Migration into the Pampas and Economic Development*

Federico Droller†

Brown University

April 2012

Abstract

I study the long-run impact of European migration in the late nineteenth century to Argentina, and shows that the historical population composition generated differences in economic outcomes. Areas with a historically higher shares of European population have significantly higher per-capita GDP, higher education rates and skilled workers, eighty years later, than areas with lower shares of European population. I use an instrumental variables approach to assess the causal relation between the share of European population on economic development. I show that industrialization and literacy rates were channels through which the effect persisted over the long-run. Counties were Europeans accounted for a higher share of the population had greater industrialization and higher literacy rates in the past.

1 Introduction

The importance of historical events on economic outcomes in the long-run has been shown by a large body of literature (Engerman & Sokoloff (1997, 2002), Acemoglu et al. (2001), Glaeser et al. (2004) and Banerjee & Iyer (2005)). But the reasons or the channels through which history matters continues to be an open debate. The discussion centered around

*I am grateful to Pedro Dal Bo, Ross Levine and David Weil for their guidance and advice. Vernon Henderson, Oded Galor, Blaise Melly and Florencia Borrescio Higa provided valuable comments and suggestions.

†Department of Economics, Brown University, Box B, 64 Waterman St, Providence, RI 02912. federico_droller@brown.edu

three main determinants of economic development: factor endowments, past political institutions and human capital. The argument follows that the appropriate natural resources, the *right* political institutions or a population with higher human capital, was decisive for economic growth and development, and long-term differences in economic development could be attributed to these initial differences.

Vast empirical evidence was shown for past factor endowments and political institutions as important determinants of current differences in economic outcomes (Acemoglu et al. (2002), Dell (2010) and Galor, Moav & Vollrath (2009)),however there is less consensus pointing to the importance of historical differences in human capital. Easterly & Levine (2009) find some support for human capital as an important determinant of economic growth in the long-run, while Gennaioli, La Porta, Lopez-de-Silanes & Shleifer (2011) find evidence that current human capital differences account for regional differences in development. Confounding effects from factor endowments, political institutions and human capital made difficult in previous studies to disentangle the effect attributable to each determinant. This paper uses variation in the assignment of European immigrants in Argentina to identify how immigrants affected economic development and through which channels the effect persisted over time.

Particularly, I study European immigration to the fertile plains of Argentina, an undeveloped area *opened* for settlement, with uniform geographical characteristics and fixed political institutions. Between 1857 and 1914[1] close to 5.5 million Europeans migrated to Argentina, arriving at the port of Buenos Aires and afterwards moving to the other cities and countryside. Most of them settled down in the Pampas, the fertile plains, practicing agriculture or living in small urban areas where economic life centered around farming.

The Argentinean government followed a policy of expanding the sovereign territory over areas occupied by native indigenous tribes, in response to the threat that these tribes casted on settlements. Military campaigns slowly gained power over the regions allowing the government to offer the area for settlement and as a factor of production. The timing of the military campaigns and the arrival of the European immigrants resulted in two

[1] 1857 is the first year the Argentine government started recording statistics for immigration and in 1914 the government conducted a census.

sources of variation that I will use to construct an exogenous measure of the share of immigration in a given region.

Using this constructed share of European population as an instrumental variable for the actual share of European population I compare counties in the fertile plains and estimate that increasing the share of European population by one standard deviation has a long-run effect of an increase in log per-capita GDP by one standard deviations, an increase approximately equal to move from the 25 percentile in the income distribution to the 75 percentile. Moreover, areas with higher share of European immigration in 1914 have higher share of population with higher education in 2001. These results provide evidence for the effect of initial differences in human capital on economic development over the long-run.

The process by which the fertile plains where settled close to one hundred years ago provides an opportunity to analyze the effect of European immigration and European's human capital on economic development, by focusing at a single country with equal political institutions across regions, and in a cross section of counties with uniform geographic characteristics.

This paper is organized as follows, Section II reviews historical accounts of the conquest of the plains and the European immigration. Section III describes the data and presents summary statistics. Section IV develops the empirical strategy and the results. Finally section V concludes.

2 The History of the Fertile Plains

2.1 The Conquest of the Plains: the Desert

It was not until end of the nineteen century that the Argentinean government gained political power over the whole territory that nowadays is Argentina. During colonial times and after independence from the Spanish Empire in 1816 most of the fertile plains where settled by several indigenous tribes that did not recognize the Argentinean government. Relationships between Argentineans and indigenous tribes were characterize by mistrust and violence. By the time of independence the situation was such that Argentineans used

to dispute land and wild livestock to the indigenous tribes, while indigenous people organized assaults into settlements and cities, stealing livestock, goods and kidnapping people. Indigenous raids attacking cities and military excursions into indigenous settlements, both ending in destruction and deaths, where common. The Argentinean government and main tribes often agreed on peace treaties, but the Argentinean government never recognized that area as an independent state, nor did it recognize indigenous people as legal owners of the land. For the Argentinean government it was a problem of civilization (Argentina) versus barbarism (the indigenous) and the solution was to conquer and settle the fertile plains (and all areas with indigenous tribes), letting the indigenous live within the rules of the government.

As described in the late nineteen century by former president Domingo Sarmiento in the Facudo:

> To the south and the north, savages lurk, waiting for moonlit nights to descend, like a pack of hyenas, on the herds that graze the countryside, and on defenseless settlements. In the solitary caravan of wagons slowly traversing the Pampas that stops to rest for a few moments, the crew, gathered around a poor fire, mechanically turn their eyes toward the south at the least murmur of wind blowing the dry grass, to bore their gaze into the profound darkness of the night, searching out the sinister bulks of savage hordes that from one moment to the next can surprise them unprepared. (Sarmiento 2004, p.46)

The threat of indigenous tribes over Argentinean settlements was not the only concern of the government regarding the national territory. For Argentina to consolidate as a nation it was necessary to delimit its frontiers, which turned necessary to occupy Patagonia, an area also claimed by neighboring country Chile (Lacoste (2002)). But it was not until the end of the civil war in 1862 that a unified national government developed systematic plans to conquer the rest of the territory, starting in 1870 until 1885.

Previous to 1870, military campaigns developed with many years of interruption and loss of domain, in particular during episodes of civil war and the war against Paraguay. Detailed information on the military campaigns and its effect on how the *frontier* be-

tween Argentineans and the indigenous tribes changed over time has been documented by Walther (1964). Figures 1-2 depict maps showing the frontier between Argentina and the indigenous tribes in 1779, 1823, 1826, 1828, 1852, 1860, 1864 and 1876. Gains of territory by the Argentinean army and looses of domain over these years were a consequence of the limited resources the government had for the multiple military conflicts they faced (Luna(1993)).

2.2 Settlement of the Fertile Plains

The end of the civil war, the re-unification and pacification of the country started a period of European migration to Argentina. Immigrants were granted the same legal rights as Argentineans, without need to naturalize or acquire citizenship.The flow of immigrants to Argentina resembles the same flow of immigrants to the USA, Canada and Australia. Europeans were escaping war, famine and persecution in Europe and migrated to countries that offered them the possibility to settle.

Figure 3 shows the time series of immigration and net immigration of Europeans to Argentina. The series starts in 1857 when the national government started recording statistics on the arrival of immigrants to its ports. The flow of migration is far from constant, nor it is a monotonic function of time. The picks and valleys in the time series of immigration can be mostly explained by the conditions in Europe, and less by the conditions in Argentina.

Immigrants settled in cities, urban areas and in the countryside, and were occupied as unskilled labor or skilled labor. Activities were diverse, ranging from farmers to construction workers, merchants and craftsmanships. As of 1895, 41 percent of the European immigrants (males, aged 15 or above) were living in urban areas, while 32 percent devoted their time to farming and 28 percent to non-farm skilled labor.

Although the decision to conquer the plains was unrelated to the immigration patterns, the timing of the expansion of the frontier over the plains overlaps with the arrival of the first European immigrants to the country, as shown in Figure 4. The correlation between the time series of immigration and the amount of land in the fertile plains under the

political power of the government is close to 0.5, the regression of immigration on the amount of land yields and R2 of 20%.

