linking is based on potentially noisy information, there is a trade-off in choosing the cutoff values $\underline{p}$ and l. On the one hand, higher values of $\underline{p}$ and l imply that a larger fraction of true matches will be discarded from the analysis. In addition, individuals who report their identifying information with high accuracy and have more uncommon names – within their place and year of birth – are more likely to be uniquely matched under a more stringent rule. On the other hand, lower values of $\underline{p}$ and l will lead to a larger sample but to a higher share of incorrect matches.

With this trade-off in mind, my baseline results are based on a sample created using a relatively conservative choice of the parameters $\underline{p}$ and l. As a result, my matching rates are lower than those typically found in recent economic history papers using US census data.

The matching rate was 11.6% for sons of natives and 13.6% for sons of immigrants. I was able to uniquely link approximately 9.5% of working-age natives and 10% of working-age immigrants. Table A.1 in the online appendix shows the matching rates that resulted from this linking process,

[7] Because the 1869 census lacks information on relationship to head of household, I used a procedure similar to the one used by IPUMS in order to identify fathers and sons. See the online appendix for further details.

[8] This decision rule is analogous to the one used by Mill & Stein (2012), Parman (2015), Feigenbaum (2016) and Feigenbaum (2017).

disaggregated by country of origin and by age group. I provide a detailed discussion on matching rates and additional sources of match failure in online appendix table A.2.[9]

Once I had completed the linking procedure, I manually digitized the economic outcomes variables using the handwritten census manuscripts available online at FamilySearch.org. In the case of working-age immigrants, children of immigrants and children of natives, I digitized the economic outcome variables for every individual in the linked sample. In the case of working-age natives, I digitized the economic outcomes only for a random sample of the linked individuals. The final sample includes about 6,000 working-age natives, 5,000 working-age immigrants, 18,000 sons of natives and 2,500 native-born sons of immigrants.

Linking passenger lists to the 1895 census

To assess the extent to which immigrants experienced occupational progress relative to Europe, I complement the above data with a sample linking male immigrants arriving to the city of Buenos Aires to the 1895 census. To construct this sample, I started with a sample of 54,036 working-age[10] immigrants who appeared in ship arrival records between 1882 and 1894.[11] These records were originally collected by the National Direction of Migration and have been digitized by CEMLA (*Centro de Estudios Migratorios Latinoamericanos*), a research center in Buenos Aires. Each record contains the name, occupation, date of arrival, port of origin and entry, civil status and age of each passenger on the ship. In this time period, about 75% of immigrants entered Argentina through the port of Buenos Aires (Dirección General de Inmigración 1925).

I then linked these immigrants to the 1895 census using a method analogous to the one described above. I was able to link 3,157 immigrants, which represents approximately a 6% matching rate. Linking these data is more challenging than linking the censuses because immigrants typically declared their original first name upon arrival but adopted a Spanish version of it while in Argentina.[12]

[9] While immigrants are expected to have a lower matching rate because of return migration, they also resided in areas of Argentina with relatively lower mortality levels. For instance, natives born in Buenos Aires and aged 18 to 35 years old exhibited a 25% mortality rate from 1869 to 1895, whereas the average native exhibited a mortality rate above 40% during the same time period. In addition, European immigrants had substantially higher literacy levels, implying that their identifying information is likely to be reported with less error.

[10] To be included in the sample, an individual had to be: (1) aged 18 to 60 years old upon arrival (2) 60 years old or less by 1895.

[11] Passenger lists started to be systematically collected following the 1876 Immigration Law, but all the individual records prior to1881 have been lost. Then, unlike Ferrie (1997), I am unfortunately unable to observe an immigrant both in the 1869 and 1895 censuses and in the passenger lists.

[12] Because some of the name changes are predictable – for instance, Italian immigrants named *Giuseppe* adopted

Table A.3 in the online appendix shows the number of individuals and linked individuals in this sample, by country of birth.

Comparing the linked samples to the population

Given the absence of numerical identifiers in the data, names provide the most important source of information in the linking procedure. The dependence on names could lead to a biased sample if having a name that is both uncommon and accurately recorded is correlated with social and economic characteristics. In this subsection, I compare individuals in the linked samples to individuals in the cross-sectional data. I provide further details on this comparison in the online appendix.

The online appendix tables A.4 to A.7 show the results of these comparisons. In these tables, I compare natives and immigrants in the linked sample to natives and immigrants in the 1869 and 1895 census cross-sections, based on the nationally representative census samples compiled by Somoza (1967). Overall, the evidence suggests some degree of positive selection of individuals into the linked census sample. First, the white-collar occupational category tends to be overrepresented in the linked sample, with the fraction of white-collar workers being 2 percentage points higher than in the cross-section among native-born working-age individuals in the 1895 census. Second, individuals in the linked sample were 10 percentage points more likely to own property and 9 percentage points more likely be literate.

In table A.8 in the online appendix, I compare immigrants in the passenger lists who were matched to an observation in the 1895 census to immigrants who were not, as well as immigrants in the linked sample in 1895 to immigrants in the 1895 cross-section. The main difference between the linked sample and the cross-section is that immigrants from Spain are overrepresented in the linked sample. For instance, while only 20% of immigrants in the passenger lists were from Spain, 35% of immigrants in the linked sample were from this country. This overrepresentation likely reflects the fact that immigrants from Spain did not change their names upon arrival to Argentina, making it easier to find them in the 1895 census.

the name *José* –, I am able to partially address this issue by performing the linking based on a Spanish version of the first name. Further details are provided in the online data appendix.

Occupations and earnings data

Similar to US national censuses of the period, 19th-century Argentine censuses lack information on individual-level earnings or income. I dealt with this feature of the data using two standard approaches in the literature. First, following Abramitzky *et al.* (2012; 2014) and Collins & Wanamaker (2014), I constructed a measure of typical earnings by occupation.[13] Second, I classified the occupational titles into occupational categories.

To create the occupational earnings measure, I employed information from a variety of historical sources. First, I used information on daily wages in blue-collar occupations in the city of Buenos Aires from Buchanan (1898). Second, I used the published census volumes to construct estimates of earnings in the commercial and industrial sectors. Third, I used the congressional reports of Correa & Lahitte (1898) to estimate earnings in the farming sector. Table A.9 in the online appendix shows the sources of income data used in the analysis. Further details on the construction of this earnings measure, including more information on the sources and assumptions are provided in the online appendix.

It is worth emphasizing that constructing a measure of typical earnings is challenging, especially for self-employed individuals. As a result, I conducted a number of sensitivity checks to assess the robustness of the results that rely on the occupational score. In particular, I paid special attention to the issue of assigning earnings to farmers.

I also classified the more than 100 occupational titles in my sample into broad occupational categories.[14] To do so, I first assigned each occupation a code from the Historical International Classification of Occupations (HISCO). This classification is based on the International Standard Classification of Occupations (ISCO) and has been adapted to deal with historical data. I then mapped each HISCO code to an occupational category using the Historical International Social Class Scheme (HISCLASS), developed by Leeuwen *et al.* (2002). Finally, I collapsed the HISCLASS scheme into four broad categories following Long & Ferrie (2013): white-collar (HISCLASS 1-5), farmer (HISCLASS 8), skilled/semi-skilled (HISCLASS 6-7,9) and unskilled (HISCLASS 10-12).

[13] Abramitzky *et al.* (2014) use median wages by occupation constructed from the 1950 US census and mean wages from the Cost of living Survey of 1901 to construct occupational earnings in the 1900, 1910 and 1920 censuses. Collins & Wanamaker (2014) constructs occupational earnings by adjusting industry wages by demographic characteristics and location.

[14] This approach is followed by Ferrie (1997; 1999), Long & Ferrie (2013), Abramitzky *et al.* (2014), among other researchers.

Table 1 shows the ten most common occupations for working-age individuals in 1895 and their corresponding occupational category, reported separately for natives and immigrants in the linked census sample.

There are three limitations associated with using occupations rather than earnings to measure assimilation. First, I am not able to capture changes in an individual's social standing that stem from changes in their economic status within a given occupation. Second, the occupational earnings measure explicitly fixes the ranking of occupations in its 1895 level. As a result, the measure is unable to capture changes in economic status that occur due to changes in the distribution of income across occupations over time.[15] Third, because the censuses lack a question on employment status, I cannot distinguish employed from unemployed individuals.

The first generation

Occupational mobility of natives and immigrants

Panel (a) in table 2 shows a transition matrix for natives, while panel (b) shows this same matrix for immigrants. Each element of these matrices represents the fraction (number) of individuals working in occupational category i in 1869 that worked in occupational category j in 1895. The last row in both matrices shows the distribution of individuals across occupational categories in 1895.

The first noticeable difference is that immigrants and natives were concentrated in different occupational groups. Farming was considerably more prevalent among natives than among immigrants, both in 1869 and in 1895. This difference is consistent with the different propensities of natives and immigrants to locate in urban areas. In 1869, 63% of the foreign-born resided in urban locations, whereas this proportion was only 28% among natives.

In panel (c) of table 2, I present a number of summary measures of mobility based on the transition matrices. The simplest measure of mobility is the fraction of individuals who switched occupational categories across years, that is the fraction of individuals off the main diagonal of the matrix. This statistic suggests that there was more occupational mobility among immigrants than

[15] Butcher & DiNardo (2002) argue that, because immigrants and natives might exhibit different skill levels, changes in the returns to skills will result in changes in immigrants' relative economic standing, even in the absence of "true" assimilation.

among natives, with 59% of immigrants and 51% of natives switching occupational categories from 1869 to 1895.

Immigrants also appear to exhibit more mobility out of the unskilled category: 72% of those initially in unskilled occupations had moved out of this category by 1895, compared to 64% among natives. The typical path out of the unskilled category was different for the two groups. Natives usually left the unskilled category by switching into farming, whereas immigrants moved into more urban occupations such as white-collar and skilled/semi-skilled jobs. While only 8% of natives in unskilled occupations in 1869 worked in white-collar jobs by 1895, this fraction is considerably higher (23%) among immigrants.

As discussed in Long & Ferrie (2013), a key shortcoming of using this simple measure to compare mobility across two matrices is that the measure does not distinguish whether differences in mobility are due to: (1) differences in the distribution of occupations across the two matrices or (2) differences in the strength of the association of the rows and columns in each of the two matrices. This distinction is important in this setting because, as discussed above, the distribution of occupations among natives differed markedly from the distribution among immigrants.

To establish whether the row-column association was stronger (i.e. there was less occupational mobility) among natives or among immigrants, I followed Long & Ferrie (2013) in completing the following steps. First, I computed the Altham (1970) statistic $d(P, Q)$, which measures the difference in the strength of the row-column association in matrices P – the mobility matrix of natives – and Q – the mobility matrix of immigrants.[16] Higher values of $d(P, Q)$ imply greater differences in this association, but are not informative regarding which of the two matrices exhibit more mobility. Next, I calculated $d(P, J)$ and $d(Q, J)$, which measure this same difference but relative to a matrix J representing full independence (a matrix of ones). Higher values of $d(P, J)$ or $d(Q, J)$ imply greater departures from independence, or less mobility.

Panel (c) of table 2 shows that $d(P, Q)$ is significantly different from zero, suggesting that the strength of the row-column association is different in the two matrices. In addition. the departure from independence is larger for natives than for immigrants –$d(P, J) > d(Q, J)$. Taken together,

[16]The Altham (1970) statistic is based on the relative odds of individuals in different occupations in 1869 find a given a job in 1895. Under conditions of perfect mobility, the relative odds are one: an occupation does not provide any relative advantage in obtaining a given occupation. More generally, given two matrices P and Q, the Altham statistic $d(P, Q)$ measures the difference in the strength of their row-column association. Importantly, it is possible to perform a likelihood-ratio test to assess whether this difference is significantly different from zero.

these results suggest that there was higher occupational mobility among immigrants than among natives.[17]

One key limitation of the existing studies of immigrant mobility in Argentina is the inability to track internal migrants. Table B.1 in the online appendix shows that this inability leads to lower rates of estimated occupational mobility among immigrants. In this table, I divide the sample of immigrants into "movers" and "stayers." Individuals are classified as "movers" if by 1895 they lived outside of their 1869 department of residence. Panels (a) and (b) show an occupational mobility matrix computed separately for stayers and movers, respectively. Panel (c) of this table shows that stayers were less likely to switch occupational categories than movers (48% versus 65%). Similarly, the Altham statistic described above indicates that the mobility matrix for stayers exhibits a greater departure from independence that the mobility matrix for movers.

Occupational earnings regressions

Next, I use the occupational earnings data to compare the rates at which natives and immigrants moved into higher paying occupations. In particular, I estimate the following model of occupational earnings:

$$\begin{aligned}\log(Occupational\,Earnings_{it}) = \beta_0 + \beta_1 Immigrant_i + \beta_2 Year1895_t \\ + \beta_3 Immigrant_i \times Year1895_t + \gamma X_{it} + \epsilon_{it}\end{aligned} \tag{1}$$

where $Occupational\,Earnings_{it}$ is the daily occupational earnings of individual i in year t, $Immigrant_i$ is an indicator variable of whether the individual was foreign-born, $Year1895_t$ is an indicator of whether the observation belongs to the 1895 census and X_{it} is a vector of individual-level characteristics. In the baseline specification, X_{it} is limited to a quartic in age. I restrict the sample to working-age individuals, defined as those being at least 18 years old and at most 35 years old in 1869, and to those with a reported occupation in both census years. Because matching rates are not constant across sending countries, in the baseline specification each observation is weighted in order to reflect the country of birth distribution in 1895 Argentina.

