Papers in Evolutionary Economic Geography

13.05

European migration, national origin and long-term economic development in the US

Andrés Rodríguez-Pose and Viola von Berlepsch

Utrecht University
Urban & Regional research centre Utrecht

European migration, national origin and long-term economic development in the US

by

Andrés Rodríguez-Pose and ***Viola von Berlepsch***

London School of Economics
Contact: a.rodriguez-pose@lse.ac.uk

Abstract

Have Irish, German or Italian settlers arriving in the US at the turn of the 20th century left an institutional trace which determines economic development differences to this day? Does the national origin of migrants matter for long-term development? This paper explores whether the distinct geographical settlement patterns of European migrants according to national origin affected economic development across US counties. It uses micro-data from the 1880 and 1910 censuses in order to identify where migrants from different nationalities settled and then regresses these patterns on current levels of economic development, using both OLS and instrumental variable approaches. The analysis controls for a number of factors which would have determined both the attractiveness of different US counties at the time of migration, as well as current levels of development. The results indicate that while there is a strong and positive impact associated with overall migration, the national origin of migrants does not make a difference for the current levels of economic development of US counties.

Keywords: Migration, National/Ethnic Origin, Institutions, Culture, Economic Development, Counties, USA

JEL classification: F22, O15, R23, N91

European migration, national origin and long-term economic development in the US

Introduction

Does the nationality of migrants arriving in any particular territory make a difference for long-term economic development? Is being settled by individuals from areas of the world with what are perceived to be better institutions, such as England, Germany or Scandinavia, a blessing for long-term economic development? Does having an Italian, Spanish or Polish institutional heritage limit a territory's potential for growth? Researchers such as Véliz (1994) Landes (1998), La Porta et al. (1998, 1999), North et al. (2000) or Acemoglu et al. (2001) believe this to be the case. In their view, the institutions that conquerors and colonists brought with them combined with their attitude to the place where they settled down have determined, to a large extent, the economic trajectory of the places they conquered. From their point of view, former British colonies in the Americas have performed better than former Spanish, Portuguese and French colonies mainly as a consequence of the superiority of the institutional heritage the British left behind.

However, while the link between the colonial heritage and economic performance has attracted considerable attention, another, generally more peaceful, way of colonising and transforming the institutions of territories, mass migration, has been somewhat overlooked by the literature. Mass migration can, nevertheless, reshape societies in a similar or even more extensive way than colonisation. Particularly in the case of North America – as in other white settlement areas, such as Australia or New Zealand – the number of Europeans arriving after independence massively exceeded that of the colonisers. The distinct geographical settlement

patterns of migrants, mostly following the footsteps of their co-nationals, has helped transform specific districts, cities, regions and even countries into replicas of the countries they left behind. In many ways, Argentina – where 60% of the population claims to be of Italian descent – can be considered as a reproduction of Italy in the southern hemisphere.

No country in recent history has attracted as many migrants as the US in the period between the late 19th and early 20th century. Between 1860 and 1920 close to 27 million Europeans left the 'old' continent to move to the US (Ward, 1987). Migrants came from all over Europe. A first migration wave in the late 19th century was dominated by Germans, Irish, Scandinavians, and English. Italians, Polish, and Russians were the largest groups in the second wave of the early 20th century. Migrants, however, rather than mingle, followed their country folk, leading to very distinct patterns of settlement by nationality. This resulted in the formation of strong ethnic communities from specific national or even local origins in different parts of the US. In sufficient numbers, migrants were able to shape the culture, institutions and identity of places in the US in light of what they had left back at home. But did these differences in migrant settlement patterns leave a legacy in terms of economic development which can still be felt today? Have US counties mainly populated by migrants from countries with more efficient institutions – say from Germany or Scandinavia – performed better than those where migrants from countries with weaker institutions – say, Italy, Poland, Ireland – settled en masse?