3 Data and Summary Statistics

This study combines current data on economic development (per-capita GDP, higher education rate and share of skilled workers) with past data on economic and social conditions (population density, productive uses of land, etc.). The unit of observation is at the county level, excluding capital cities and the urbanized greater Buenos Aires. The sample covers the four states that hold the fertile plains: Buenos Aires, Santa Fe, Córdoba and Entre Ríos. The southwest section of the fertile plains lays in the state of La Pampa, this state is not included in the sample since by the time of the European immigration it was not a state, but part of the national territory. Statistical information is not as exhaustive for territories as it is for states. Moreover, the state of La Pampa changed all the county boundaries over time. Working with four states allows me to control for unobservable fixed variables at the state level. Though county boundaries have slightly changed over time, it is still possible to match older counties to new counties. New counties were mostly founded on previously unoccupied land, but there were cases where old counties split into two or more counties. When a new county can not be linked to an old county, the observation is drop from the sample. There are 197 counties in the sample, where 31 are new counties not linked to an old county. From the remaining 166 counties, 25 are capital cities or large urban areas and 5 are counties without current information on economic outcomes, leaving a sample of 136 counties in four states.

Historical information comes from four sources: the 1895 and 1914 Argentinean censuses, the Migration Office from Argentina and Walther (1964). Both censuses contain detailed information at the county level on population characteristics and economic activities. I digitalized data on all variables used from the censuses: total population, foreign born population and population living in urban areas. Moreover, the 1914 census includes an agricultural and livestock census, which was used to construct a variable on the economic activities performed at the county level. Somoza and Lattes (1967) computerized

samples of historical 1895 census microdata, from which individual level data on nationality, age, sex and occupation can be obtained. The 1895 individual level data is used for the construction of the occupation and literacy variables.

The Argentine Office of Migration records since 1857 all non-Argentine incoming and outgoing population. Detailed data on the number of migrants and country of origin since 1857 until 1914 was digitalized for this study.

Data on the territory under the political power of the Argentine government comes from Walther (1964). Walther detailed description of the military campaigns are summarize with a series of maps that show for different years the actual *frontier* between the territory under the Argentinean government and the native tribes' territory. Walther's work is based on military and historical documents. I complement these maps with Gallo (1983) and Tell (2008) who provide more detailed information for the states of Córdoba and Santa Fe.

The Argentinean Statistical Office (INDEC) computes GDP at the national and state level, but not at the county level. In 1994 INDEC conducted the National Economic Census (CNE) censing all business at the county level, except for the agricultural sector, recording the value of production, costs, investment, etc. Per-capita GDP is constructed by combining CNE's gross product data with yearly agricultural output estimates from the Ministry of Agriculture (see Appendix). For those state statistical offices that compute GDP at the county level, Buenos Aires and Santa Fe, the correlation between CNE's gross product with state's GDP at the county level is 95%, the correlation between CNE's gross product augmented by the agricultural output estimates and state's GDP is also 95%. The regression of state's GDP on the CNE's gross product augmented by agricultural output has an R2 of 90.34. I will use CNE's gross product augmented by agricultural output as a proxy for GDP at the county level.

Further, I will use data from the 1935 Industrial Census, which documents the number of industrial establishments, the value of the production, the number of workers and the usage of energy at the county level.

Data on higher education rates and share of skilled workers is from the 2001 Pop-

ulation Census and is publicly available from the Argentine Statistical Office. Finally, geo-referenced data on the quality of the soil comes from the National Institute for Agriculture and Livestock Technology (INTA). INTA provides geo-referenced detailed data on the quality of the soil and elaborates an index that assigns a greater value to better soils. This index of land quality refers the geographical conditions of the soil (like ground composition and rain) and not to the technologies used for cultivation. I combine the geo-referenced data provided by INTA with the county boundaries and compute an area weighted average of the land-quality index.

Table 1 shows the summary statistics for the variables used in this study. The average (and median) share of European population is 29% and a standard deviation of 14%, with counties with less than 1% and up to 56% of its population of European origin. GDP per capita averages slightly above 6.700 dollars, where the bottom 25% of the counties have less than 3.560 dollars and top 25% of the counties have a per-capita GDP above 9.000 dollars. On average 10.4% of the population 25 years of age and older have completed more than 12 years of education (completed secondary school and started or finished tertiary or university degrees). Of those individuals reporting an occupation in 2001, on average 18% work in high skilled jobs.

4 Estimation Strategy and Results

I will compare log per-capita GDP, higher education rates and the share of skilled workers between counties with different population composition. I start by running a regression of the dependent variable on the share of European population:

$$y_i = \alpha + \beta SE_i + X_i\gamma + \eta_s + \epsilon_i \tag{1}$$

Where y_i is the dependent variable in county i, SE_i is the share of European population in the total population in county i, X_i are controls for county i characteristics in 1914, and η_s are state fixed effects. County characteristics include population density, share of the population living in urban areas (2000 or more inhabitants), share of productive land

used for agriculture, land-quality and (log) distance to the city of Buenos Aires. Since the city of Buenos Aires is the capital city of the country, the main port of entry (for traded goods and immigrants) and the most densely populated city, proximity to this political and economic relevant city may have independent effects on development.

Table 2 documents OLS results of regressing log per-capita GDP in 1994 on the share of European population in 1914, equation (1). Column 1 only controls for state fixed effects, columns 2-5 add controls for the share of productive land used for agriculture, population density, urbanization rate, distance to the city of Buenos Aires and land quality. Column 7 controls for all variable at the same time, while in column 8 fixed effects are not included. The basic OLS regression shows that the share of Europeans in 1914 has a positive and significant coefficient. In column 7 distance to the city of Buenos Aires has (as expected) a negative coefficient (though not different from zero) reflecting progress coming from proximity to the most important political and economic city. Land quality has a positive effect on development, as does the share of productive land used for agriculture. Population density enters negatively, while urbanization rate positively. These results show the importance of cities for development, but that densely populated areas hinder economic growth.

Following column 7, the preferred specification, a one standard deviation in the share of Europeans increases per-capita GDP by 0.67 standard deviation. As this result shows, European immigration positively affected economic development in the long-run, since close to eighty years after the arrival of European immigrants differences in economic performances can be found across counties depending on the pattern of settlement. But, as discussed above, my results might be biased if the European immigrants selected themselves into the counties depending on an unobservable or omitted variable. To deal with this potential problem I will use instrumental variables to account for the possible endogeneity of the share of European population.

4.1 Instrumental Variable Approach

European migration to the different counties in the fertile plains may not have been random. Immigrants may had information in hand to choose one destination in favor of another, for example, previously settled immigrants may have sent letters or went back to the home country to attract the rest of the family to the newly settled area across the ocean. Even differences in infrastructure, access to railroad or size of the cities in the plains may have played a roll for immigrants when deciding where to settle. To account for the possible endogeneity in where European immigrants settled once they arrived to Argentina, I will construct an exogenous measure of the share of immigrants in each county and use it as an instrumental variable for the actual share of immigrants in a given county.

In order to construct an exogenous measure of the share of immigrants in a give county I will exploit two sources of variation: a) the availability of previously unoccupied land to anyone willing to settled in. And b) the time series of immigration to Argentina between 1857 and 1914. As will be discussed below the variation in both, available land for settlement and arrival of immigrants, can be used to construct a measure of the share of immigrant in a given county, allocating immigrants and Argentineans to counties depending on the year of arrival.

The History of the Instrument

Using historical information on the military campaigns followed by the Argentine government, I am able to assign to each county a year in which (at least half of) the land was available to settlers.

From historical records (Walther (1964)) I am able to trace the area under the political power of the Argentine government for this period. Walther (1964) documents for the years 1779, 1823, 1826, 1860, 1864, 1869 and 18764 the end result of military excursions and the boundary that resulted of these expeditions between the Argentine government and the indigenous tribes, by 1884 the Argentine government controlled the rest of the fertile plains. I assume that no land is conquered or lost until the next military campaign, an assumption very close to the actual events.