There are two main coefficients of interest in equation 1: β_1, which captures baseline differences

[17] I also computed $d(P,Q)^i$, which measures the row-column association in matrices P and Q while excluding the elements in the main diagonal of the matrix. Using this alternative measure, I also find higher mobility among immigrants than among natives, although the difference is now less stark.

in the occupational earnings of natives and immigrants, and β_3, which captures differences in their occupational earnings growth.

This specification differs from the standard in the immigrant assimilation literature, which uses years since migration as the independent variable of interest (Abramitzky *et al.* 2014, Borjas 1985, Chiswick 1978, Lubotsky 2007). It is not possible to estimate the standard specification in this context because the censuses lack information on year of arrival to Argentina. This lack of information also prevents me from following cohorts of immigrants over time, as in Borjas (1985) and Minns (2000). Note, however, that estimating this regression on the panel data enables me to keep the composition of the sample constant across census years. In doing so, I am able to disentangle changes in the social standing of immigrants from changes in the composition of the immigrant pool.[18]

The first column of table 3 presents the results of the baseline specification. This specification suggests that natives had relatively higher occupational earnings in 1869, but that immigrants upgraded their occupations faster. In particular, the growth in occupational earnings among immigrants was 6% faster than among natives. This evidence is consistent with the finding of immigrants' higher rates of movements out of the unskilled category documented in the previous subsection.

Explaining the differences between immigrants and natives

I explore two main alternative hypotheses; other than labor market assimilation,;that could explain the faster growth in occupational earnings among immigrants. First, I test whether this finding could be attributed to the greater propensity of immigrants to locate in areas of Argentina that were experiencing faster progress, particularly urban areas. Second, I test whether the finding is driven by an overall increase in returns to skill that disproportionately benefited immigrants, who had on average higher human capital levels than natives.

To explore the first possibility, I perform two different exercises. First, in column 2 of table 3, I restrict the sample to individuals who in 1869 resided in the provinces of Buenos Aires (including the city of Buenos Aires), Entre Ríos and Santa Fe. These three provinces hosted more than 95%

[18]Since it is not possible to follow cohorts of immigrants over time, the composition bias in this context also differs from the one that would arise when using US data. In particular, a comparison of my results with those estimated in the repeated cross-section would not be informative about the selection of return migrants, but rather about the *net* change in the composition of the immigrant pool from 1869 to 1895.

of the European immigrants in 1869. When restricted to this set of provinces, I find that European immigrants performed worse than natives in the baseline year, but still exhibited higher relative occupational earnings growth. Second, in column 3 of table 3, I include department of residence fixed effects and an interaction between department of residence fixed effects and a 1895 census year indicator. That is, I compare immigrants residing in the same departments and allow the department of residence effects to differ based on the census year. The results are similar to the ones that I obtain in the previous specification.[19] Finally, the results (not reported) are similar if I restrict the sample to: (1) individuals residing in urban locations in 1869 or to (2) individuals residing in urban locations both in 1869 and in 1895. Hence, the evidence suggests that immigrants experienced faster growth in occupational earnings also *within* urban areas.

As noted in the introduction, European immigrants had higher human capital levels than natives, as captured by literacy rates. Accordingly, the higher growth in occupational earnings among immigrants might reflect a general increase in the returns to skill from 1869 to 1895 rather than assimilation. To test this possibility, in column 4 of table 3 I estimate a version of equation 1 in which I include two additional controls: a literacy indicator and an interaction between a literacy indicator and a 1895 census year indicator. Again, I find faster occupational earnings growth among immigrants than among natives.

In the last column of table 3, I test whether the above explanations combined could account for immigrants' relatively faster growth in occupational earnings. To do so, I include both the interaction between literacy and the 1895 census year indicator and the department of residence fixed effects interacted with the 1895 census indicator, as well as the main effects of both groups of variables. The evidence still suggests faster occupational upgrading among immigrants than among natives.

Heterogeneity by sending country

Immigrants from different sending countries differed both in terms of their human capital levels and in terms of their cultural and linguistic similarity to natives. Hence, the assimilation experience of

[19] The results in this specification should be interpreted with caution, as place of residence is an endogenous choice. Indeed, geographic mobility might be a strategy for occupational upgrading. I note, however, that the results (not reported) are also similar if I instead interact the department of residence fixed effects based on place of residence in 1869 with the 1895 year indicator.

the average immigrant might mask differences across sending countries. To explore this possibility, I estimate a version of equation 1 in which I include an indicator variable for each of the sending countries included in my sample, as well as an interaction of each of these country indicators with an 1895 census indicator. This specification captures differences both in country-specific baseline occupational earnings and in country-specific occupational earnings growth.

In figure 2, I plot the coefficients corresponding to each of the countries around a 95% confidence interval. Upon arrival, immigrants from every major sending country (with the exception of Italy) appear to do better than natives, although the difference with respect to natives is not statistically significant for French immigrants. The evidence is broadly consistent with immigrants from countries with higher levels of average human capital doing better upon arrival. In particular, the ranking of countries based on average occupational earnings mostly matches the ranking of countries based on the average literacy of immigrants in Argentina. In addition, the evidence suggests that immigrants from every major sending country (with the exception of Switzerland) experienced faster occupational upgrading than natives. Note, however, that the interaction between the country of origin indicator and the 1895 census indicator is not statistically significant for the English or the German immigrants.

Alternative specifications and robustness

In this subsection, I show that the finding of immigrants' higher occupational earnings growth relative to natives is robust to: (1) how earnings are assigned to farmers, (2) using an alternative measure of occupational status based on access to property as the dependent variable and (3) the procedure used to create the linked sample.

Assigning an earnings measure to farmers is challenging for a variety of reasons. First, it is hard to distinguish in the census between owners and operators of farms and farm employees. Although the 1895 census contains a question on whether the person holds real estate property, which could be useful for distinguishing the two, this question is not available in the 1869 census. In addition, the farming sector encompasses a wide range of economic realities, ranging from small farms to large-scale production. Hence, relying on typical earnings is more challenging than in other occupations where within-occupation earnings dispersion is likely to be less prominent. While my baseline measure of occupational earnings introduces a distinction based on the reported occupational title

between small (who I estimate to be close to the median in terms of occupational earnings) and large farmers (who are at the top of the income distribution) and also incorporates farm laborers as a separate category, this distinction is most likely not sufficiently rich to capture the different realities within the farming sector.

In column 1 of table 4, I exclude individuals who were employed as farmers in either 1869 or 1895 from the sample. The table shows that the finding of higher occupational upgrading for immigrants remains unchanged in this case. However, I now find that immigrants outperform natives even upon arrival. Despite the challenges in measuring occupational earnings for farmers, excluding them could introduce biases if immigrants and natives exhibit differential rates of movement into and out of farming. In the context of Argentina, immigrants' occupational upgrading relative to natives is exaggerated once farmers are excluded, as moving into farming was a more frequent avenue for upward mobility for natives than for immigrants, as shown above.

Another concern with the finding of faster occupational earnings growth among immigrants is that it might be sensitive to how the earnings measure was constructed. If my measure systematically overestimated earnings in occupations to which immigrants were more likely to transition, I would find that immigrants exhibited faster occupational upgrading than natives. As an alternative approach, I computed a measure of occupational status based on access to property. The 1895 census includes the question "*¿Posee propiedad raíz?*" – "Do you own real estate property?."[20] I used this information to compute an alternative measure of occupational status: the fraction of individuals within a given occupation that owned real estate property in 1895. For instance, this measure takes a value of 0.2 for *jornaleros* (day laborers) and a value of 0.7 for *comerciantes* (storekeepers), which means that 20% of *jornaleros* and 70% of *comerciantes* in my sample held real estate property in 1895.[21] Column 2 in table 4 shows that the finding of faster occupational upgrading among immigrants is robust to using the log of this measure as the dependent variable.

Finally, I assess the robustness of the results to the linking procedure used to create the sample. There are two main concerns related to the linking procedure. First, the linked samples are not fully representative of the population. Second, the fraction of false positives might be higher among immigrants than among natives.

[20] As the 1869 census lacks a question on access to property, I cannot directly use the property variable as the outcome.

[21] The correlation between this measure and my measure of occupational earnings is of about 0.7 in the data.

To alleviate the first concern, in column 3 of table 4 I show that my results are similar when I reweight the sample to account for differences in observable characteristics (in addition to country of birth) with respect to the census cross-section.[22] This evidence suggests that selection into the linked sample, at least based on observable characteristics, is unlikely to drive the results. In addition, column 4 shows that the results are also similar when I do not reweight the sample to account for differences in matching rates across sending countries.

The second concern is that the fraction of false positives in the linked sample might be higher among immigrants than among natives. If that were the case, relying on a linked sample would mechanically overestimate the extent of mobility among immigrants; although not necessarily the extent of *upward* mobility. While this possibility cannot be fully ruled out , I can replicate my analysis in a sample where this issue is likely to be less prevalent: immigrants with infrequent names. In particular, I re-estimate equation 1 using immigrants with whose names fall below the median in the first name frequency distribution within their country of birth, while keeping the full sample of natives. Column 5 shows that the finding of faster occupational upgrading is robust to excluding this group of immigrants from the sample.

Finally, in column 6 I replicate the analysis but focusing on the sample of immigrants whose identifying information matches perfectly, while again keeping the full sample of natives.[23] Overall, this evidence suggests that the result of faster occupational upgrading is unlikely to be driven by features of the linking procedure.[24] Note, however, that restricting the sample to immigrants with uncommon names (which likely exacerbates positive selection of immigrants into the linked sample) causes immigrants to perform better than natives in the baseline year.

Occupational mobility from arrival to 1895

In the previous subsection, I documented the extent to which immigrants moved up in the occupational ladder as they spent time in Argentina. Yet, another question is whether immigrants

[22]To compute the sample weights, I pool the 1895 census cross-section and the linked sample and estimate a probit model of the probability of being an observation in the linked sample. I then reweight my sample by the inverse of this linkage probability.

[23]I define a perfect match as one in which both the first name and the last name agree perfectly, but I allow the year of birth to differ by at most one year. Because the two censuses took place in different times of the year, the difference in estimated year of birth could be one even if an individual accurately reported his age in both censuses.

[24]The results are also similar if I (1) keep only natives with infrequent names and keep all the immigrants, (2) keep only natives who are perfect matches and keep all the immigrants, (3) drop both immigrants and natives with common names, (4) drop both immigrants and natives who are not exact matches (results not reported).

were able to progress relative to their pre-migration occupations. To answer this question, I use the sample following immigrants from their arrival to the city of Buenos Aires to the 1895 census. In this sample, I observe an individual's occupation in Europe as declared upon arrival and his occupation in 1895.[25]

Table 5 shows a transition matrix in which rows represent occupations in Europe and columns represent occupations in the 1895 census. Overall, about 80% of immigrants who entered Argentina from 1882 to 1894 switched their occupational category by 1895. Yet, the data strongly reject independence between an occupation upon arrival and an occupation in 1895 (p-value< 0.01).

Given the absence of a comparable full ranking of occupations in Argentina and each of the sending countries, it is not possible to assess the fraction of immigrants that downgraded or upgraded their occupations relative to Europe. However, under the assumption that the unskilled category is the least desirable, the data show that occupational upgrading occurred for a large fraction of those who had held unskilled occupations in Europe; less than 25% of those who entered the country as unskilled workers were still in those occupations by 1895. The fraction of immigrants moving out of unskilled occupations is substantially higher than documented by Ferrie (1997) in the Antebellum US. Ferrie (1997) finds that about half of the immigrants arriving to the US in the 1840-1850 period were still working as unskilled workers by 1860. Indeed, even when excluding Irish immigrants, who had the worst outcomes among all immigrant groups, from the US data, the evidence suggests higher rates of upward occupational mobility in Argentina.[26]

Another contrast relative to the findings of Ferrie (1997) is that immigrants who declared skilled/semi-skilled occupations were quite unlikely to experience occupational downgrading; only 14% of those initially in the skilled/semi-skilled category were working in unskilled occupations in 1895. In the US case, the fraction experiencing downgrading was considerably higher; as high as 30% among skilled workers. Overall, this evidence suggests that immigrants entering Argentina benefited from relatively better opportunities for occupational upgrading than those entering the

[25] A limitation of these data is that the accuracy of the occupation declared upon arrival has been questioned. The main issue is that immigrants might have answered their intended occupation in Argentina rather than their last occupation in Europe. According to Devoto & Benencia (2003), the most likely bias is that immigrants declared occupations that they deemed would be perceived as desirable by the Argentine authorities. If immigrants indeed exaggerated the quality of their occupations in Europe upon arrival, then the rates of upward occupational mobility that I document would likely be a lower bound.