This paper addresses these questions using historical microdata – provided in the Integrated Public Use Microdata Series (IPUMS) database – of the US censuses which coincided with the greatest waves of European migration (1880, 1910) in order to uncover the settlement patterns of migrants of different national origins across the more than 3,000 US counties. We look at the national origin of migrants living in any given US county and at their percentage in the local population of the county at the turn of the 19th century so as to assess whether a greater presence of migrants from specific national origins has influenced and/or continues to

influence development patterns across US counties to this day. We further control for other factors which may have shaped economic development in the individual counties of the US, using income per capita, population, the unemployment rate, the educational attainment of the population, the percentage of blacks, female labour participation, and employment in agriculture as independent variables.

Our expectation, based on the literature, is that the presence of large contingents of migrants with similar national and/or ethnic origins would have left a distinct cultural and institutional imprint, which may have affected subsequent economic development. However, the results of the analysis show that this is not the case. While there is a strong and positive correlation between where migrants settled and current levels of development, this correlation seems to be completely independent of the national origin of migrants. Specific migrant origins – with the only exception of the English, precisely those with the least problems of adaptation to the new environment – do not make a difference for long-term economic development whatsoever. This holds for the first and second wave of migration. Hence being settled by Germans, Irish, Scandinavians, Poles, or Italians has always been favourable for subsequent economic development. The main difference is simply between counties which received a large influx of migrants, which tend to be significantly richer today, and those that did not.

In order to achieve these results, the paper is structured as follows. In the next section we examine the European settlement patterns in the US, paying special attention to differences according to national origin. The third section analyses the links between the formation of national enclaves, institutions, and economic development from a theoretical perspective. The empirical model and data are presented in section four, which is followed by the discussion of the results of the regression analysis. The final section concludes and highlights some potential policy implications.

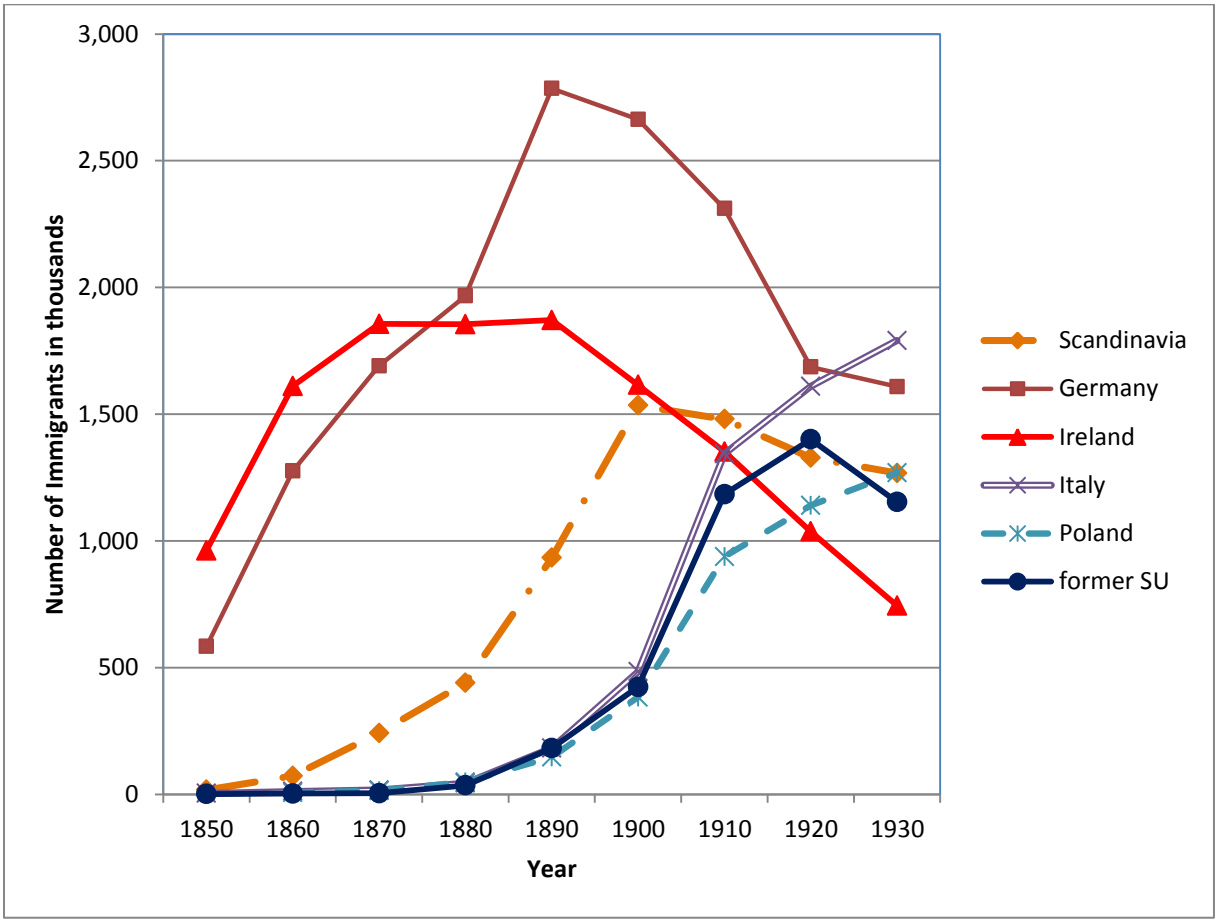
European settlement patterns in the United States