For the availability of land for settlement to be exogenous in my analysis, the underlying assumption implies that the governments decision to conquer these vast tracks of land was independent of the arrival of European immigrants to the country. It follows from the above historical description that this was indeed the case. Military and safety issues prompted the government to take power of this region, starting years before the first wave of European immigrants arrived.

For the time series of immigration to be exogenous in my analysis it must be the case that Europeans decided to migrate to Argentina for reasons unrelated to the success or failure of the military campaigns in conquering new land. Europeans decided to migrate because of the condition in their home country, moving mostly to the Americas and Australia. The migration pattern to Argentina resembles that of the USA, the correlation in the time series of migration to Argentina and the USA is 0.795[2]. The decision to migrate to Argentina did not depend on labor and settling possibilities in a given county.

Moreover, the military campaigns ended in 1884 in the fertile plains, when slightly less than 850,000 European immigrants arrived between 1857 and 1880, in comparison to 5.5 million immigrant arrived by 1914.

For the identifying assumption to be correct constructed immigration affects the dependent variable (per capita GDP, higher education, etc.) only through the actual share of European immigration, and has no effect over other channels.

The Construction of the Instrument

The instrument is constructed by assigning Argentinean and European population to each county and simulating the process of population growth, given the fertility and mortality rates, over the years 1857 to 1914. The assignment of population to counties is only possible when the county is under the political power of the Argentine government. Historical records on the military campaigns and its results provide me with information on whether a given county was or was not under the political power of the Argentine

[2]Data on USA migration from Historical Statistics of the United States, Millennial Edition On Line, edited by Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, Cambridge University Press 2006. http://hsus.cambridge.org/HSUSWeb/toc/tableToc.do?id=Ad1-2

government.

Starting in 1857 Europeans will be distributed uniformly across counties, varying the quantity of immigrants each county is assigned by the year of arrival. Argentineans, on the other hand, are initially present in counties under the political power of the Argentine government by 1857, but not in counties conquered after 1857. The population growth of Argentineans and Europeans is given by the fertility rate and the mortality rate.

Europeans arrive every year[3] and move uniformly to any county that is under the political power of Argentina, and once they settled they never move again. Europeans die at rate δ and reproduce at rate ρ, although children born to Europeans in Argentina are considered as Argentineans.

The initial Argentinean population in 1857 comes from the 1869 census, adjusted by the population growth rate to the year 1857. Argentineans die at rate δ and reproduce at rate ρ. There is a fraction ϕ of Argentineans that each year decides to move to a new county. I assume they move equally to all the counties that belong to Argentina.

The mortality rate, the fertility rate and the fraction of Argentineans that move each year are computed from the 1869, 1895 and 1914 censuses. The mortality rate is computed to be equal to 2.2%[4]. The fertility rate is computed to be equal to 5.3%[5]. The moving rate for Argentineans, ϕ, is computed to be equal to 1.95%[6].

[3] From 1857 until 1914.

[4] I compare the stock of Europeans in 1914 with the flow of Europeans from 1857 to 1914 and assuming that Europeans die at a constant rate δ I solve for δ such that $\sum_{t=1857}^{1914}(1-\delta)^{1914-t}\cdot x_t = X_{1914}$, where x_t is the number of Europeans that arrived at time t, and X_{1914} is the stock of Europeans in 1914.

[5] Given the Argentinean population from 1869 and 1914 censuses, and given that children of Europeans are considered Argentineans, I solve for ρ such that:
$w_{1870} = (1-\delta+\rho)\cdot w_{1869} + \rho x_{1869}$,
$w_{1871} = (1-\delta+\rho)\cdot w_{1870} + \rho x_{1870} = (1-\delta+\rho)^2\cdot w_{1869} + (1-\delta+\rho)\cdot\rho x_{1870} + \rho x_{1869}$,
$\vdots$
$w_{1914} = (1-\delta+\rho)^{1914-1869}\cdot w_{1869} + \sum_{t=1869}^{1914-1}(1-\delta+\rho)^{1914-1-t}\cdot\rho x_t$,
where w_t is the number of Argentineans at time t.

[6] Using individual-level data from 1895 census I estimate the fraction of Argentineans living in a different province than the one in which they were born. Define $\pi_{i,a}$ as the fraction of people aged a born in county i, who still live in county i.

$$\pi_{i,a} = \frac{p^i_{i,a}}{\sum_j p^i_{j,a}},$$

where $p^i_{i,a}$ is the number of people born in county i who live in county i, and $p^i_{j,a}$ is the number of people born in county i who live in county j.
Then,

$$\pi^i_{i,a} = (1-\phi_a)^a.$$

I will compute ϕ_a for all ages and then compute the average ϕ weighting by the fraction of people in each

The number of Europeans in each county in 1914 is defined as:

$$X_i = \sum_{t=1857}^{1914} \frac{1}{N_t}(1-\delta)^{1914-t} x_t \cdot \mathbb{1}_i\{t \geq D\}. \tag{2}$$

The number of Argentineans in each county in 1914 is defines as:

$$W_i = W_{i1857}(1-\delta+\rho-\phi)^{57} + \sum_{t=1857}^{1914} \frac{1}{N_t}(1-\delta+\rho-\phi)^{1914-t}(\phi w_t + \rho x_t) \cdot \mathbb{1}_i\{t \geq D\}, \tag{3}$$

where X_i and W_i are the constructed number of Europeans and Argentineans in county i in 1914, respectively. W_{i1857} is the initial number of Argentineans in a given county. $\mathbb{1}_i\{\cdot\}$ is an indicator whether county i belongs to Argentina, and D is the year in which county i started to be under the political power of the Argentine government. $N_t = \sum_i n_{it}$ is number of counties under the Argentinean political power at time t and n_{it} equals 1 if county i belongs to Argentina at time t, 0 otherwise.

The constructed share of Europeans population is defined as $X_i/(X_i+W_i)$, and is used as IV for the actual share of European population.
I run the following specification for the first stage:

$$SE_i = \alpha + \psi CSE_i + X_i\gamma + \eta_p + \epsilon_i \tag{4}$$

Where CSE_i is the constructed share of European immigration.

Figure 5 shows the first-stage correlation between the share of European population and the constructed share of European immigration. Figure 6 shows the first-stage correlation when control variables and fixed effects are included. Both figures show a strong positive correlation between the actual share and the constructed share of European immigration.

cohort.

$$\phi = \sum_{i=1}^{I}\sum_{a=1}^{99} \frac{p_{i,a}}{\sum_i \sum_a p_{i,a}} \cdot (1-\pi_{i,a}^{1/a}),$$

where $p_{i,a}/\sum_i \sum_a p_{i,a}$ is the fraction of a years old in the population.

4.2 The long-run effect of European immigration

Table 3 shows the first-stage regression, equation (4). In column 1 no controls nor fixed effects are included, in column 2 state fixed effects are included, while column 3 also adds the control variables, X_i. The coefficient on the constructed share of immigration remains positive and significant across specifications,confirming the result presented in figures 5 and 6. An F-test of the coefficient ψ shows a strong first-stage with a statistic greater than 30 for the full specification in column 3, and weak identification is ruled out by the Kleibergen-Paap test.

Table 4 shows in columns 1-3 panel A. IV results and in panel B. OLS results for different dependent variables, where the constructed share of immigration is used as instrumental variable for the actual share of European population. In column 1 the dependent variable is per-capita GDP in 1994, panel B. repeats the result from table 2 column 7, panel A. shows the IV regression. The coefficient on the share of Europeans in column 1 shows a long-run effect of the share of European population on per-capita GDP, a one standard deviation in the share of European population increases per-capita GDP by one standard deviations. The point estimate of 5.65 is slightly higher than the OLS estimate of 3.72, suggesting a negative bias in the selection of Europeans to counties.

Column 2 and 3 of Table 4 examine census data on higher education and high skilled occupations in 2001, respectively. Results also show a positive an significant effect of European immigration on these variable. One standard deviation in the share of European immigration raise the share of population with higher education by 0.52 and the share of workers in high skilled occupations by 0.64 standard deviation, both effects are significant at the 10% and 1% level, respectively. Panel B. columns 2 and 3 show OLS results for these last two dependent variables.