[26] For instance, the fraction of British and German immigrants moving out of unskilled occupations was about 60% in the US (Ferrie 1997).

US.[27]

Similar to Ferrie's (1997) and consistent with historical accounts of the period (Germani 1966), I find that those who declared farming occupations upon arrival were fairly unlikely to work as farmers in Argentina, at least in the short run. Indeed, about 66% of these immigrants resided in urban locations by 1895 (35% in the city of Buenos Aires) compared to 72% and 40% among all immigrants in the sample. This finding confirms that, for many European immigrants, the Transatlantic move was accompanied by a shift from rural to urban occupations (Alonso 2007).

The second generation

Intergenerational occupational mobility

In this section, I compare the intergenerational occupational mobility of the native-born sons of immigrants – the "second generation" – and the sons of natives. Panels (a) and (b) of table 6 show an intergenerational mobility matrix for sons of natives and for sons of immigrants, respectively. Rows in each of these matrices represent a father's occupation in 1869 and columns show the occupation of his son in 1895. Each element of these matrices represents the fraction (number) of fathers in occupational category i in 1869 whose sons worked in occupational category j in 1895.

The first noticeable difference between the two groups is that movements out of unskilled occupations were considerably more frequent among the children of immigrants. About half of the sons of natives whose father worked in unskilled occupations themselves worked in these occupations, compared to only 25% among the children of immigrants. The opposite pattern occurs in the white-collar category, where persistence was much prominent among sons of immigrants than among sons of natives.

In panel (c) of table 6, I follow Collins & Wanamaker (2015b) in reporting a counterfactual exercise simulating the distribution of sons of natives across occupational categories had they been exposed to the transition matrix of the sons of immigrants. This exercise reveals two main findings. First, the fraction of sons of natives in unskilled occupations would have been substantially lower under this counterfactual: 17% versus 28%. Second, the fraction in white-collar occupations would

[27] In table B.2 in the online appendix, I show that this pattern was similar for Italian and Spanish immigrants, the two largest sending countries. The sample size does not allow me to perform a similar comparison for the other sending countries.

have been almost twice as high: 33% versus 17%.[28] Taken together, these results suggest that the children of European immigrants faced better prospects for mobility out of unskilled jobs than the children of natives.[29]

Panel (d) of table 6 provides summary measures of mobility analogous to the ones presented in section . The table shows that the fraction of sons switching occupational categories is almost identical in both groups: About 55% of sons are in a different occupational category relative to their father. Yet, using the Altham measure of mobility discussed in section reveals slightly *higher* mobility among sons of natives than among sons of immigrants.

The economic performance of the second generation

Next, I compare the economic outcomes of the native-born sons of immigrants and the sons of natives. I begin by estimating the following model:

$$Son\,Outcome_i = \beta_0 + \beta_1 Second\,Generation_i + \gamma X_i + \epsilon_i \tag{2}$$

where $Son\,Outcome_i$ is an economic outcome of the son in father-son pair i, $Second\,Generation_i$ is an indicator variable that takes a value of one for native-born individuals whose father was an immigrant, X_i is a vector of household characteristics (limited to a quartic in son's and father's age in the baseline specification), and ϵ_i is an error term, clustered at the level of the father to account for the presence of siblings in the data.

The key difficulty in estimating the above equation is that the censuses do not include a question on parental place of birth. I hence infer this information by linking males from the 1895 census to their childhood household in 1869, where they were observed living with their parents. The sample is restricted to native-born males 26 to 44 years old in the 1895 census, that is, those who had already been born and that were young enough to still be living in their childhood household in 1869.

I first compare the literacy rate of second-generation immigrants to that of the children of natives. A son is defined as being literate if in 1895 he answered affirmatively to the question:

[28] I observe a similar pattern when I restrict the sample to individuals initially residing in the Provinces of Buenos Aires, Entre Ríos and Santa Fe, where most European immigrants and their children resided. See table B.3 in the online appendix.

[29] Feigenbaum (forthcoming) documents a similar patter of higher mobility among the grandchildren of immigrants in the early 20th-century Iowa.

"Do you know how to read and write?". The question does not specify whether the individual was literate in Spanish or in another language. In column 1 of table 7, X_i includes only a quartic in father's and son's age. This specification shows that the sons of immigrants were substantially more likely to be literate as adults than the sons of natives, an advantage above 35 percentage points relative to a mean literacy rate of about 50% among natives.

As first-generation immigrants themselves were more likely to be literate than natives, one simple explanation for the above finding is the transmission of human capital from parents to their children. I test this hypothesis in column 2, where I control for parental literacy and also include indicators corresponding to the occupational category of the father in 1869.[30] The point estimate goes down by about a third, suggesting that differences in parental background – beyond place of birth – play a role in explaining the advantage of the second generation. However, the difference remains large (about 23 percentage points) and statistically significant.

An alternative explanation for the observed difference in literacy rates is that, as discussed above, immigrants were more likely to reside in urban locations than natives. If, for instance, urban locations were better served by schools, then the children of immigrants would have been more likely to acquire formal education and thus be literate as adults. Moreover, the children of immigrants were disproportionately located in the relatively more prosperous areas of Argentina, the Littoral region. To explore this possibility, I estimate equation 2 adding a full set of enumeration district fixed effects, based on the family place of residence in 1869. In addition to controlling for the urban/rural status of the family's place of residence, these fixed effects control for other observable and unobservable factors that vary at the enumeration district level and that might have influenced literacy rates in adulthood.[31] The results in column 3 again show that second-generation immigrants did better in terms of literacy, an advantage of about 15 percentage points, even when compared to sons of natives who grew up in similar locations.

Columns 4 to 6 in table 7 repeat the above analysis using the log of occupational earnings as the outcome variable. The evidence also suggests a substantial advantage of the children of immigrants relative to the sons of natives. This difference is robust to controlling for father's literacy and occupational status in 1869 (in column 5), as well as for observable and unobservable enumeration

[30] The results are similar if instead of occupational category indicators I include occupational earnings as a control variable.

[31] There were 3,045 different enumerators in the 1869 census, each covering about 600 individuals.

district characteristics (in column 6).

In columns 7 to 9, the dependent variable is an indicator that takes a value of one if the individual held real estate property in 1895. Interestingly, I do not observe an advantage of the sons of immigrants in terms of access to property as adults in the baseline specification (column 7) or when controlling for father's literacy and occupational category (column 8). However, second-generation immigrants also do better in terms of access to property, an advantage of 3.2 percentage points relative to a mean of 33% among sons of natives, when I include enumeration district fixed effects (column 9). This pattern is consistent with the fact that property was more difficult to access in the relatively more urban and prosperous locations where European immigrant families resided in Argentina.

I next explore how the children of immigrants from different sending countries fared as adults. Figure 3 shows the results of estimating equation 2, now using separate country of origin indicators instead of a single second-generation indicator. In each of the figures, I plot the coefficient corresponding to each of the countries around a 95% confidence interval. Panel (a) shows the results for literacy, panel (b) for log occupational earnings and panel (c) for access to property. In the three cases, I control for father's literacy and occupational category in 1869, as well as for enumeration district fixed effects as determined by the family place of residence in 1869. The figures show that the children of immigrants from every major sending country fared on average better as adults than the children of natives. The differences relative to the sons of natives are especially prominent with respect to literacy and occupational earnings. Overall, these results suggest that the advantage of second-generation immigrants was not confined to any particular sending country.

How can we account for the advantages of second-generation immigrants? One possible interpretation is that the children of immigrants benefited from higher "ethnic capital" than the children of natives (Borjas 1992). According to the "ethnic capital" hypothesis, the human capital of a child is influenced not only by the human capital of her own parents, but also by the average human capital of the ethnic group to which the child belongs. As immigrants had on average higher human capital than natives, their children might have been able to take advantage of the higher human capital within their ethnic network. In table B.4 in the online appendix, I formally test this possibility by re-estimating equation 2 including the average literacy rate of a child's father's country of birth, as measured among immigrants from that country in Argentina in 1869– as an

additional variable. The table shows that, conditional on father's background characteristics, there is a positive association between the average literacy of the ethnic group in the father's generation and literacy rates and occupational earnings of the second-generation. Note that this evidence should be interpreted as only suggestive because of the well-known "reflection problem" (Manski 1993) in measuring social interactions.

Conclusions

I provided evidence on the mobility and economic outcomes of European immigrants and their children in 19th-century Argentina. To do so, I constructed longitudinal data linking males across the two national censuses of population of the period and passenger ship lists of immigrant arrivals to the city of Buenos Aires.

I first looked at the labor market assimilation of first-generation immigrants. I found that immigrants exhibited faster occupational upgrading than natives. This finding does not seem to be driven by immigrants from any particular origin. Rather, I documented that immigrants from most of the main European sending countries outpaced natives in terms of occupational upgrading. Immigrants also appear to have in many cases upgraded their pre-migration occupations: Most of those entering the country as unskilled workers upgraded their occupation after a relatively short period of time. A comparison with similarly constructed US data suggests that European immigrants in Argentina experienced relatively higher rates of upward occupational mobility.

Focusing on second-generation immigrants, I documented that the sons of European immigrants from every major sending country outperformed the sons of natives along a number of important economic outcomes. In addition, the majority of the sons of unskilled immigrant workers were able to upgrade their occupations, in many cases to white-collar jobs.

Overall, my findings give further support to an optimistic view of the economic performance of European immigrants in 19th-century Argentina. Furthermore, I showed that this optimism also largely extends to the second generation. These findings provide an economic rationale of why some Europeans chose to migrate to Argentina instead of to the US. While real wages were likely lower in Argentina (Williamson 1995), the prospects for upward economic mobility might have been better.

In comparing the experiences of Italians in Argentina and in the US, Klein (1983) argues that

"The sharp differences in the Italian immigrant experience within Argentina and the United States were fully perceived by both the immigrants themselves and virtually all contemporary observers." But why did these differences exist? My paper cannot offer a definite answer to this question. One plausible interpretation points to differences in the social structure of the two countries at the onset of Mass Migration. When the Age of Mass Migration started, Argentina had a relatively small population and no significant skilled working class. In contrast, the US featured a larger and more skilled workforce (Baily 1983). Hence, immigrants moving to Argentina were in a better position to become a dominant force of the economy than those going to the US.

References

Abramitzky, Ran, Boustan, Leah Platt, & Eriksson, Katherine. 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.

Abramitzky, Ran, Boustan, Leah Platt, & Eriksson, Katherine. 2013. Have the poor always been less likely to migrate? evidence from inheritance practices during the age of mass migration. *Journal of development economics*, **102**, 2–14.

Abramitzky, Ran, Boustan, Leah Platt, & Eriksson, Katherine. 2014. A nation of immigrants: Assimilation and economic outcomes in the age of mass migration a nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of political economy*, **122**(3).

Alberdi, Juan Bautista. 1852. *Bases y puntos de partida para la organización política de la república argentina*.

Alonso, Blanca Sánchez. 2007. The other europeans: immigration into latin america and the international labour market (1870–1930). *Revista de historia económica/journal of iberian and latin american economic history (second series)*, **25**(03), 395–426.

Alsina, Juan A. 1898. La inmigración europea en la república argentina...

Altham, Patricia ME. 1970. The measurement of association of rows and columns for an r× s contingency table. *Journal of the royal statistical society. series b (methodological)*, 63–73.

Álvarez, Beatriz, & Nicolini, Esteban A, Esteban Alberto. 2010. Income inequality in the north-west of argentina during the first globalization. methodology and preliminary results. *Ponencia presentada en ii encuentro anual de la asociación española de historia económica. madrid*, **8**.

Argentina, Dirección General de Estadistica. 1895. *Censo de los empleados administrativos, funcionarios judiciales personal docente de la república argentina correspondiente al 31 de diciembre de 1894*. Compañia sud-americana de billetes de banco.

Baily, Samuel L. 1983. The adjustment of italian immigrants in buenos aires and new york, 1870-1914. *The american historical review*, 281–305.

Borjas, George J. 1985. Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of labor economics*, 463–489.

Borjas, George J. 1992. Ethnic capital and intergenerational mobility. *Quarterly journal of economics*, **107**(1), 123–150.

Boustan, Leah Platt, Kahn, Matthew E, & Rhode, Paul W. 2012. Moving to higher ground: migration response to natural disasters in the early twentieth century. *The american economic review*, **102**(3), 238–244.

Buchanan, William I. 1898. La moneda y la vida en la república argentina. *Cuadernos del cish*, **3**.

Butcher, Kristin F, & DiNardo, John. 2002. The immigrant and native-born wage distributions: Evidence from united states censuses. *Industrial & labor relations review*, **56**(1), 97–121.

Chiswick, Barry R. 1978. The effect of americanization on the earnings of foreign-born men. *Journal of political economy*, **86**(5), 897–921.

Collins, William J, & Wanamaker, Marianne H. 2014. Selection and economic gains in the great migration of african americans: New evidence from linked census data. *American economic journal: Applied economics*, **6**(1), 220–252.