Between 1860 and 1920, millions of Europeans left their homes for lands that would allow them, in different degrees, to escape poverty, famine, drought, political upheavals and religious persecution. Two thirds of the reported 40 million migrants who left Europe for the 'new world' ended up in the United States (US) (Bertocchi and Strozzi, 2006). Accordingly, the number of foreigners living in the US increased from around 4 million in 1860 to just under 14 million in 1920 (Gibson and Jung, 2006).

Two main migration waves can be distinguished during that period (Hatton and Williamson, 1994). The first wave – also known as the old migration – covers the period ranging from the pre-Civil War years to around 1890. These so-called "pioneers of the century of immigration" (Daniels 1990: 121) mainly stemmed from northern and western European countries. The earliest contingents were dominated by the English. The English were followed by Irish, Germans and Scandinavians. 4.7 out of 5.5 million foreign-born living in the US declared these countries as their origin in the 1870 census (Alexander, 2007).

The second wave of migration started towards the end of the 19th century. It was more intense than the first wave and was characterized by a radical shift in the national origin of migrants. The English, German, Irish and Scandinavians of the first wave were replaced by southern and eastern Europeans. Italians, Polish, Russians, Greeks or Portuguese entering the US soon outnumbered those coming from first wave countries. By 1896 there were already more migrants from second than from first wave countries. The number of southern and eastern Europeans settling in the US between 1891 and 1920 doubled that of western and northern Europeans (Alexander, 2007). This shift in migration waves becomes evident when plotting the evolution of migrants from different national origins.

Figure 1: Immigration from six European Countries into the United States 1850 – 1930



Source: Gibson and Jung, 2006; own calculations

In Figure 1 we examine the presence of migrants in the US stemming from three first wave – Germany, Ireland, Scandinavia – and three second wave – Italy, Poland and the former Soviet Union – European countries or groups of countries. The number of Irish living in the US peaked as early as 1870 and their number remained stable until 1890, reflecting the large influx of Irish migrants in the 1860s, 1870s and 1880s. From 1890 onwards the overall size of Irish nationals in the US declined, as they became naturalized and their children were born American. The number of Germans – the largest migrant community in the US – reached its peak in 1890, but fell rapidly after 1900. Scandinavians arrived en masse in the 1880s and peaked in 1900. Until 1890, the large majority of foreigners living in the US stemmed from

these countries, plus Great Britain and, to a lesser extent, from other western European nations such as France or the Netherlands. Southern and eastern Europeans in 1890 represented a mere anecdote (Figure 1). Yet within the space of a few years the panorama had changed radically. Following two decades of rapid immigration from southern and eastern Europe, the number of Italians living in the US already matched that of the Irish. Russians and Poles were not far behind. By 1930, Italians had become the largest foreign community, surpassing the Germans, while Poles and peoples from the former Soviet Union represented the third and fifth largest foreign group respectively (Figure 1).