The results in columns 1-3 panel A. show a causal effect of European immigration over the long-run, Europeans affected the degree of economic development as measured by GDP, higher education and skilled workers. Next I will examine the channels through which these differences arose and persisted over time.

4.3 The effect of European immigration: the channels of persistence

Why did Europeans affect economic outcomes close to a century after their arrival? How did their initial effect on the economy propagate and persist over time? To answer these questions I will next investigate the channels through which the effect of European immigration persisted over time.

I will focus on two main channels of persistence: Industrialization and Human Capital. Industrialization has been widely understood as an important factor in a country's development, were those countries that industrialized earlier rank higher in todays development, per-capita income and living standards. In the process of development human capital becomes relevant and investment in human capital acquisition reinforces through technological advancements, enabling economies to achieve economic growth (Galor 2005).

4.3.1 Industrialization

Table 5 examines the nationality of the owners and workers of industrial establishments in Argentina in 1895, 1913 and 1935. In 1895 81% of these establishments were own by foreigners, while 59% of the workers employed were immigrants. Close to twenty years later, in 1913, 65% of the industrial establishments were run by foreigners, while 49% of the employed workers were from foreign origin. Industry at that time was mostly centered around the production of garment, food, wooden, metal and chemical products, and construction. Table 5 also shows that still in 1935 58%[7] of the industrial establishments were under the ownership of foreign citizens.

Next I examine the number of workers and establishments, value of production[8] and installed energy capacity of industrial establishments in 1935. The 1935 industrial census records these information at the establishment level and county level. I analyze the number of establishments per person, percent of skilled workers in the population, per-capita value of production and energy in horse power per person[9]. Table 6 examines the relationship between the share of European immigration in 1914 and these variables in 1935,

[7]For 1935 detailed data on the nationality of ownership shows that at least 52% of the industrial establishments were owned by Europeans.

[8]In 1935 peso currency.

[9]For the per person variables I consider the 1914 population, since it is the closest population census.

panel A. shows IV estimates, while panel B. shows OLS estimates. The share of European population has a positive and significant effect on all industrial variables. Following columns 1-4, panel A., one standard deviation (SD) in the share of European population raises the value of industrial production by 0.83 SD, the share of skilled workers by 0.81 SD, the number of factories per person by 1.02 SD and the energy in horse power per person by 0.54 SD.

Tables 5 and 6 show the importance of European population in the process of industrialization, between 1895, 1914 and 1935 the fraction of industrial firms owned by Europeans has been above 50%, industrial workers have been overwhelmingly of European origin and counties that happened to have a greater share of their population of European origin experienced greater industrial output and assigned more resources to industry: workers and investment in energy production, as measured by the 1935 census. Moreover, as documented in table 6 columns 5, the share of European population has a significant effect on industrial production as measured by the Economic census in 1994.

Consistent with the results presented in the previous section, counties where the share of European population is greater experience more industrial output, a higher share of skilled workers and greater investments in installed energy in 1935.

4.3.2 Human Capital: Literacy rates in 1914

Did Europeans bring human capital or did they fostered the acquisition of human capital by the population at a large? As a measure for human capital in 1914 I will examine literacy rates, the percent of the population that is able to read and write. The previous analysis considered all Europeans as equal in contrast to the Argentine population, but differences within Europeans might also be relevant for explaining differences in development across counties. The industrial revolution started in the 18th century in England, while Italy's economy remained agrarian and differences in human capital were also present, literate population varied significantly across countries. In 1851 England and Wales between 30-33% of the population was illiterate, while in 1857 Spain 75% of the population was illiterate, and only 20% of the population in Prussia and Scotland was illiterate by 1849 and 1851 respectively (Cipolla (1969)).

The 1914 census shows these differences, table 7 examines literacy rates by nationality for immigrants in Argentina. Literacy varies quiet significantly between Argentineans and Europeans and also within Europeans. While the Argentinean population is on average 63.2% literate, Germans are 88.2% literate and immigrants from Italy, Spain and France (the greatest immigration groups) are 59.6%, 67.4% and 79.3% respectively. When weighted by population, on average Europeans are 64.2% literate and the population as a whole is 63.3% literate.

In table 8 I examine the relationship between the literacy rate in 1914 at the county level and the share of European population, column 1 panel A. shows IV estimates, while panel B. shows OLS estimates. As column 1 panel B. shows, at first Europeans appear not to have any effect on the literacy rates, a result that might be expected given the small difference between literacy rates of immigrants and natives. But once the endogenous distribution of immigrants is accounted for, the share of European immigration has a positive and significant effect on literacy rates, as column 1 panel A. shows. In particular, one standard deviation in the share of European population raises literacy rates by 0.21 SD.

The question that tables 7 and 8 raise is whether this difference in literacy rates across counties can be explained by a composition effect, namely if by substituting a less literate Argentinean by a more literate European, one can expect such an increase in literacy. As documented in table 7 on average Europeans are 1.1% more literate than Argentineans, implying that switching 1% European population for 1% Argentinean population will automatically raise literacy by 1.1%. The effect of 7.6% shown in table 8 column 1 is far greater than 1.1%, implying an effect beyond the composition effect, a positive externality on literacy rates.

European Immigration and Human Capital formation in 1914

Were counties with a higher share of European population more literate because of school availability? Did the Argentinean government promote education in areas with more Europeans to assimilate them to the native population? Are counties with higher literacy the results of public financed education, or the result of private financed educa-

tion?

Since mid-eighteenth century schools were built through the country by the government, offering free public education to all individuals in school-age (6 to 14 years old). These schools were mostly in urban areas or highly densely populated areas. Private schools were also present and offered religious learning and/or were present in areas without public schools. As shown below, areas with a higher share of European immigrants are associated with a higher number of private schools per schooling age population and a lower number of public schools.

In table 8, columns 2 and 3 I regress the number of private schools and public schools per 1000 school-age population on the share of European immigrants, controlling for county characteristics. Census data on schools in 1914 lists school's location and the school-age population in each county, from which I construct the number of schools per 1000 school-age children, on average there are 5.3 public schools and 0.85 private schools in each county per 1000 school-age population, with a standard deviation of 2.32 and 0.71, respectively. Column 2 panel A. shows IV estimates of regressing the number of private schools on the share of Europeans, results show a positive, although not significant, effect of immigrants on the quantity of schools, one standard deviation in the share of immigrants increases by 0.48 SD the number of private schools per school-age population.

However, one concern would be that the government invested in public education to facilitate the assimilation of immigrants. In column 3, I regress the number of public schools per school-age population on the share of European population, the share of European population has a negative and significant effect on the number of public schools. One standard deviation in the share of European population reduces the number of public schools by 0.95 standard deviations, equal to more than two schools. These findings are consistent with literacy rates being higher in areas with more Europeans because of education promoted by Europeans, and not because of educational policies pursued by the national government.

5 Conclusion

The period between mid eighteen hundred and the First World War saw an unprecedented flow of European immigrants to Argentina, mostly to the rural and urban areas across the fertile plains. Areas where Europeans accounted for a greater share of the total population managed to develop more than areas with fewer Europeans, as measured close to one hundred years later.

Why were areas with a higher share of European immigrants able to develop more than areas where Europeans represented a fewer share of the population? As I have discussed above, political institutions and geographical conditions can not explain these differences in development, the answer has to be found in the role played by human capital. When compare to native Argentineans, Europeans brought human capital, knowledge or skills complementary to industrial production. Europeans started most of the industrial activities and provided for most of the industrial (skilled and unskilled) workers.
Moreover, where Europeans accounted for a greater share of the population, the population had higher literacy rates. This higher literacy rates can not be explained by differences in literacy of Europeans and Argentineans, Europeans had a positive effect on literacy rates beyond what can be attributed to a composition effect. Higher literacy rates can not be explained by an effort of the national government to educate and assimilate immigrants, since public schools were less available in counties were Europeans accounted for a higher share of the population, although the number of private schools was higher. Private schools were created either by Argentineans or immigrants, but private schools' availability was greater in counties with a higher share of Europeans. Europeans generated a positive externality on the society as a whole, providing for greater literacy rates, even when public schools were less available.