Collins, William J, & Wanamaker, Marianne H. 2015a. The great migration in black and white: New evidence on the selection and sorting of southern migrants. *The journal of economic history*, **75**(04), 947–992.

Collins, William J, & Wanamaker, Marianne H. 2015b. Up from slavery? intergenerational mobility in the shadow of jim crow.

Conde, Roberto Cortés. 1979. El progreso argentino: 1880-1914.

Correa, A, & Lahitte, Emilio. 1898. Investigación parlamentaria sobre agricultura, ganadería, industrias derivadas y colonización. *Anexo b, buenos aires*.

Da Orden, M.L. 2005. *Inmigración española, familia y movilidad social en la argentina moderna: una mirada desde mar del plata (1890-1930)*. Colección La Argentina plural. Editorial Biblos.

de la Fuente, Diego Gregorio. 1898. *Segundo censo de la república argentina, mayo 10 de 1895*. Vol. 3. Taller tip. de la Penitenciaria nacional.

Devoto, Fernando, & Benencia, Roberto. 2003. *Historia de la inmigración en la argentina*. Sudamericana.

Diaz-Alejandro, Carlos F. 1970. Essays on the economic history of the argentine republic.

Dirección General de Inmigración. 1925. *Resumen estadístico del movimiento migratorio en la república argentina: años 1857-1924*. Talleres Gráficos del Ministerio de Agricultura de la Nación.

Dorfman, Adolfo. 1942. *Evolución industrial argentina*. Losada.

Feigenbaum, James J. 2016. Intergenerational mobility during the great depression.

Feigenbaum, James J. 2017. *Automated census record linking: A machine learning approach*.

Feigenbaum, James J. forthcoming. Multiple measures of historical intergenerational mobility: Iowa 1915 to 1940. *Economic journal*.

Ferenczi, Imre. 1929. International migration statistics. *Pages 47–76 of: International migrations, volume i: Statistics*. NBER.

Ferrie, Joseph P. 1997. The entry into the us labor market of antebellum european immigrants, 1840–1860. *Explorations in economic history*, **34**(3), 295–330.

Ferrie, Joseph P. 1999. Yankeys now: Immigrants in the antebellum us 1840-1860.

Germani, Gino. 1966. Mass immigration and modernization in argentina. *Studies in comparative international development (scid)*, **2**(11), 165–182.

King, Miriam L, & Magnuson, Diana L. 1995. Perspectives on historical us census undercounts. *Social science history*, 455–466.

Klein, Herbert S. 1983. The integration of italian immigrants into the united states and argentina: A comparative analysis. *The american historical review*, 306–329.

Knights, Peter R. 1991. Potholes in the road of improvement? estimating census underenumeration by longitudinal tracing: U.s. censuses, 1850-1880. *Social science history*, **15**(4), pp. 517–526.

Kosack, Edward, & Ward, Zachary. 2014. Who crossed the border? self-selection of mexican migrants in the early twentieth century. *The journal of economic history*, **74**(04), 1015–1044.

Leeuwen, MHD van, Maas, Ineke, & Miles, Andrew. 2002. *Hisco: Historical international standard classification of occupations*. Leuven: Leuven University Press.

Long, Jason. 2005. Rural-urban migration and socioeconomic mobility in victorian britain. *The journal of economic history*, **65**(01), 1–35.

Long, Jason, & Ferrie, Joseph. 2013. Intergenerational occupational mobility in great britain and the united states since 1850. *The american economic review*, **103**(4), 1109–1137.

Lubotsky, Darren. 2007. Chutes or ladders? a longitudinal analysis of immigrant earnings. *Journal of political economy*, **115**(5), 820–867.

Maddison, Angus. 2007. *The world economy volume 1: A millennial perspective volume 2: Historical statistics*. Academic Foundation.

Manski, Charles F. 1993. Identification of endogenous social effects: The reflection problem. *The review of economic studies*, **60**(3), 531–542.

Míguez, Eduardo José. 1993. La movilidad social de nativos e inmigrantes en la frontera bonaerense en el siglo xix: datos, problemas, perspectivas. *Estudios migratorios latinoamericanos*, **8**(24), 139–170.

Mill, Roy, & Stein, Luke CD. 2012. *Race, skin color, and economic outcomes in early twentieth-century america*. Tech. rept. Working Paper, Stanford University December.

Minns, Chris. 2000. Income, cohort effects, and occupational mobility: a new look at immigration to the united states at the turn of the 20th century. *Explorations in economic history*, **37**(4), 326–350.

Otero, Hernán. 1994. Redes sociales primarias, movilidad espacial e inserción social de los inmigrantes en la argentina. los franceses de tandil, 1850-1914. *Estudios migratorios latinoamericanos*, **9**(28), 521–548.

Panettieri, José. 1965. *Los trabajadores en tiempos de la inmigración masiva en argentina 1870-1910*. Ph.D. thesis, Facultad de Humanidades y Ciencias de la Educación.

Panettieri, José. 1998. El informe buchanan: Primer estudio sobre salarios y precios en la argentina, 1886/1896. *Sociohistórica*, **3**(4).

Parman, John. 2015. Childhood health and sibling outcomes: Nurture reinforcing nature during the 1918 influenza pandemic. *Explorations in economic history*, **58**, 22–43.

Provincia de Buenos Aires. 1883. Censo general de la provincia de buenos aires, 1881. *Buenos aires*, 239.

Ruggles, Steven, Sobek, Matthew, Fitch, Catherine A, Hall, Patricia Kelly, & Ronnander, Chad. 1997. *Integrated public use microdata series: Version 2.0*. Historical Census Projects, Department of History, University of Minnesota.

Salisbury, Laura. 2014. Selective migration, wages, and occupational mobility in nineteenth century america. *Explorations in economic history*, **53**, 40–63.

Sofer, Eugene F. 1982. From pale to pampa: A social history of the jews of buenos aires.

Somoza, Jorge L. 1967. *Muestras de los dos primeros censos nacionales de población, 1869 y 1895*. Instituto Torcuato Di Tella, Centro de Investigaciones Sociales.

Szuchman, Mark. 1981. *Mobility and integration in urban argentina: Córdoba in the liberal era*. University of Texas Press.

Taylor, Alan M. 1994. Mass migration to distant southern shores. *Migration and the international labor market, 1850-1939*, 91–115.

Tolcachier, Fabiana Sabina. 1995. Movilidad socio-ocupacional de los israelitas en el partido de villarino, 1905-1950. *Estudios migratorios latinoamericanos*, **10**(31), 633–672.

Williamson, Jeffrey G. 1995. The evolution of global labor markets since 1830: background evidence and hypotheses. *Explorations in economic history*, **32**(2), 141–196.

Winkler, William E. 1988. Using the em algorithm for weight computation in the fellegi-sunter model of record linkage. *Page 671 of: Proceedings of the section on survey research methods, american statistical association*, vol. 667.

Winkler, William E. 1990. String comparator metrics and enhanced decision rules in the fellegi-sunter model of record linkage.

Tables and figures

Table 1: Frequent occupations, 1895 Census

(a) Natives

Occupation	Frequency	%	Occupational group
Farmer	2943	38.73	Farmer
Laborer	1249	16.44	Unskilled
Breeder	536	7.05	Farmer
Storekeeper	330	4.34	White collar
Independent means	238	3.13	White collar
Carpenter	166	2.18	Skilled/semiskilled
Carter	156	2.05	Skilled/semiskilled
Public employee	115	1.51	White collar
Construction worker	110	1.45	Skilled/semiskilled
Shepherd	97	1.28	Unskilled
Total Top 10	5940	78.17	
Total	7599	100	

(b) Immigrants

Occupation	Frequency	%	Occupational group
Farmer	964	17.42	Farmer
Storekeeper	873	15.77	White collar
Laborer	650	11.74	Unskilled
Independent means	302	5.46	White collar
Carpenter	212	3.83	Skilled/semiskilled
Construction worker	196	3.54	Skilled/semiskilled
Breeder	153	2.76	Farmer
Shoemaker	117	2.11	Skilled/semiskilled
Carter	89	1.61	Skilled/semiskilled
Clerk	72	1.30	White collar
Total Top 10	3628	65.55	
Total	5535	100	

Notes: This table shows the ten most frequent occupations in 1895 among working-age natives and immigrants in the sample linking the 1869 and 1895 national censuses.

Table 2: Occupational mobility of natives and immigrants, 1869 (rows) and 1895 (columns)

(a) Natives

1869	White Collar	Farmer	Skilled/semi-skilled	Unskilled	Total
White Collar	**0.48**	0.35	0.07	0.10	100
	(**406**)	(298)	(55)	(87)	(846)
Farmer	0.10	**0.68**	0.06	0.15	100
	(239)	(**1584**)	(151)	(358)	(2332)
Skilled/semi-skilled	0.13	0.37	**0.35**	0.15	100
	(140)	(397)	(**371**)	(163)	(1071)
Unskilled	0.08	0.46	0.10	**0.36**	100
	(160)	(890)	(203)	(**700**)	(1953)
Total	0.15	0.51	0.13	0.21	100
	(945)	(3169)	(780)	(1308)	(6202)

(b) Immigrants

White Collar	**0.57**	0.18	0.13	0.11	100
	(**613**)	(196)	(138)	(120)	(1067)
Farmer	0.18	**0.47**	0.16	0.19	100
	(78)	(**204**)	(68)	(81)	(431)
Skilled/semi-skilled	0.29	0.15	**0.42**	0.15	100
	(474)	(247)	(**691**)	(245)	(1657)
Unskilled	0.23	0.27	0.23	**0.27**	100
	(338)	(404)	(342)	(**411**)	(1495)
Total	0.32	0.23	0.27	0.18	100
	(1503)	(1051)	(1239)	(857)	(4650)

(c) Summary measures of mobility

	M	d(P,J)	d(Q,J)	d(P,Q)	$d(P,Q)^i$
	(1)	(2)	(3)	(4)	(5)
Natives (P)	0.51	17.23 ***		6.64 ***	2.79 *
Immigrants (Q)	0.59		13.01 ***		

Notes: Panel (a) presents an occupational transition matrix for native-born males. Each cell in this matrix shows the percentage and number (between brackets) of individuals in the linked sample in each occupational category in 1869 (rows) and 1895 (columns). Panel (b) presents the same matrix for foreign-born individuals. Occupations were classified based on the HISCLASS scheme. White-collar (HISCLASS 1-5), farmer (HISCLASS 8), skilled/semi-skilled (HISCLASS 6-7,9) and unskilled (HISCLASS 10-12). Panel (c) reports summary measures of mobility. M (column 1) corresponds to the fraction of individuals off the main diagonal of the matrix. $d(P,J)$ and $d(Q,J)$ (columns 2 and 3) correspond to the distance between matrices P and Q and a matrix representing full independence, respectively. $d(P,Q)$ (column 4) represents the distance between matrices P and Q and $d(P,Q)^i$ (column 5) represents this same distance after excluding the elements of the main diagonal from each matrix. For each of these distances, I performed a test of the hypothesis that $d(i,j)=0$. Significance levels are indicated by $***p<0.01$, $**p<0.05$, $*p<0.1$.

Table 3: Occupational earnings regressions

	Baseline	Geography		Literacy	Both
	(1)	(2)	(3)	(4)	(5)
Immigrant	-0.0320***	-0.103***	-0.0786***	-0.1000***	-0.114***
	(0.00810)	(0.0136)	(0.0135)	(0.00822)	(0.0127)
Immigrant X 1895	0.0585***	0.0691***	0.0586***	0.0548***	0.0471***
	(0.0118)	(0.0199)	(0.0186)	(0.0121)	(0.0179)
Observations	21462	12458	21462	20350	20350

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Robust standard errors in parentheses. All the regressions control for a fourth order polynomial on age and are weighted by the country of birth distribution in 1895. See the online appendix for details on how the income measure was computed. Column 1 reports the baseline specification. In column 2, I restrict the sample to individuals residing in the provinces of Buenos Aires (including the city of Buenos Aires), Entre Ríos and Santa Fe in 1869. In column 3, I include department of residence fixed effects and the interaction between these fixed effects and a 1895 census year indicator. In column 4, I control for literacy and an interaction between literacy and a 1895 census year indicator. In column 5, I control both for department fixed effects and their interaction with the 1895 indicator, and for literacy and its interaction with the 1895 indicator.

Table 4: Occupational earnings regressions, robustness

	Excluding Farmers	Occupational status (property)	Linking			
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant	0.0488***	-0.151***	-0.0378*	-0.00827	0.0187	0.0466***
	(0.0108)	(0.0123)	(0.0217)	(0.00752)	(0.0134)	(0.0161)
Immigrant X 1895	0.102***	0.132***	0.0809***	0.0618***	0.0623***	0.0809***
	(0.0155)	(0.0167)	(0.0233)	(0.0112)	(0.0191)	(0.0231)
Observations	11170	20642	20074	21462	14492	13996

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Robust standard errors in parentheses. All the regressions control for a fourth order polynomial on age and are weighted by the country of birth distribution in 1895 (with exception of columns 3 and 4). In column 1, I exclude individuals employed as farmers in either 1869 or 1895 from the sample. In column 2, I report the results of estimating the baseline equation using an alternative dependent variable based on access to property. In column 3, I reweight the sample to account for differences in observable characteristics between the linked sample and the population. In column 4, I report the results without reweighting the sample by the country of birth distribution in 1895. In column 5, I exclude immigrants with common first names (frequency higher than median within country of birth) from the sample. In column 6, I drop all the immigrants from the sample who are not perfect matches.