Although the motives for migrating varied from individual to individual, there were a series of common traits to a large majority of migrants. Most migrants during the period of analysis tended to be male, single, young, poor and unskilled (Carlsson, 1976). This was the case regardless of national origin, as described, for example, by Bodnar (1992) for Swedish migrants or Fitzpatrick (1983), when describing Irish migration. English migrants also fitted the bill, but tended to be slightly more skilled and spoke the language (Erickson, 1972). The majority of those leaving their homelands between 1850 and 1920 came from rural or proto-industrial areas "where there were many agriculturalists but little agriculture" (Kamphoefner, 1976: 182). A combination of rapid population growth in their places of origin coupled with a system of partible inheritance, decreasing profits of farming, and, in some cases, a progressive exhaustion of the soil, drove landless farmhands to take a one-way passage to America. In Sweden, for instance, by 1870, 48% of the agricultural workforce was landless (Daniels, 1990). Further restrictions in access to land in the 'old world' and a growing incapacity to absorb a rapidly growing labour force pushed many Germans, Irish, Poles and southern Italians to migrate in search of new and different employment opportunities in the US (Fitzpatrick, 1983).

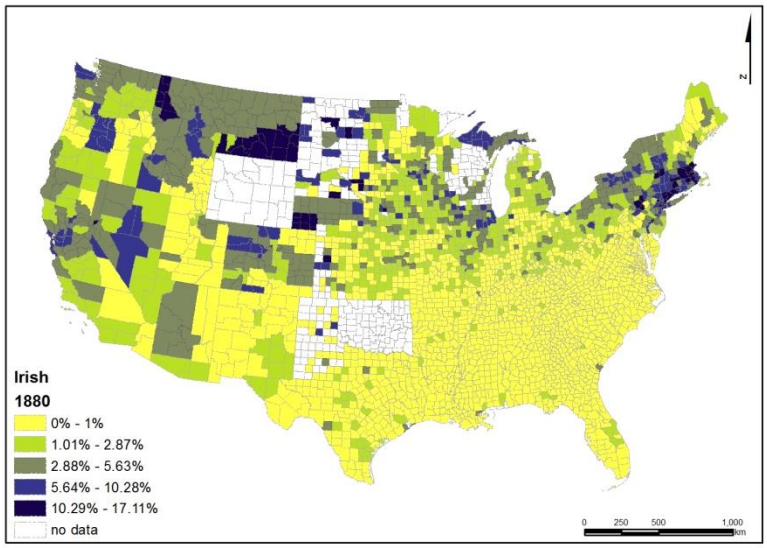
Young, poor, unskilled with limited formal education and, in the majority of cases, no knowledge of the language: the characteristics of the average migrant represented a particularly adverse cocktail for a rapid integration in their new society. With the partial exception of the English – who already spoke the language, were more familiar with local institutions and the system, and were often more skilled and formally educated – migrants initially relied heavily on the support of their country folk. Those who had previously settled in the US already knew the system and were a rich source of information about crucial issues such as living conditions, access to land, jobs and wages. They acted as a magnet for new migrants not just from their countries of origin, but also from their same region and village. Newcomers flocked to where their relatives, friends or acquaintances had settled before. This chain migration led to the formation of national and/or ethnic and even local enclaves across the US and of tightly-knit ghettos in cities (e.g. Vedder and Gallaway, 1972; Levy and Wadycki, 1973; Dunlevy and Gemery, 1977). "Regardless of when and where they entered the country, most people knew exactly where they were going, and most were headed for locales where kith and kin already lived and worked" (Alexander, 2007: 28).

The concentration of migrants in these national and/or ethnic communities across counties and states generated a distinct geography of migration according to national origin.

The Irish, who represented 7 out of 10 migrants in 1860 (Daniels, 1990), rarely wandered well beyond the main ports of entry. The majority clustered almost exclusively in and around north-eastern cities, predominantly in New England. By 1870, 15% of the population of the 50 largest US cities claimed to have Irish roots. In the nine largest cities of New England, New York state and New Jersey more than 20% of the population was Irish. In Boston, for instance, 64,793 people – 56% of the foreign born population – declared Ireland to be their place of birth in 1880 (Daniels, 1990). Figure 2 shows the location of Irish migrants by county in 1880, the height of Irish presence in the US. The counties with the highest incidence of

Irish migrants tend to be located in the Boston-Philadelphia axis, with particularly strong concentrations around Boston-Providence, New York and Hartford. Irish presence could also be detected in Illinois, northern Indiana and out West in sparsely populated places in south eastern Montana or in Nevada.