These results point to the importance of human capital for economic development. The study of the fertile plains of Argentina, an area with equal political institutions and uniform geographical characteristics, shows that there is a long-term impact of initial differences in human capital on economic development.

References

[1] Acemoglu, D., S. Johnson, and J.A. Robinson (2001): "The colonial origins of comparative development: an empirical investigation," *American Economic Review*, 91:1369–401.

[2] Acemoglu, D., S. Johnson, and J.A. Robinson (2002): "Reversal of fortunes: geography and institutions in the making of the modern world income distribution," *Quarterly Journal of Economics*, 117:1231–94.

[3] Banerjee, A., and L. Iyer (2005): "History, institutions and economic performance: the legacy of colonial land tenure systems in India," *American Economic Review* 95: 1190–213.

[4] Cipolla, C.M. (1969): *Literacy and Development in the West.* Baltimore: Penguin Books.

[5] Cruzate, G., L. Gomez, M. J. Pizarro, P. Mercuri, S. Banchero, G. Angueira, D. Prieto, J. Lopez and G. Ibarraza (1990): "Suelos de la República Argentina, Versión 1.0." Versión digital corregida y ajustada mediante técnicas actuales de ingeniería SIG en base a la información original vectorizada a partir de los mapas de suelos provinciales que integran el *Atlas de Suelos de la República Argentina*, INTA.

[6] Dell, M., (2010): "The Persistent Effects of Peru's Mining Mita," *Econometrica* 78(6), 1863–1903.

[7] Dirección General de Inmigración (1925): *Resumen Estadístico del Movimiento Migratorio en la República Argentina. Años 1857 - 1924.* Buenos Aires: Talleres Graficos del Ministerio de Agricultura de la Nacion.

[8] Easterly, W., and R. Levine(2009): "The European Origins of Economic Development," mimeo, Brown University.

[9] Engerman, S.L., and K.L. Sokoloff (1997): Factor endowments, institutions, and differential paths of growth among New World economies: a view from economic

historians of the United States. In *How Latin America Fell Behind*, ed. S Harber, pp. 260–304. Stanford, CA: Stanford Univ. Press.

[10] Engerman, S.L., and K.L. Sokoloff (2002): Factor endowments, inequality, and paths of development among New World economies. *NBER*, Working Paper 9259.

[11] Gallo, E. (1983): *La Pampa gringa: la colonización agrícola en Santa Fe (1870-1895).* Buenos Aires: Editorial Sudamericana, 1983.

[12] Galor, Oded (2005): "From Stagnation to Growth: Unified Growth Theory," *Handbook of Economic Growth.*

[13] Galor, O., O. Moav and D. Vollrath (2009): "Inequality in land ownership, the emergence of human capital promoting institutions and the great divergence," *Review of Economic Studies* 76: 143–79.

[14] Gennaioli, N., R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2011): "Human Capital and Regional Development," *NBER Working Paper* No. 17158.

[15] Glaeser, E.L., R. La Porta, F. Lopez-De-Silanes and A. Shleifer (2004): "Do institutions cause growth?" *Journal of Economic Growth*, 9: 271–303.

[16] Lacoste, Pablo (2002): "La guerra de los mapas entre Argentina y Chile: una mirada desde Chile," *Historia (Santiago)*, ISSN 0717–7194, vol. 35, pages 211–249.

[17] Luna, Félix (1993):"Breve Historia de los Argentinos," Buenos Aires: Planeta.

[18] República Argentina (1938): *Censo Industrial de 1935.* Buenos Aires: Dirección General de Estadística de la Nación.

[19] República Argentina (1898): *Segundo Censo de la República Argentina, Mayo 10 de 1895.* Buenos Aires: Taller Tipográfico de la Penitenciaria Nacional.

[20] República Argentina (1916): *Tercer Censo Nacional de la República Argentina, Levantado el 1^{ro} de Junio de 1914.* Buenos Aires: Talleres Gráficos de L. J. Rosso y Cía..

[21] Sarmiento, D. F. (2004): *Facundo : Civilization and Barbarism.* Ewing, NJ, USA: University of California Press.

[22] Somoza, J. L., and A. E. Lattes (1967): *Muestras de los dos primeros censos nacionales de población, 1869 y 1895.* Buenos Aires: Instituto Torcuato Di Tella, Centro de Investigaciones Sociales, Documento de Trabajo no 46.

[23] Tell, S. (2008): *Córdoba rural, una sociedad campesina (1750-1850).* Buenos Aires: Prometeo Libros, Asociación Argentina de Historia Económica, 2008.

[24] Walther, J. C. (1964): *La Conquista del Desierto.* Buenos Aires, Círculo Militar, 1934.

Appendix

1 GDP at the county level

The Argentinean Statistical Office (INDEC) computes GDP at the national- and *provincial*- (state) level, but not at the county level. Provincial Statistical Offices may or may not compute GDP at the county level, varying in methodology and accuracy of their procedures. Only two *provinces*, Buenos Aires and Santa Fe, compute GDP at the county level, out of four *provinces* that covers this study: Buenos Aires, Santa Fe, Córdoba and Entre Ríos.

During 1994 INDEC conducted the National Economic Census ("Censo Nacional Económico" or CNE), censing all business at the county level. The census gathered information on production, employment, revenue, costs, investment, etc. The unit of observation was the *premises*, the physical space used for an economic activity. The statistical unit of observation was the firm. All businesses in Oil and Natural Gas, Mining, Manufacturing Industries, Electricity, Gas and Water, Retail and Wholesale, Financial Intermediation, Communication, Eterprise Service Providers and Personal Service Providers, were censed.

The CNE records the value of the gross product at the county level, for the above mentioned sectors. This number by itself proxies for GDP at the county level. Indeed, for those *provinces* that compute GDP at the county level, the correlation between the two measures is 94.68% for all counties in Buenos Aires and Santa Fe, and 95.95 % for the sample of counties used in this study (106 observations).

Since agriculture is not being censed and because of its importance for this area, not accounting for its value may be particularly problematic. Yearly agricultural output esti-

mates are performed by the Ministry of Agrigulture (MAGyP) at the county level. Adding the value of the agricultural output to the CNE gross product accounts for this relevant sector, although with one drawback. Since I only observe the value of the agricultural output, I am not discounting for intermediate goods and services used by the agricultural sector, thus overstating the value of the agriculture.

For each *province* I consider the major agricultural products (Wheat, Soybean, Corn, etc.) in 1994, accounting for 84% to 96% of the sowed area. For each county I compute the value of agricultural output as the sum of each crop times its price (prices are from FAOstat). The CNE gross product augmented by the agricultural output is used as a proxy for the GDP. For those *provinces* that compute GDP at the county level, the correlation between the two variables is 94.65%for all counties in Buenos Aires and Santa Fe, and 95.05 % for the sample of counties used in this study (106 observations).

The correlation between the two proxies for GDP with- and without agricultural output is 99.40% for all counties and 96.56% for the sample of counties used in this study (106 observations). The regression of actual GDP on the proxy for GDP has an R2 of 90.34 (adding *province* fixed effects does not change this result) a coefficient of 1.04 with a standard error of 0.033.