Table 5: Occupational mobility of immigrants, declared upon arrival (rows) and in 1895 census (columns)

	Argentina, 1895				
Declared upon arrival	White Collar	Farmer	Skilled/Semi-skilled	Unskilled	Row total
White Collar	**0.59**	0.06	0.24	0.11	100
	(**166**)	(17)	(68)	(31)	(282)
Farmer	0.28	**0.21**	0.28	0.23	100
	(230)	(**172**)	(225)	(185)	(812)
Skilled/semi-skilled	0.23	0.12	**0.52**	0.14	100
	(115)	(60)	(**260**)	(69)	(504)
Unskilled	0.27	0.18	0.33	**0.23**	100
	(175)	(117)	(217)	(**151**)	(660)
Column total	0.30	0.16	0.34	0.19	100
	(686)	(366)	(770)	(436)	(2258)

Notes: Each cell in the table shows the percentage and the number (between brackets) of individuals in the linked sample in each occupational category upon arrival to Argentina (rows) and in 1895 (columns). Occupations were classified based on the HISCLASS scheme. White-collar (HISCLASS 1-5), farmer (HISCLASS 8), skilled/semi-skilled (HISCLASS 6-7,9) and unskilled (HISCLASS 10-12).

Table 6: Intergenerational mobility of children of natives and immigrants, fathers in 1869 (rows) and sons in 1895 (columns)

(a) Sons of natives

Fathers, 1869	White Collar	Farmer	Skilled/semi-skilled	Unskilled	Total
White Collar	**0.44**	0.27	0.12	0.17	100
	(**888**)	(548)	(251)	(337)	(2024)
Farmer	0.12	**0.49**	0.11	0.28	100
	(1159)	(**4709**)	(1019)	(2631)	(9518)
Skilled/semi-skilled	0.16	0.33	**0.26**	0.25	100
	(419)	(879)	(**701**)	(655)	(2654)
Unskilled	0.10	0.31	0.13	**0.45**	100
	(366)	(1116)	(481)	(**1629**)	(3592)
Total	0.16	0.41	0.14	0.30	100
	(2832)	(7252)	(2452)	(5252)	(17788)

(b) Sons of immigrants

White Collar	**0.69**	0.11	0.11	0.09	100
	(**562**)	(92)	(89)	(70)	(813)
Farmer	0.19	**0.51**	0.12	0.18	100
	(87)	(**232**)	(52)	(81)	(452)
Skilled/semi-skilled	0.48	0.10	**0.29**	0.12	100
	(398)	(86)	(**243**)	(98)	(825)
Unskilled	0.36	0.18	0.20	**0.26**	100
	(151)	(77)	(83)	(**109**)	(420)
Total	0.48	0.19	0.19	0.14	100
	(1198)	(487)	(467)	(358)	(2510)

(c) Counterfactual distribution: Sons of natives

White Collar	Farmer	Skilled/semi-skilled	Unskilled	Total
0.33	0.34	0.16	0.18	100

(d) Summary measures of mobility

	M	d(P,J)	d(Q,J)	d(P,Q)	d(P,Q)^{i}
	(1)	(2)	(3)	(4)	(5)
Sons of natives (P)	0.55	12.41 ***		8.21 ***	4.56 *
Sons of immigrants (Q)	0.54		15.07 ***		

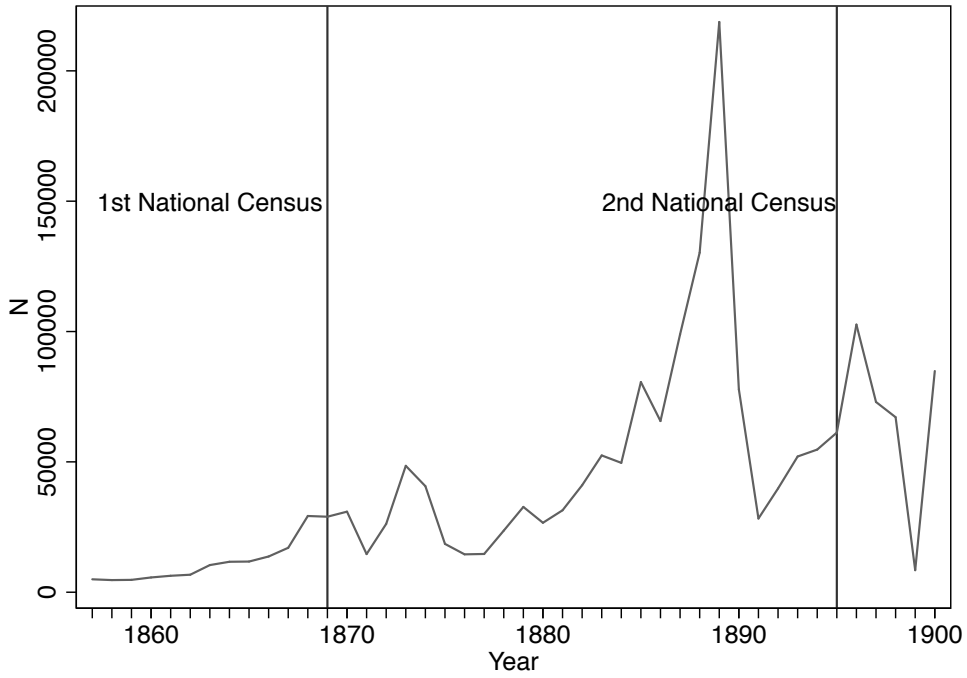
Notes: Panel (a) presents a transition matrix for sons of natives. Each cell in this matrix shows the percentage and the number (between brackets) of fathers in occupation i in 1869 with sons in occupation j in 1895. Panel (b) presents the same matrix for native-born sons of immigrants. Panel (c) presents the counterfactual occupational distribution for sons of natives had they been exposed to the transition matrix of sons of immigrants. Occupations were classified based on the HISCLASS scheme. White-collar (HISCLASS 1-5), farmer (HISCLASS 8), skilled/semi-skilled (HISCLASS 6-7,9) and unskilled (HISCLASS 10-12). Panel (d) reports summary measures of mobility. M (column 1) corresponds to the fraction of individuals off the main diagonal of the matrix. $d(P,J)$ and $d(Q,J)$ (columns 2 and 3) correspond to the distance between matrices P and Q and a matrix representing full independence, respectively. $d(P,Q)$ (column 4) represents the distance between matrices P and Q and $d(P,Q)^i$ (column 5) represents this same distance after excluding the elements of the main diagonal from each matrix. For each of these distances, I performed a test of the hypothesis that $d(i,j)=0$. Significance levels are indicated by $***p<0.01$, $**p<0.05$, $*p<0.1$.

Table 7: Economic outcomes of the second generation

	Literacy			Occupational Earnings			Access to Property		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Second generation	0.363***	0.234***	0.151***	0.201***	0.127***	0.124***	0.00737	-0.0169*	0.0329***
	(0.00641)	(0.00720)	(0.00981)	(0.00745)	(0.00809)	(0.0117)	(0.00817)	(0.00879)	(0.0117)
Household characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Enumeration District FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	23802	23509	23509	22528	22248	22248	24584	24277	24277

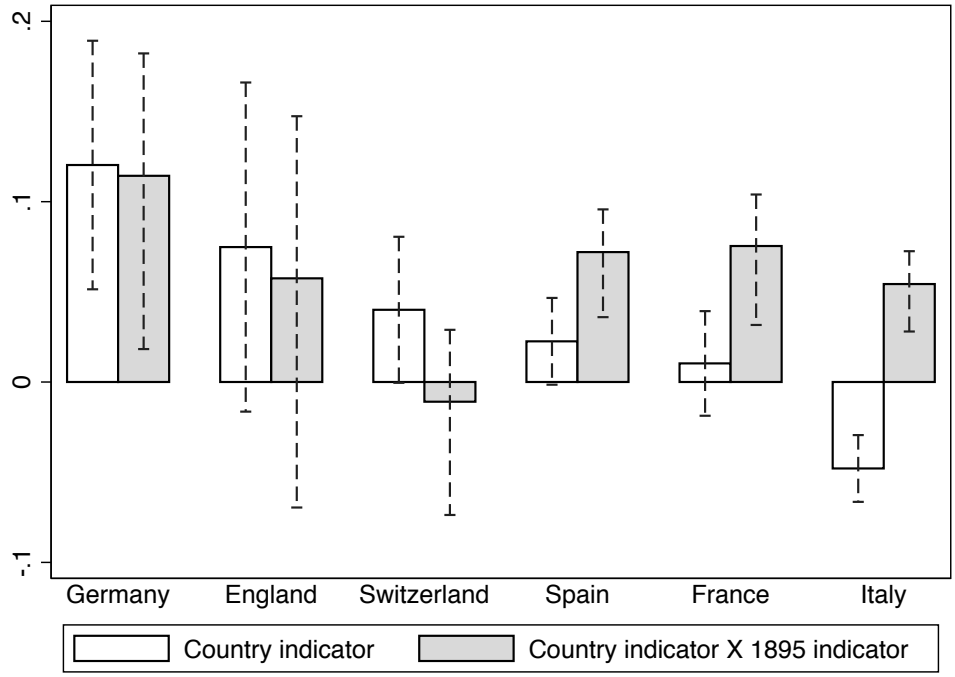
Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Robust standard errors clustered at the level of the father in parentheses. This table shows the estimated coefficients in a regression comparing the children of immigrants and the children of natives in the 1895 census. The dependent variable is an indicator for literacy in columns 1 to 3, log occupational earnings in columns 4 to 6 and an indicator of access to property in columns 7 to 9. Parental place of birth is inferred by linking individuals from the 1895 census to their childhood households in 1869. All regressions control for a quartic in son's and father's age. Household characteristics include the literacy of the father and a full set of dummies corresponding to occupational classes of the father. Enumeration district fixed effects are based on a family place of residence in 1869.

Figure 1: Number of new immigrant arrivals, 1857-1900



Source: Argentina, Anuario de la Dirección General de Estadística (1908)

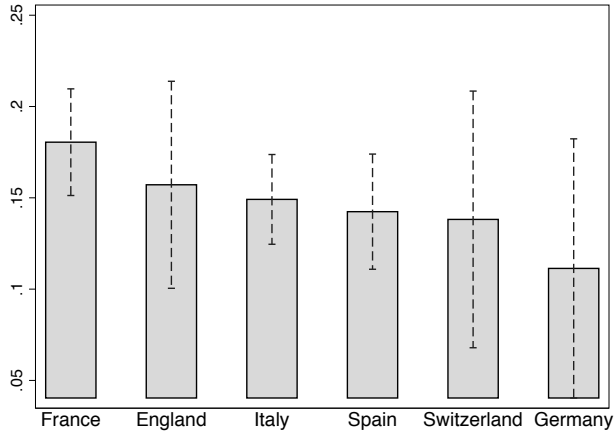
Figure 2: Occupational earnings regression, by country of origin



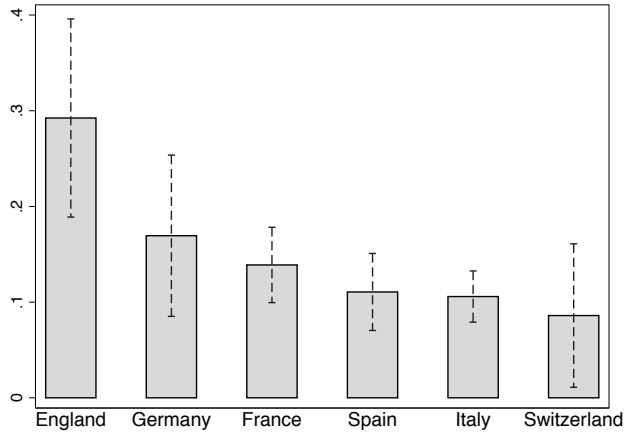
Notes: I estimate a version of equation 1 in which I replace the $Immigrant_i$ indicator with separate indicators for each of the countries included in my sample, as well as an interaction of each of these country indicators with an 1895 census indicator. This figure shows the estimated coefficient around a 95% confidence interval corresponding to each country of origin indicator and the coefficient corresponding to the interaction between the country of origin indicator and the 1895 census indicator. Omitted category are natives.