Figure 2: Percentage of Irish Migrants in the US, 1880

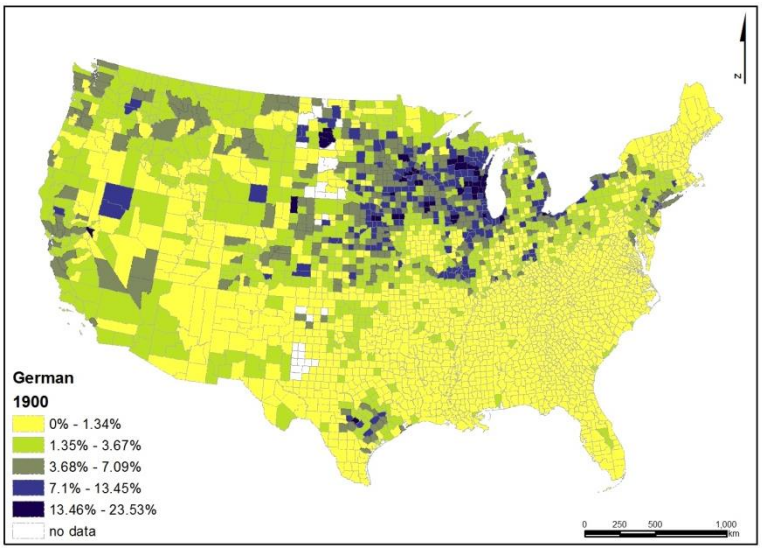


Source: Own elaboration

Germans looked for land and, in contrast to the Irish, avoided the cities of the North East. The first German migrants headed for rural areas in New York state and later spread out to rural states south of the Great Lakes and in the mid-west, most notably within the so called 'German triangle' (Daniels, 1990) between Saint Louis, Cincinnati and Milwaukee. Only two fifths ended up settling in cities, where they concentrated in ethnic enclaves, such as 'Kleindeutschland' in New York or 'Over the Rhine' in Cincinnati. These ethnic quarters "replicated the culture of the homeland" (Daniels, 1990: 150).

Figure 3 shows the location of German migrants in 1900, ten years after their numbers peaked.[1] Illinois, Iowa, Minnesota and, above all, Wisconsin were heavily settled by Germans, who also made up large communities in Michigan, Pennsylvania, Ohio and central Texas.