Figure 1

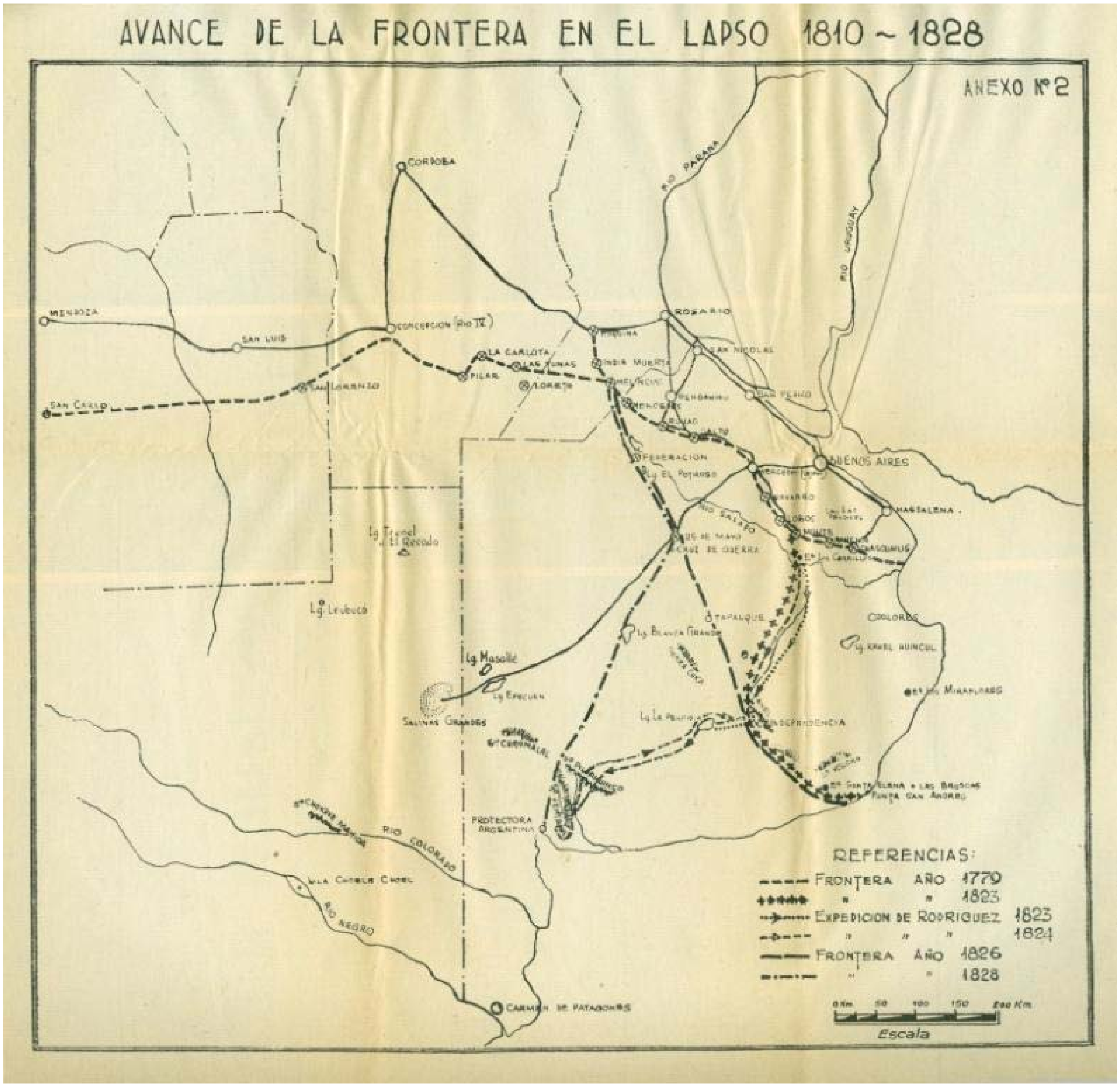
AVANCE DE LA FRONTERA EN EL LAPSO 1810 ~ 1828
ANEXO Nº 2
CORDOBA
RIO PARANA
RIO URUGUAY
MENDOZA
SAN LUIS
ROSARIO
SAN NICOLAS
LA CARLOTA
PILAR
LORETO
SAN CARLOS
FEDERACION
BUENOS AIRES
MAGDALENA
RIO SALADO
TAPALQUE
Lg. Leubucó
Lg. Trenel
Lg. Masallé
Lg. Epecuen
SALINAS GRANDES
RIO COLORADO
RIO NEGRO
PROTECTORA ARGENTINA
CARMEN DE PATAGONES
REFERENCIAS:
FRONTERA AÑO 1779
" " 1823
EXPEDICION DE RODRIGUEZ 1823
" " " 1824
FRONTERA AÑO 1826
" " 1828
Escala

Figure 2

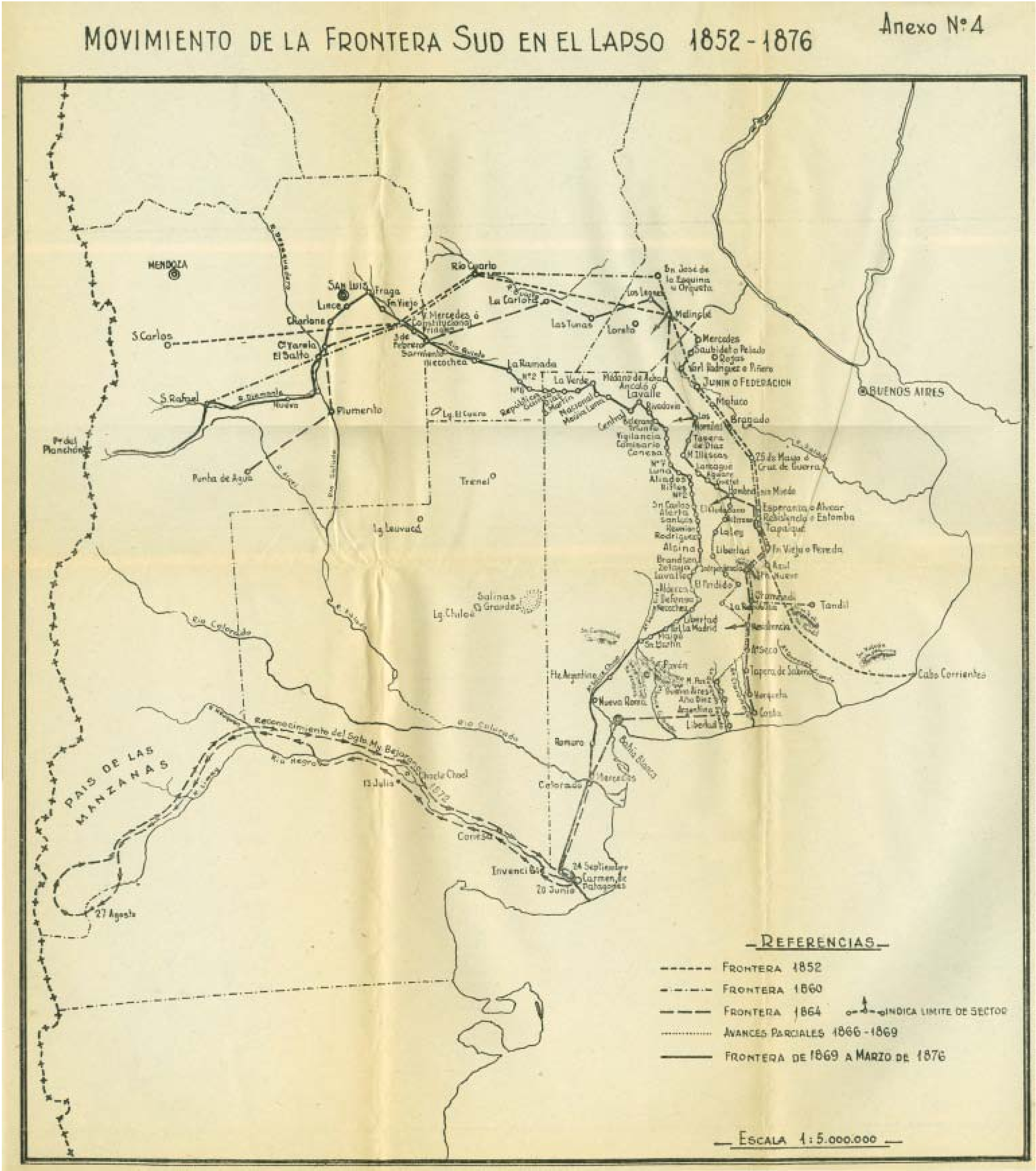
MOVIMIENTO DE LA FRONTERA SUD EN EL LAPSO 1852 - 1876
Anexo N°4
MENDOZA
SAN LUIS
Río Cuarto
BUENOS AIRES
Tandil
Cabo Corrientes
Bahía Blanca
Carmen de Patagones
PAIS DE LAS MANZANAS
Río Colorado
Río Negro
REFERENCIAS.
FRONTERA 1852
FRONTERA 1860
FRONTERA 1864
INDICA LIMITE DE SECTOR
AVANCES PARCIALES 1866 - 1869
FRONTERA DE 1869 A MARZO DE 1876
ESCALA 1:5.000.000

Figure 3: Immigration Time Series.