Figure 3: Economic outcomes of the second generation, by father's country of origin

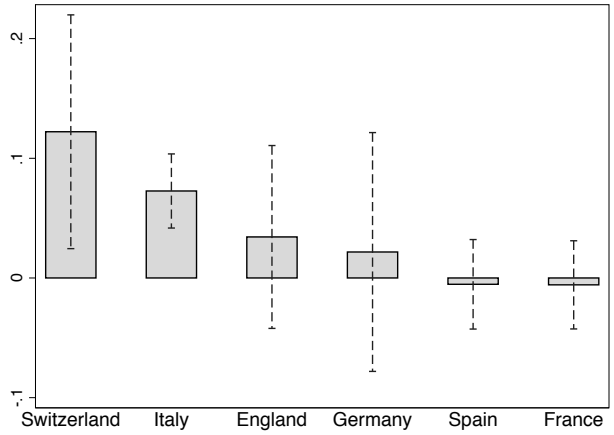
(a) Literacy



(b) (Occupational earnings



(c) Access to property



Notes: I estimate a version of equation 2 in the main text in which I replace the $SecondGeneration_i$ indicator with separate indicator variables for each of the countries included my sample. The figure shows the estimated coefficient around a 95% confidence interval corresponding to each country of origin indicator. The dependent variable in panel (a) is a literacy indicator, in panel (b) is the log of occupational earnings and in panel (c) is an indicator of access to real estate property. Omitted category are sons of natives.

Online Appendix

Linking Algorithm

This section describes the procedure implemented to create the sample linking males across the 1869 and 1895 censuses used throughout the paper. The procedure is similar in spirit to the one employed in Mill & Stein (2012). I started by identifying a set of potential matches for each individual in the 1869 census. To be considered a potential match for an 1869 record, a record in the 1895 census had to satisfy the following criteria:

1. Born in the same province of birth (country of birth in the case of the foreign born)
2. Same first letter of (standardized) first name
3. Same first letter of (standardized) last name
4. Born in the same year $\pm$ 5 years

For those pairs of individuals that met this criteria, I measured their similarity in terms of three identifying variables: first name, last name and year of birth –estimated based on the reported age in the census. To calculate the similarity in first and last names, I used the Jaro-Winkler string distance (Winkler 1990). The Jaro-Winkler distance is a measure of the similarity between two strings. The measure is normalized such that a score of 0 represents two identical strings and a score of 1 represents two strings without any common characters. I measured the similarity between the estimated years of birth by computing the absolute value of their difference.

The next step of the procedure is to summarize the distance in first name, last name and year of birth into a single linking score. There are two main approaches to computing these scores in the literature. The first option is to manually classify a subsample of the potential matches into matches and non-matches and use this subsample to train a classification model. This approach is the one followed by Parman (2015) and Feigenbaum (2016). An alternative approach is to compute the linking scores without a training sample. The core intuition of the approach is to assume an statistical distribution for the vector of distances. Then, I find the parameters of that underlying statistical distribution that maximize the probability of observing the distances in the sample. The procedure is similar to the one used in the US Census (Winkler 1988).

Figures A.1 illustrates the data used to create this sample. Panel (a) of this figure shows an Italian immigrant ("Carlos Bonazola") in the 1869 census. Panel (b) shows the same immigrant in the 1895 census.

The procedure used to match immigrants from the passenger lists to the 1895 census had to be adjusted because a substantial fraction of immigrants adopted a Spanish version of their first names while in the country but yet reported their original names upon arrival. With this in mind, I modified the linking strategy in the following way. In addition to computing the distance in the first names as originally reported, I also computed the distance between a Spanish version of the first name reported upon arrival and the first name in the 1895 census. So, for instance, if an Italian immigrant declared the name "Giuseppe Renzi" upon arrival to Argentina, the Spanish version-based string distance measure would take a value of 0 when compared to an Italian named "José Renzi" in the 1895 census. I hence computed the linking score using as inputs four measures of similarity between the records: distance in reported first name, distance in Spanish version of first name, distance in last name and distance in year of birth.

Accounting for match failure

There is a trade-off in the linking procedure between efficiency –matching a large fraction of the observations– and accuracy –avoiding incorrect matches. In my baseline sample linking working-age natives and immigrants, I find a match –defined as a potential match with a linking score above the $\underline{p}$ threshold– for about 30% of the sample and a unique match –defined as a potential match with a linking score that is both above the $\underline{p}$ threshold and sufficiently better than the second best match ($\frac{p_1}{p_2} > l$)– for about 10%. In the sample linking sons of natives and sons of immigrants, I obtain a match for about 37% of the individuals and a unique match for about 12%. Mill & Stein (2012) use a similar procedure to linking records and report matching rates that are similar to mine.

Table A.2 compares the observed matching rates with the predicted matching rates after subtracting mortality, census underenumeration and return migration. I perform the analysis separately for (1) working-age natives, (2) sons and (3) working-age immigrants. The main reason for match failure is mortality during the intercensal period. Based on the censuses full count, I estimated that about 44% of the natives in my 1869 sample were dead by 1895. Hence, the overall matching rate for natives is capped at 56%. Estimating the mortality rate for immigrants is harder

because there is no information on year of arrival to the country on the census. I hence approximate the mortality rate of immigrants using the mortality rate of natives born in the province of Buenos Aires, which hosted most of the immigrants in my sample. In addition, despite censuses are intended to be a full count of the population, there is non-trivial underenumeration in historical censuses. I am not aware of estimates of underenumeration in 19th-century Argentina censuses, but estimates for a similar time period using US data find underenumeration rates ranging from 7.4% to as large as 23% (King & Magnuson 1995). An additional source of match failure is return migration. Assuming independence among these three sources of underenumeration, the predicted matching rate ranges from 45% to 53% for natives and 23% to 37% for foreigners. The remaining difference between the predicted and the observed matching rate corresponds to errors in the enumeration process that are too severe to be accommodated by my linking procedure. For instance, individuals that misreport their age by more than five years or that have the first letter of their first name misspelled will be missed by my linking procedure.

One challenge in linking the passenger lists to the census is that many immigrants changed their first names while in the country. While my linking procedure takes that explicitly into account, I am only able to accommodate name changes in which the immigrant adopted the Spanish version of their first names. Because of this difficulty, the matching rates are lower in this case than in previous case.

Comparing the linked samples to the population

I start by comparing working-age immigrants and natives in the linked sample to individuals in the 1869 cross section. To perform this comparison, I take advantage of two nationally representative samples of the 1869 and 1895 censuses compiled by Somoza (1967) using the original census manuscripts. Individuals in the linked sample might differ from individuals in the 1869 census cross-section for two conceptually different sets of reasons. First, there is attrition due to mortality and return migration. These sources of attrition are unrelated to the linking procedure, but are also unlikely to be random. Second, there is attrition directly driven by the linking procedure.

Panel (a) of table A.4 shows that there are a number of differences between natives in the 1869 census cross-section and natives in the linked sample. The most salient difference is the higher literacy rates of natives in the linked sample. The distribution across regions of Argentina, as well

as urban/rural status are similar across the two samples. In addition, the fraction of individuals by occupational category is also fairly similar, although the white-collar category is overrepresented in the linked sample. Overall, the evidence suggests some degree of positive selection for natives in the linked sample. Panel (b) of A.4 repeats the analysis for the working-age foreign born. In this case, the differences between individuals in the linked sample and those in the cross section are quantitatively smaller.

I then compare immigrants and natives in the linked sample to individuals in the 1895 census cross-section. This second comparison has the advantage that, in the case of natives, the survivors in the 1895 census cross-section should be similar to the individuals in my linked sample in the absence of biases introduced by the linking procedure.[32] However, the comparison is problematic for immigrants because the stock of immigrants in 1895 includes more recent arrivals, who might be different from the long-term –arrived before 1869– immigrants who are the subject of my study. Unfortunately, it is not possible to restrict the sample to those immigrants who entered Argentina before a given year, as the 1895 census contains no information on an immigrant's year of arrival to the country.[33]

Panel (a) of table A.5 shows that natives in the linked sample look different than natives in the 1895 cross-section along a few dimensions. In particular, they are more likely to be literate and to own property, two characteristics that suggest positive selection of individuals into the linked sample. The distribution of individuals across occupational categories is roughly similar, although, similar to the evidence in table A.4, individuals in the linked sample are underrepresented in the unskilled category. Panel (b) of table A.5 also suggests positive selection of immigrants into the linked sample, although in this case the differences could also stem from differences between long-term immigrants and more recent arrivals.

In table A.6, I compare the fathers –natives and immigrants– of individuals linked to the 1895 census to fathers in the 1869 cross section. In table A.7, I compare all sons –individuals 26 to 44

[32]This statement assumes that natives were a "closed" population. In particular, this statement will not be accurate if there are Argentine born individuals who lived abroad in 1869 but returned to the country in the intercensal period. Quantitatively, this factor is unlikely to matter as the number and fraction of Argentine individuals living abroad was probably very low –below 1% percent according to the 1895 census.

[33]Assuming that immigrants had the same mortality rates than natives and that return migration was 30% in the intercensal period, I estimate that about 70% of the immigrants residing in Argentina in 1895 and in the relevant age cohort arrived after 1869.

years old in 1895– in the linked data to sons in the cross-section.[34] In both cases, the comparison shows a similar pattern to the ones above: individuals in the linked sample are more likely to be literate and more likely to own property. However, the sons of farmers are overrepresented in the linked sample.

In panel (a) of table A.8, I compare immigrants in the passenger lists who were matched to an observation in the 1895 census to immigrants who were not. In panel (b), I compare immigrants in the linked sample in 1895 to immigrants in the 1895 cross section. Immigrants in the passenger lists are older than immigrants in he linked sample but look similar to them in terms of civil status and occupational categories. The main difference between the linked sample and the cross-section is that immigrants from Spain are overrepresented in the linked sample. This overrepresentation likely reflects the fact that immigrants from Spain did not change their name upon arrival to Argentina and it is hence easier to find them in the 1895 census. Compared to immigrants in the census cross-section (panel (b) of table A.8), individuals in the linked sample from passenger lists are younger, more likely to reside in urban locations and underrepresented among farmers.

Identifying fathers and sons in the data

There are two main challenges in identifying fathers and sons in the 1869 census. The first challenge is that not all fathers and sons live in the same household. Sons might have left their childhood household by the time of the census. I minimize this possibility by focusing on children who are young enough (18 years old or younger) in 1869 to presumably still be living with their parents by the time of the census. Second, the father might be either absent or dead at the moment of the census.

A second challenge is that even if fathers co-reside with their sons, the data lack household identifiers. In particular, the 1869 census does not include a question on the relationship of each household member to the head of the household. However, because members of the same household were recorded in the census forms consecutively and father and sons share their last names, it is possible to identify for each individual a set of potential fathers. More precisely, for each male

[34] Because the census cross section does not contain information on parental place of birth, it is not possible to distinguish between sons of natives and sons of immigrants in this comparison. In addition, it is not possible to distinguish in the 1895 census cross-section between those individuals who were residing with their father in 1869 and those who did not.

under the age of 18, I identified the set of potential fathers as anyone who met all of the following the criteria: (1) same last name, (2) recorded consecutively in the census forms (either on the same page or in the one immediately before or in the one immediately after), (3) had an age difference of at least 16 years but no more than 50 years.

The procedure for identifying fathers and sons is similar to the one used by IPUMS to impute relationships among different household members in the 1850, 1860 and 1870 US censuses, where the question on relationship to head of household is also unavailable (Ruggles *et al.* 1997).

Occupations and earnings data

This subsection describes the sources and assumptions used to compute the occupational earnings measure used throughout the paper. This measure varies across but not within occupations and is intended to capture the typical earnings in a given occupation.

Salaried workers

The data on daily wages of blue-collar workers comes from two sources: Buchanan (1898) and the 1881 census of the province of Buenos Aires (Provincia de Buenos Aires 1883). Buchanan was the economic aggregate of the US Embassy in Buenos Aires and systematically collected wage data for workers in this city. His report contains yearly information on the typical wages on 95 occupations from 1886 to 1896. These data have been used extensively in historical research (Dorfman 1942, Panettieri 1965; 1998) and are considered to be accurate (Conde 1979). In particular, for each of the occupations in the Buchanan's data, I take a simple average of the wages in 1894 and 1896. I complement these data with the 1881 census of the Province of Buenos Aires, which contains information from this province on the wages in 65 different occupations.

Because of its greater level of detail and the availability of data closer to 1895, my baseline results use the information from Buchanan (1898) to assign a wage to those occupations in which this information is available from both Buchanan (1898) and the 1881 census. In those cases where the information is only available in the 1881 census, I use wages in the census scaled– to account for changes in the price level that took place over this time period– by the mean wage in Buchanan (1898).

The data on the wages of public employees comes from the 1893 national census of public

employees, which contains the full roster of public employees and their corresponding monthly wages (Argentina, Dirección General de Estadistica 1895). I computed average wages in the public sector by dividing the sum of the wages of male public employees by their total number.

Wages in the above sources were sometimes reported on a monthly basis. In these cases, I converted the data to daily values by dividing the monthly wage by 25 working days.