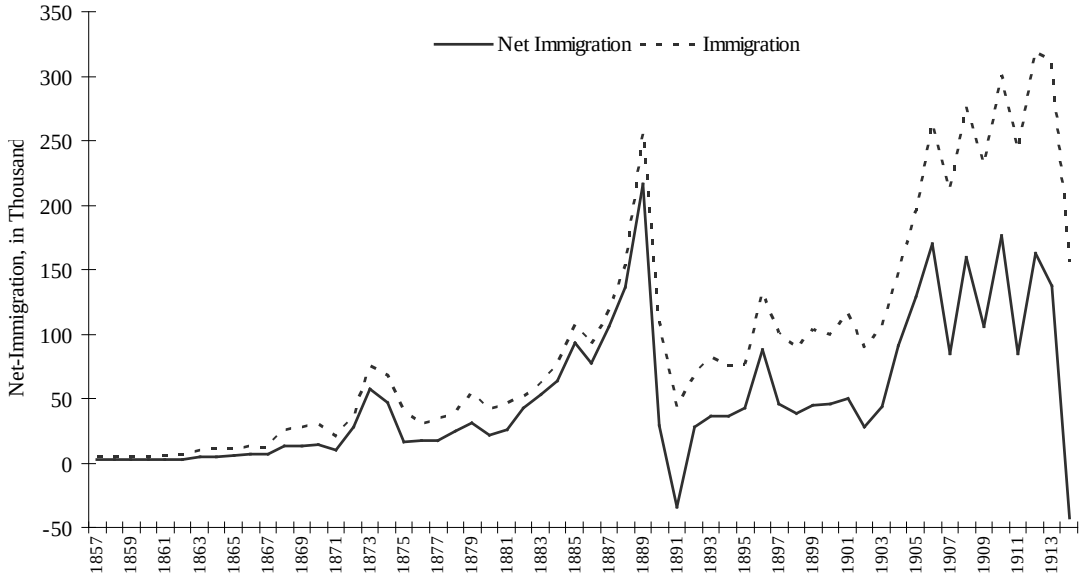

Figure 4: Cumulative Net-Immigration and Area for settlement.

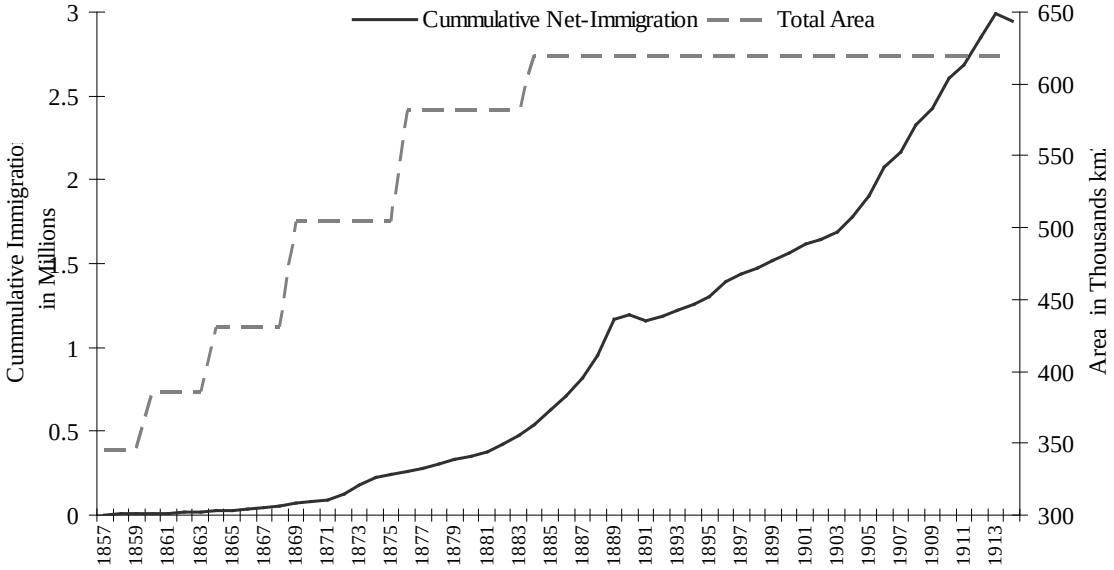

Figure 5: 1st Stage correlation between the share of European population and the constructed share of European immigration.

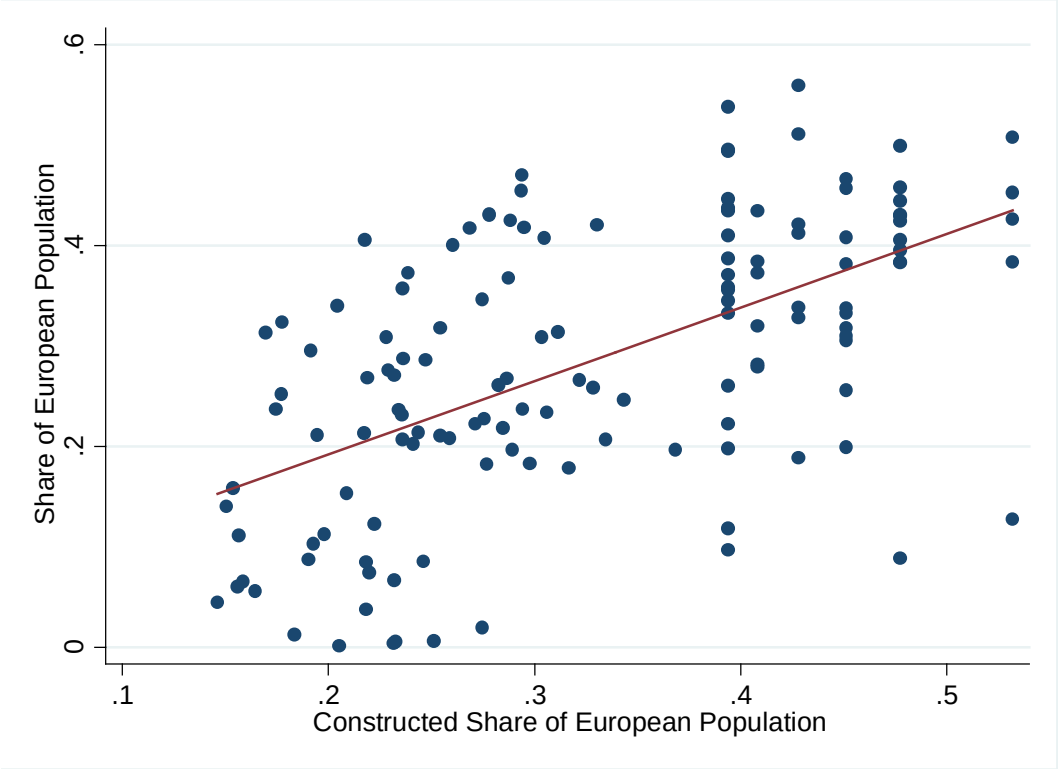

Figure 6: 1st Stage correlation between the share of European population and the constructed share of European immigration, control variables and fixed effects included.