Business owners

To estimate the average income of storekeepers –"comerciantes"-, I complemented the wage data with information on the size of the capital stock in the commercial sector, obtained from the third volume of the 1895 national census (de la Fuente 1898). Conceptually, the earnings of a storekeeper could be decomposed into the returns to capital and the returns to labor. Based on this insight, average occupational earnings were computed as the sum of the earnings of a store clerk –the returns to labor– and the estimated per capita returns to capital, assuming a 8% net annual return on capital.[35] I computed per capita capital as the ratio between the total capital stock in the commercial sector –as reported in the census– and the total number of individuals who declared working as shopkeepers in the census More precisely, earnings of shopkeepers s were estimated as: $Earnings_s = Labor\ Income_s + 0.08\frac{Capital_s}{\#Storekeepers}$

Average earnings of the owners of industrial firms –"industriales" and "fabricantes" in the census– were similarly estimated by adding the labor income –in this case, the earnings of a foreman– and the returns to capital in the industrial sector. Data on capital in the industrial sector is also from the 1895 national census.

Farmers

Farmers constitute the most challenging category to measure typical earnings. I estimated the income of farmers using the following procedure. I first divided farmers in the agricultural sector into two groups: (1) *hacendados* and *estancieros* and (2) *agricultores*. The *hacendados* category corresponds to those holding the largest plots of land. Overall, less than 10% of all the farmers in my sample are in the *hacendados* category. The second group in the agricultural sector –*agricultores*. encompasses the vast majority of farmers. I then estimated the earnings of farmers using the

[35] In assuming an 8% return to capital, I follow Álvarez & Nicolini (2010).

information provided in the Congressional report of the farming sector prepared by Correa & Lahitte (1898). This report includes information on the typical revenue and expenditure in inputs of farms of different size.

Others

Finally, I assigned the mean earnings within their corresponding HISCLASS to those occupations for which I could not find information in the above sources.

Table A.9 shows the share of observations in the sample linking the 1869 and 1895 censuses that correspond to each of the above data sources in my sample linking working-age natives and immigrants across the 1869 and 1895 censuses.

Table A.1: Matching rate by country of origin, 1869 to 1895 censuses

(a) Working-age individuals

Country of Origin	N	Matched	Uniquely Matched	Matched (%)	Uniquely matched (%)
England	2861	274	159	0.096	0.056
France	11342	2662	975	0.235	0.086
Germany	1815	165	134	0.091	0.074
Italy	25851	10231	2728	0.396	0.106
Spain	15065	5368	1638	0.356	0.109
Switzerland	1821	234	190	0.129	0.104
Total	58755	18934	5824	0.322	0.099
Natives	182982	48593	17352	0.266	0.095

(b) Sons

Country of Origin	N	Matched	Uniquely Matched	Matched (%)	Uniquely matched (%)
England	1189	313	154	0.263	0.130
France	5310	1723	733	0.324	0.138
Germany	740	192	123	0.259	0.166
Italy	9481	3481	1350	0.367	0.142
Spain	5629	2270	663	0.403	0.118
Switzerland	583	139	86	0.238	0.148
Total	22932	8118	3109	0.354	0.136
Natives	183532	71141	21241	0.388	0.116

Notes: This table computes the matching rates in the sample linking the 1869 and 1895 censuses of population. In panel (a), I compute the matching rate among working-age individuals (18 to 35 years old in 1869) and in panel (b) I repeat the same for the sample of sons (aged 17 years old or less in 1869). In both cases, I compute the matching rate separately for each of the countries of origin included in my sample.

Table A.2: Accounting for match failure, 1869 to 1895 censuses

	Natives		Foreigners
	< 18 years old	≥ 18 years old	
	(1)	(2)	(3)
Mortality	0.42	0.42	0.25
Census underenumeration	0.9-0.22	0.9-0.22	0.9-0.22
Return migration	.	.	0.3-0.5
Predicted matching rate	0.45 -0.53	0.45 -0.53	0.29 -0.48
Share matched	0.36	0.27	0.32
Share uniquely matched	0.11	0.09	0.10

Notes: This table reports the reasons for match failure in the sample linking the 1869 and 1895 censuses. The predicted matching rate is computed assuming independence among the factors leading to match failure. Mortality is estimated based on the census data and using only the observations corresponding to natives in the relevant age cohort. Return migration estimates are from Alsina (1898). Estimates of census underenumeration are based on estimates from the US spanning the same time period (Knights 1991), as no references were found for the case of Argentina.

Table A.3: Matching rate by country of origin, passenger lists to 1895 census

Country of Birth	N	Matched	Uniquely Matched	Matched (%)	Uniquely matched (%)
England	2499	90	63	0.036	0.025
France	9857	568	416	0.058	0.042
Germany	2187	95	82	0.043	0.037
Italy	27023	2378	1390	0.088	0.051
Spain	11045	3799	1164	0.344	0.105
Switzerland	1425	53	42	0.037	0.029
Total	54036	6983	3157	0.129	0.058

Notes: The sample includes men at least 18 years old upon arrival and that would be at most 60 by 1895. The number of matched cases refers to men who were matched to an observation in the 1895 census. Men are classified as unique if they are unique in the passenger lists by their combination of name, country of birth and age.

Table A.4: Linked sample and cross-sectional data, 1869 census

(a) Working-age natives

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	25.23	25.29	0.44
Literate	0.26	0.39	0.00
II: Place of residence			
Urban	0.28	0.29	0.12
East	0.46	0.44	0.02
West	0.10	0.14	0.00
North	0.44	0.42	0.02
III: Occupation			
White collar	0.11	0.14	0.00
Farmer	0.38	0.37	0.17
Skilled/semiskilled	0.18	0.17	0.07
Unskilled	0.33	0.32	0.05
Observations	10530	6908	

(b) Working-age immigrants

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	26.67	26.34	0.00
Literate	0.65	0.67	0.00
II: Place of residence			
Urban	0.70	0.70	0.72
East	0.98	0.97	0.07
West	0.00	0.00	0.40
North	0.02	0.03	0.02
III: Occupation			
White collar	0.22	0.23	0.38
Farmer	0.12	0.09	0.00
Skilled/semiskilled	0.38	0.35	0.01
Unskilled	0.28	0.32	0.00
IV: Country of birth			
England	0.05	0.03	0.00
France	0.18	0.17	0.16
Germany	0.03	0.02	0.00
Italy	0.47	0.47	0.65
Spain	0.26	0.28	0.04
Switzerland	0.00	0.03	0.00
Observations	3477	5210	

Notes: Panel (a) compares working-age natives in the sample linking the 1869 and 1895 censuses to working-age natives in the 1869 census cross-section. Panel (b) repeats the same exercise for the working-age foreign born. The census cross-section corresponds to the sample described in Somoza (1967).

Table A.5: Linked sample and cross-sectional data, 1895 census

(a) Working-age natives

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	50.48	51.10	0.00
Literate	0.37	0.46	0.00
II: Place of residence			
Urban	0.23	0.26	0.00
East	0.52	0.44	0.00
West	0.11	0.14	0.00
North	0.37	0.42	0.00
South	0.01	0.00	0.12
III: Property			
Owns property	0.41	0.51	0.00
IV: Occupation			
White collar	0.14	0.16	0.02
Farmer	0.50	0.51	0.71
Skilled/semiskilled	0.11	0.12	0.06
Unskilled	0.25	0.21	0.00
Observations	2904	6854	

(b) Working-age immigrants

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	49.82	51.99	0.00
Literate	0.72	0.76	0.00
II: Place of residence			
Urban	0.61	0.66	0.00
East	0.94	0.95	0.56
West	0.02	0.01	0.08
North	0.03	0.04	0.22
South	0.01	0.00	0.02
III: Property			
Owns property	0.30	0.47	0.00
IV: Occupation			
White collar	0.26	0.33	0.00
Farmer	0.28	0.23	0.00
Skilled/semiskilled	0.27	0.26	0.61
Unskilled	0.19	0.18	0.42
V: Country of birth			
England	0.03	0.03	0.05
France	0.13	0.17	0.00
Germany	0.02	0.02	0.29
Italy	0.58	0.47	0.00
Spain	0.23	0.28	0.00
Switzerland	0.00	0.04	0.00
Observations	2976	4889	

Notes: Panel (a) compares working-age natives in the sample linking the 1869 and 1895 censuses to natives in the 1895 census cross-section. Panel (b) repeats the same exercise for the working-age foreign born. The census cross-section corresponds to the sample described in Somoza (1967).

Table A.6: Linked sample and cross-sectional data, 1869 census, fathers

(a) Native fathers

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	35.34	38.39	0.00
Literate	0.26	0.33	0.00
II: Place of residence			
Urban	0.27	0.26	0.00
East	0.47	0.44	0.00
West	0.11	0.14	0.00
North	0.42	0.42	0.93
III: Occupation			
White collar	0.10	0.12	0.00
Farmer	0.45	0.53	0.00
Skilled/semiskilled	0.17	0.15	0.00
Unskilled	0.28	0.20	0.00
Observations	15577	19366	

(b) Immigrant fathers

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	34.06	39.98	0.00
Literate	0.63	0.67	0.00
II: Place of residence			
Urban	0.70	0.67	0.00
East	0.97	0.95	0.00
West	0.01	0.01	0.59
North	0.02	0.04	0.00
III: Occupation			
White collar	0.23	0.29	0.00
Farmer	0.15	0.21	0.00
Skilled/semiskilled	0.37	0.33	0.00
Unskilled	0.25	0.18	0.00
IV: Country of birth			
England	0.05	0.05	0.05
France	0.20	0.22	0.00
Germany	0.03	0.04	0.22
Italy	0.47	0.46	0.29
Spain	0.25	0.20	0.00
Observations	5244	3627	

Notes: Panel (a) compares the native fathers in the sample linking the 1869 and 1895 censuses to natives in the 1869 census cross-section. Panel (b) repeats the same exercise for the foreign born fathers. The census cross-section corresponds to the sample described in Somoza (1967).

Table A.7: Linked sample and cross-sectional data, 1895 census, sons

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	33.96	33.08	0.00
Literate	0.48	0.57	0.00
II: Place of residence			
Urban	0.28	0.31	0.00
East	0.54	0.53	0.02
West	0.09	0.11	0.00
North	0.36	0.36	0.91
South	0.01	0.00	0.00
III: Property			
Owns property	0.27	0.34	0.00
IV: Occupation			
White collar	0.18	0.21	0.00
Farmer	0.38	0.37	0.34
Skilled/semiskilled	0.12	0.15	0.00
Unskilled	0.32	0.28	0.00
Observations	5966	21646	

Notes: This table compares sons in the linked sample to sons in the 1895 census cross-section. The census cross-section corresponds to the sample described in Somoza (1967). Note that, because the census cross-section does not include information on parental place of birth, it is not possible to perform this analysis separately for sons of natives and sons of immigrants.

Table A.8: Linked sample and cross-sectional data, passenger lists to 1895 census

(a) Passenger lists

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	32.40	31.59	0.00
Married	0.39	0.40	0.26
IV: Occupation			
White collar	0.11	0.12	0.06
Farmer	0.35	0.34	0.39
Skilled/semiskilled	0.20	0.21	0.37
Unskilled	0.28	0.28	0.74
V: Country of birth			
England	0.05	0.02	0.00
France	0.18	0.14	0.00
Germany	0.04	0.03	0.00
Italy	0.50	0.45	0.00
Spain	0.20	0.35	0.00
Observations	52892	2497	

(b) 1895 Census

Variable	Cross Section (1)	Panel (2)	p-val (3)
I: Demographic			
Age	39.82	39.00	0.00
Literate	0.77	0.84	0.00
II: Place of residence			
Urban	0.62	0.72	0.00
East	0.94	0.93	0.01
West	0.02	0.02	0.17
North	0.04	0.04	0.07
South	0.01	0.01	0.36
III: Property			
Owns property	0.21	0.22	0.68
IV: Occupation			
White collar	0.26	0.29	0.00
Farmer	0.23	0.15	0.00
Skilled/semiskilled	0.27	0.32	0.00
Unskilled	0.19	0.18	0.14
V: Country of birth			
England	0.03	0.02	0.00
France	0.11	0.14	0.00
Germany	0.02	0.03	0.05
Italy	0.59	0.45	0.00
Spain	0.24	0.35	0.00
Observations	9005	2529	

Notes: Panel (a) compares immigrants in the linked sample to immigrants in the passenger lists. Panel (b) compares immigrants in the linked sample to immigrants in the 1895 census.

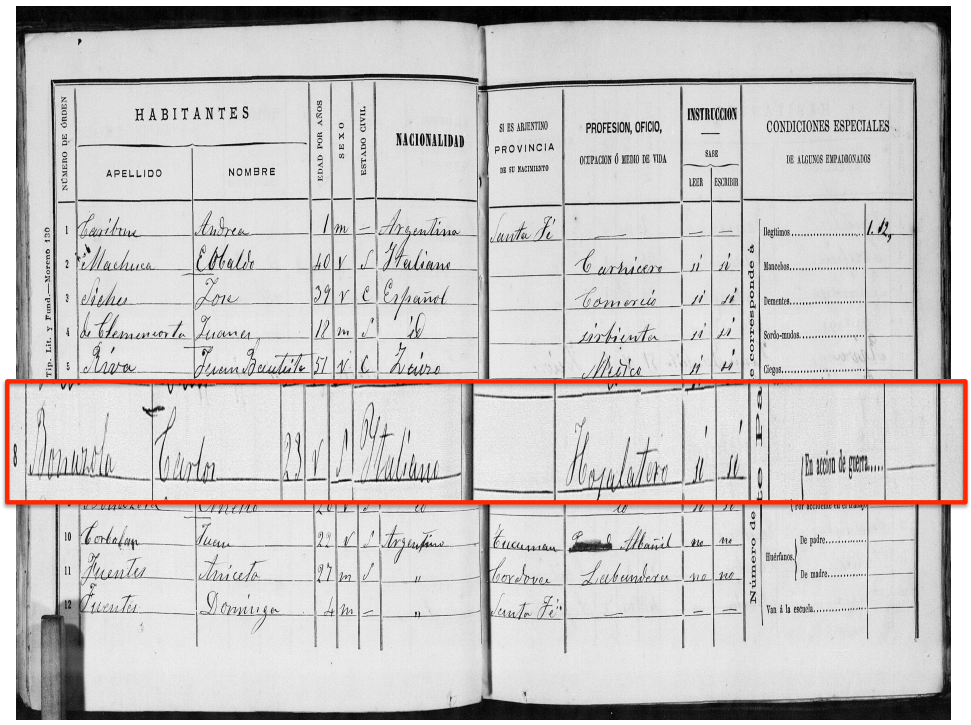
Table A.9: Sources of income data, 1869 to 1895 census sample (working-age)

Source	N	%
Buchanan (1898)	8463	32.12
Pcia. de Buenos Aires census (1881)	1836	6.97
National census (1895)	2998	11.38
Census of public employees (1894)	171	0.65
Correa and Lahitte (1898)	7457	28.30
Imputed based on HISCLASS	5424	20.59
Total Observations	26349	100

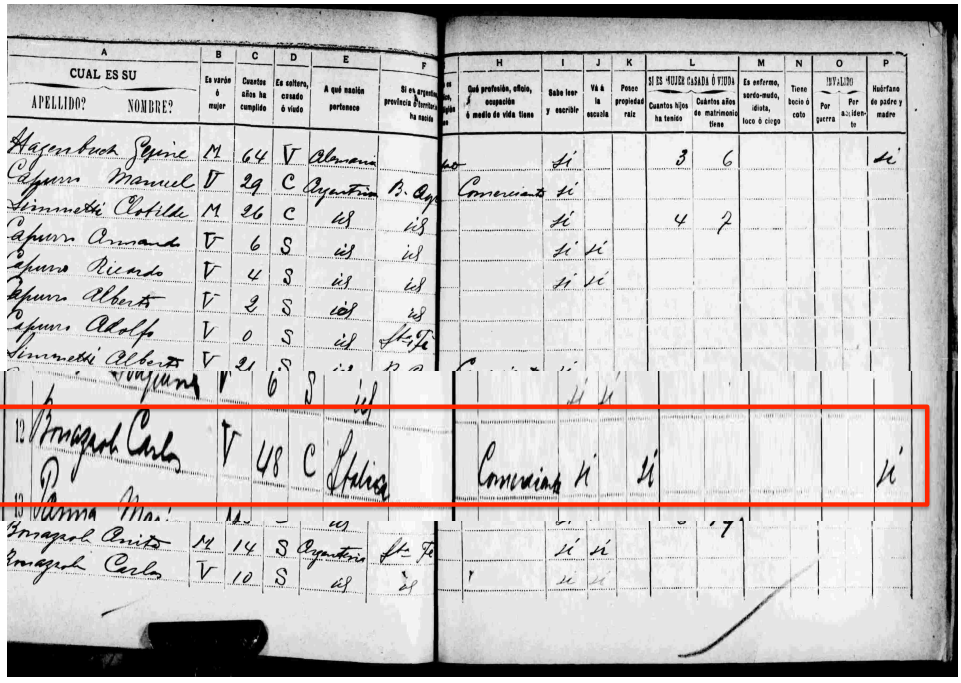
This table describes the sources used to assign an earnings score to each of the occupations in the sample linking working-age males across the 1869 and 1895 censuses.

Figure A.1: Illustration of the linking procedure: an Italian immigrant in 1869 and 1895

(a) 1869 census



(b) 1895 census



This figure shows the same Italian immigrant as in the previous figure but in the 1895 census. He now works as a 'comerciante' (shopkeeper).

Additional tables

Table B.1: Occupational mobility of movers and stayers, 1869 (rows) and 1895 (columns)

(a) Stayers

1869	White Collar	Farmer	Skilled/semi-skilled	Unskilled	Total
White Collar	**0.73**	0.10	0.11	0.06	100
	(**348**)	(49)	(53)	(28)	(478)
Farmer	0.17	**0.62**	0.12	0.09	100
	(24)	(**87**)	(17)	(13)	(141)
Skilled/semi-skilled	0.36	0.03	**0.50**	0.11	100
	(252)	(22)	(**346**)	(77)	(697)
Unskilled	0.31	0.14	0.27	**0.29**	100
	(127)	(56)	(110)	(**119**)	(412)
Total	0.43	0.12	0.30	0.14	100
	(751)	(214)	(526)	(237)	(1728)

(b) Movers

White Collar	**0.45**	0.25	0.14	0.16	100
	(**265**)	(147)	(85)	(92)	(589)
Farmer	0.19	**0.40**	0.18	0.23	100
	(54)	(**117**)	(51)	(68)	(290)
Skilled/semi-skilled	0.23	0.23	**0.36**	0.17	100
	(222)	(225)	(**345**)	(168)	(960)
Unskilled	0.19	0.32	0.21	**0.27**	100
	(211)	(348)	(232)	(**292**)	(1083)
Total	0.26	0.29	0.24	0.21	100
	(752)	(837)	(713)	(620)	(2922)

(c) Summary measures of mobility

	M	d(P,J)	d(Q,J)	d(P,Q)	d(P,Q)^{i}
	(1)	(2)	(3)	(4)	(5)
Stayers (P)	0.48	22.52 ***		16.19 ***	7.41 *
Movers (Q)	0.65		8.97 ***		

Notes: Panel (a) presents an occupational transition matrix for foreign-born individuals who by 1895 resided in the same department of residence as in 1869. Each cell in this matrix shows the percentage and number (between brackets) of individuals in the linked sample in each occupational class in 1869 (rows) and 1895 (columns). Panel (b) presents the same matrix for foreign-born individuals who by 1895 resided in a different department of residence than in 1869. Occupations were classified based on the HISCLASS scheme. White-collar (HISCLASS 1-5), farmer (HISCLASS 8), skilled/semi-skilled (HISCLASS 6-7,9) and unskilled (HISCLASS 10-12). Panel (c) reports summary measures of mobility. M (column 1) corresponds to the fraction of individuals off the main diagonal of the matrix. $d(P,J)$ and $d(Q,J)$ (columns 2 and 3) correspond to the distance between matrices P and Q and a matrix representing full independence, respectively. $d(P,Q)$ (column 4) represents the distance between matrices P and Q and $d(P,Q)^{i}$ (column 5) represents this same distance after excluding the elements of the main diagonal from each matrix. For each of these distances, I performed a test of the hypothesis that $d(i,j) = 0$. Significance levels are indicated by $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table B.2: Occupational mobility of immigrants, declared upon arrival (rows) and in 1895 census (columns). Italy and Spain.

(a) Italians

	Argentina, 1895				
Declared upon arrival	White Collar	Farmer	Skilled/Semi-skilled	Unskilled	Row total
White Collar	**0.49**	0.06	0.34	0.11	100
	(**34**)	(4)	(24)	(8)	(70)
Farmer	0.23	**0.22**	0.32	0.24	100
	(80)	(**76**)	(110)	(82)	(348)
Skilled/semi-skilled	0.19	0.10	**0.60**	0.11	100
	(46)	(24)	(**147**)	(28)	(245)
Unskilled	0.22	0.19	0.34	**0.25**	100
	(76)	(67)	(117)	(**86**)	(346)
Column total	0.23	0.17	0.39	0.20	100
	(236)	(171)	(398)	(204)	(1009)

(b) Spaniards

	Argentina, 1895				
Declared upon arrival	White Collar	Farmer	Skilled/Semi-skilled	Unskilled	Row total
White Collar	**0.65**	0.06	0.15	0.13	100
	(**74**)	(7)	(17)	(15)	(113)
Farmer	0.33	**0.16**	0.25	0.26	100
	(100)	(**49**)	(77)	(81)	(307)
Skilled/semi-skilled	0.30	0.16	**0.32**	0.22	100
	(44)	(24)	(**48**)	(33)	(149)
Unskilled	0.33	0.14	0.32	**0.22**	100
	(78)	(34)	(76)	(**52**)	(240)
Column total	0.37	0.14	0.27	0.22	100
	(296)	(114)	(218)	(181)	(809)

Notes: Each cell in the table shows the percentage and the number (between brackets) of individuals in the linked sample in each occupational class upon arrival to Argentina (rows) and in 1895 (columns). Occupations were classified based on the HISCLASS scheme. White-collar (HISCLASS 1-5), farmer (HISCLASS 8), skilled/semi-skilled (HISCLASS 6-7,9) and unskilled (HISCLASS 10-12).

Table B.3: Intergenerational mobility of natives and immigrants, fathers in 1869 (rows) and sons in 1895 (columns). Buenos Aires, Entre Rios and Santa Fe.

(a) Sons of natives

Fathers, 1869	White Collar	Farmer	Skilled/semi-skilled	Unskilled	Total
White Collar	**0.55**	0.16	0.13	0.15	100
	(**491**)	(145)	(116)	(137)	(889)
Farmer	0.18	**0.30**	0.11	0.41	100
	(394)	(**673**)	(240)	(917)	(2224)
Skilled/semi-skilled	0.32	0.11	**0.28**	0.29	100
	(148)	(50)	(**133**)	(138)	(469)
Unskilled	0.12	0.21	0.13	**0.55**	100
	(157)	(279)	(170)	(**749**)	(1355)
Total	0.24	0.23	0.13	0.39	100
	(1190)	(1147)	(659)	(1941)	(4937)

(b) Sons of immigrants

White Collar	**0.69**	0.10	0.11	0.10	100
	(**494**)	(75)	(78)	(70)	(717)
Farmer	0.19	**0.51**	0.11	0.18	100
	(84)	(**223**)	(50)	(80)	(437)
Skilled/semi-skilled	0.48	0.10	**0.29**	0.12	100
	(376)	(79)	(**228**)	(97)	(780)
Unskilled	0.35	0.18	0.20	**0.26**	100
	(148)	(77)	(83)	(**109**)	(417)
Total	0.47	0.19	0.19	0.15	100
	(1102)	(454)	(439)	(356)	(2351)

(c) Counterfactual distribution: Sons of natives

	White Collar	Farmer	Skilled/semi-skilled	Unskilled	Total
	0.35	0.31	0.15	0.18	100

(d) Summary measures of mobility

	M	d(P,J)	d(Q,J)	d(P,Q)	d(P,Q)^{i}
	(1)	(2)	(3)	(4)	(5)
Natives (P)	0.59	14.72 ***		8.37 ***	5.00 *
Immigrants (Q)	0.55		15.04 ***		

Notes: Panel (a) presents an occupational transition matrix for sons of natives. Each cell in this matrix shows the percentage and the number (between brackets) of fathers in occupation i in 1869 with sons in occupation j in 1895. Panel (b) presents the same matrix for sons of immigrants. Panel (c) presents the counterfactual occupational distribution for sons of natives had they been exposed to the transition matrix of sons of immigrants. Occupations were classified based on the HISCLASS scheme. White-collar (HISCLASS 1-5), farmer (HISCLASS 8), skilled/semi-skilled (HISCLASS 6-7,9) and unskilled (HISCLASS 10-12). I restrict the sample to father-son pairs in which the father resided in the provinces of Buenos Aires –including the city of Buenos Aires-, Entre Rios or Santa Fe in 1869. Panel (d) reports summary measures of mobility based on the occupational transition matrices. M corresponds to the fraction of individuals off the main diagonal of the matrix. $d(P,J)$ and $d(Q,J)$ correspond to the distance between matrices P and Q and a matrix representing full independence, respectively. $d(P,Q)$ represents the distance between matrices P and Q and $d(P,Q)^i$ represents this same distance after excluding the elements of the main diagonal from each matrix. For each of these distances, I performed a test of the hypothesis that $d(i,j)=0$. Significance levels are indicated by $***p<0.01$, $**p<0.05$, $*p<0.1$.

Table B.4: Ethnic capital and the economic outcomes of the second-generation

	Literacy	Earnings	Access to property
	(1)	(2)	(3)
Ethnic capital	0.325**	0.673***	-0.300
	(0.144)	(0.213)	(0.198)
Observations	23509	22248	24277

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Robust standard errors clustered at the level of the father in parentheses. "Ethnic capital" is the average literacy rate –as measured in the 1869 census– among individuals from the same country of birth of the father. Each regression controls for a quartic in son's and father's age, own father's literacy and enumeration district fixed effects, as determined by the place of residence of the family in 1869.