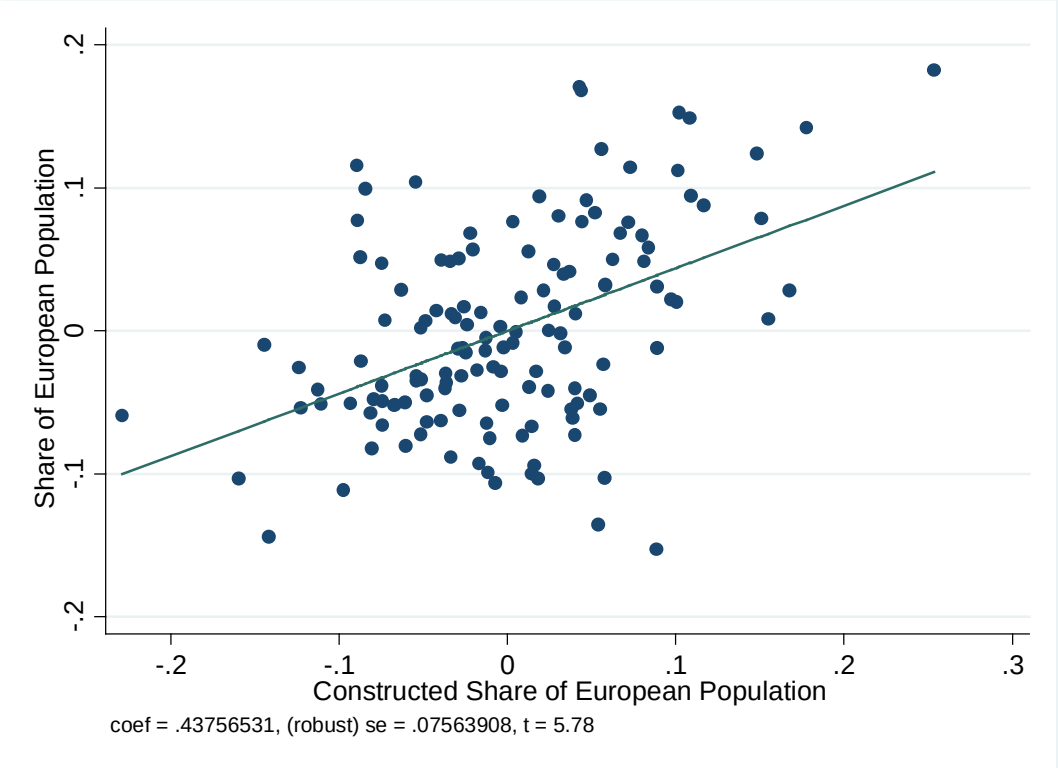

Table 1

Variable	Number of Observations	Mean	Standard Deviation	5th Percentile	25th Percentile	50th Percentile	75th Percentile	95th Percentile
Share of European population, 1914	136	0.29	0.14	0.04	0.20	0.30	0.41	0.49
GDP per-capita, 1994	136	6754	4190	1288	3560	6418	9213	13301
log GDP per-capita, 1994	136	8.59	0.78	7.16	8.18	8.77	9.13	9.50
Share of pop. w/higher education, 2001	136	0.10	0.02	0.07	0.09	0.10	0.11	0.13
Share of skilled workers, 2001	136	0.18	0.04	0.09	0.15	0.18	0.21	0.23
log industrial output per-capita, 1994	136	5.40	1.45	2.61	4.49	5.52	6.49	7.26
log industrial output per-capita, 1935	136	4.40	1.14	2.97	3.87	4.38	4.99	6.20
Share of skilled workers, 1935	136	1.99	2.06	0.26	0.89	1.40	2.45	5.60
Number of factories per-capita, 1935	136	3.69	2.16	0.66	2.16	3.31	4.94	7.30
Energy in H.P. per-capita, 1935	136	0.10	0.14	0.01	0.05	0.07	0.10	0.29
Literacy rate, 1914	136	0.63	0.05	0.55	0.58	0.65	0.67	0.68
Number of private schools per-1000 school age p	136	0.85	0.71	0.00	0.35	0.74	1.20	2.28
Number of puclic schools per-1000 school age p	136	5.33	2.32	2.70	3.63	4.78	6.56	9.72
Distance to BA City	136	5.71	0.65	4.41	5.29	5.84	6.16	6.59
Land Quality Index	136	45.43	16.39	20.58	33.06	45.10	54.30	72.90
Percent of Land used for Agriculture	136	0.28	0.23	0.01	0.07	0.26	0.44	0.72
Population Density	136	6.67	5.53	1.52	2.78	5.22	7.88	17.38
Urban Rate	136	0.33	0.18	0.00	0.22	0.32	0.44	0.63

Table 2

Dependent Variable:	OLS							
	log per capita GDP, 1994							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
European population / total population, 1914	4.699*** (0.518)	4.830*** (0.523)	4.051*** (0.591)	3.781*** (0.618)	4.709*** (0.539)	4.595*** (0.478)	3.720*** (0.644)	2.698*** (0.520)
Distance to BA City		-0.139 (0.085)					-0.113 (0.097)	-0.005 (0.081)
Land Quality			0.011*** (0.004)				0.010* (0.005)	0.007 (0.005)
Percent of Land used for Agriculture in 1914				0.697*** (0.251)			0.546* (0.314)	0.919*** (0.275)
Population Density in 1914					-0.001 (0.009)		-0.029*** (0.009)	-0.025*** (0.009)
Urban Rate in 1914						0.216 (0.301)	0.501 (0.331)	0.596* (0.344)
Constant	7.412*** (0.228)	8.286*** (0.522)	7.105*** (0.228)	7.400*** (0.230)	7.413*** (0.229)	7.384*** (0.251)	7.818*** (0.593)	7.247*** (0.460)
Province Fixed Effects	yes	yes	yes	yes	yes	yes	yes	no
Observations	136	136	136	136	136	136	136	136
R-squared	0.513	0.518	0.544	0.533	0.510	0.511	0.558	0.543

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3

Dependent Variable:	First Stage European population / total population		
	(1)	(2)	(3)
Simmulated Share of European Immigration	0.731*** (0.090)	0.469*** (0.089)	0.438*** (0.076)
Distance to BA City			0.041*** (0.013)
Land Quality			0.001 (0.001)
Percent of Land used for Agriculture in 1914			0.248*** (0.035)
Population Density in 1914			0.003* (0.001)
Urban Rate in 1914			0.141*** (0.040)
Constant	0.046 (0.031)	0.046 (0.031)	-0.337*** (0.079)
Province Fixed Effects	no	yes	yes
Observations	136	136	136
R-squared	0.322	0.453	0.777

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4

Dependent Variable:	log per-capita GDP, 1994	share of population with hgher education, 2001	share of population with high skilled occupations,
	(1)	(2)	(3)
		Panel A. IV	
European population / total population, 1914	5.655***	0.065*	0.183***
	(1.178)	(0.036)	(0.057)
		Panel B. OLS	
European population / total population, 1914	3.720***	0.052***	0.171***
	(0.644)	(0.020)	(0.035)
Control variables	yes	yes	yes
Province Fixed Effects	yes	yes	yes
Observations	136	136	136

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5

	year	Foreigners	Argentineans	Share of Foreigners
		number		
Ownership	1895	19,610	4,504	0.81
	1913	31,483	15,763	0.65
	1935	30,322	20,663	0.58
Workers	1895	103,291	72,391	0.59
	1913	200,578	209,623	0.49

Table 6

Dependent Variable:	log value of industrial production, 1935	share of skilled workers in the	factories per person, 1935	energy in h.p. per person, 1935	log value of industrial production, 1994
	(1)	(2)	(3)	(4)	(5)
			Panel A. IV		
European population / total population, 1914	6.742***	11.988**	15.794***	0.693**	5.582**
	(2.113)	(4.832)	(4.530)	(0.306)	(2.749)
			Panel B. OLS		
European population / total population, 1914	4.502***	7.101***	8.962***	0.351**	4.342***
	(1.293)	(2.257)	(2.102)	(0.168)	(1.392)
Control variables	yes	yes	yes	yes	yes
Province Fixed Effects	yes	yes	yes	yes	yes
Observations	136	136	136	136	136

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7

Nationality	Literacy rate
Argentina	63.2%
Average Europe	64.2%
Average Population	63.3%
Austria	69.2%
France	79.3%
Germany	88.2%
Great Britain	90.9%
Italy	59.6%
Russia	55.7%
Spain	67.4%
Switzerland	86.9%

Table 8

Dependent Variable:	share of literate population, 1914	Private Schools x 1000 school-age population,	Public Schools x 1000 school-age population,
	(1)	(2)	(3)
		Panel A. IV	
European population / total population, 1914	0.076**	2.426	-15.714***
	(0.031)	(1.499)	(4.101)
		Panel B. OLS	
European population / total population, 1914	-0.003	2.639***	-9.345***
	(0.015)	(0.734)	(2.212)
Control variables	yes	yes	yes
Province Fixed Effects	yes	yes	yes
Observations	136	136	136

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1