Figure 3: Percentage of German Migrants in the US, 1900



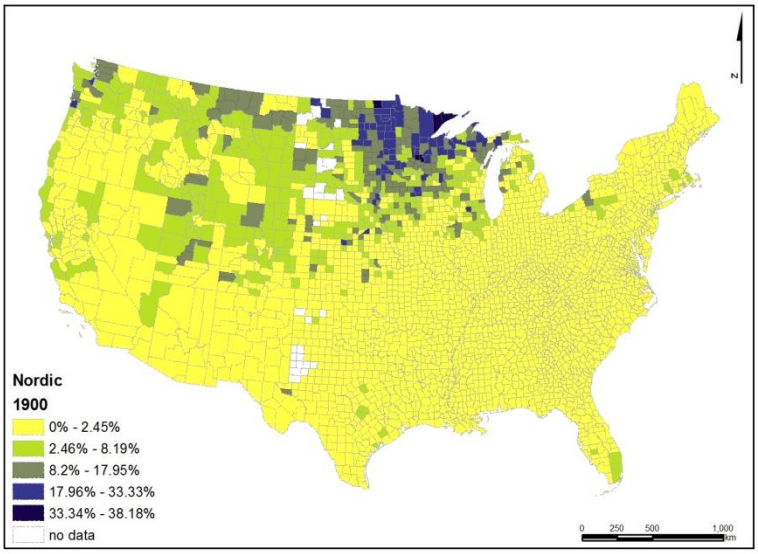
Source: Own elaboration

The third largest group in the first migration wave was made of Scandinavians. Although smaller than German or Irish migration, a total 2.15 million Scandinavians – more than half of them Swedes – ended up in the US, making 7% of the country's population at the time (Daniels, 1990). Early Swedish, Danish and Norwegian migrants clustered in rural areas in the upper wheat belt in the Midwest, especially Minnesota, the Dakotas, Wisconsin and Nebraska. By the 1890s and 1900s some of them had also settled down along the border with Canada, all the way to Washington state, and were colonising cities, such as Chicago, Minneapolis and Jamestown (New York) (Figure 4). In 1900, 9% of the population of Chicago stemmed from Sweden, making it the second largest Swedish city in the world at the

[1] The figures for the 1890 census are not available.

time (Daniels, 1990). They, however, remained the most rural group of foreign migrants (Ward, 1987).

Figure 4: Percentage of Scandinavian Migrants in the US, 1900



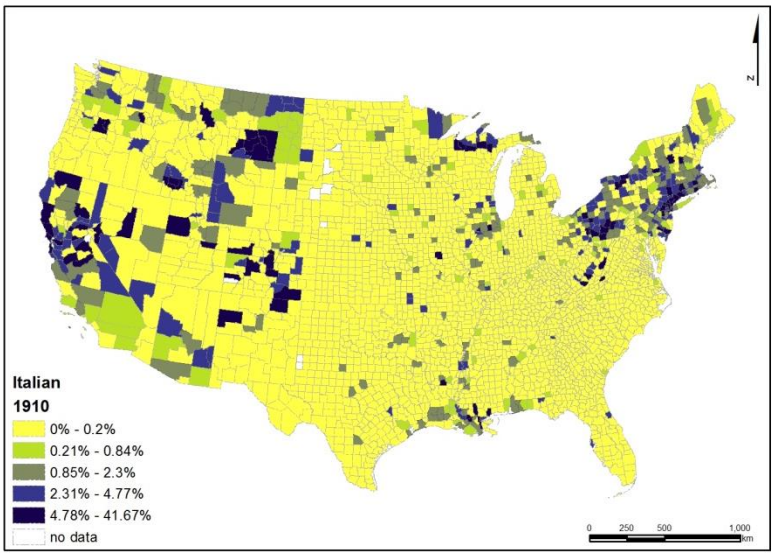
Source: Own elaboration

The English – perhaps as a consequence of their easier capacity to become integrated in their new environment – became less geographically clustered and were the most evenly distributed across the geography of the US (Ward, 1987). Although early English settlers gravitated to eastern and north-central states, later migrants spread out to states west of the Mississippi River (Alexander, 2007).

The second wave of southern and eastern Europeans also displayed strong tendency to converge into specific parts of the US and, particularly, into urban areas, were 88% of them clustered (Daniels, 1990). Italians, with 4.1 million new entrants between 1880 and 1920 represented the largest group of new wave immigrants. They concentrated in and around New York City. By 1920, almost four hundred thousand Italians lived there (Daniels, 1990). This amounted to almost one quarter of all Italians living in the US. They formed strong Italian quarters and enclaves in the city and in the neighbouring state of New Jersey. Other north

eastern cities, such as Boston, Philadelphia or Providence contained large Italian communities (Figure 5). California and, to a much lesser extent, other western states also displayed a non-negligible Italian presence. In these areas however, Italians focused more on agriculture than on small shop businesses or manual labour, as was the case in the north eastern cities (Alexander, 2007). Other southern Europeans – Greeks, Portuguese and Spaniards – came in much smaller numbers. Greeks tended to stay in the cities of New England and, in particular, in Massachusetts, with some presence in the Mid-West, fundamentally around Chicago. The Portuguese and the Spaniards flocked to California and Nevada, although some strong Portuguese fishing communities developed in Rhode Island and southern Massachusetts. Maps indicating the location of the remaining national and ethnic groups at their peak are included in Annex I.

Figure 5: Percentage of Italian Migrants in the US, 1910

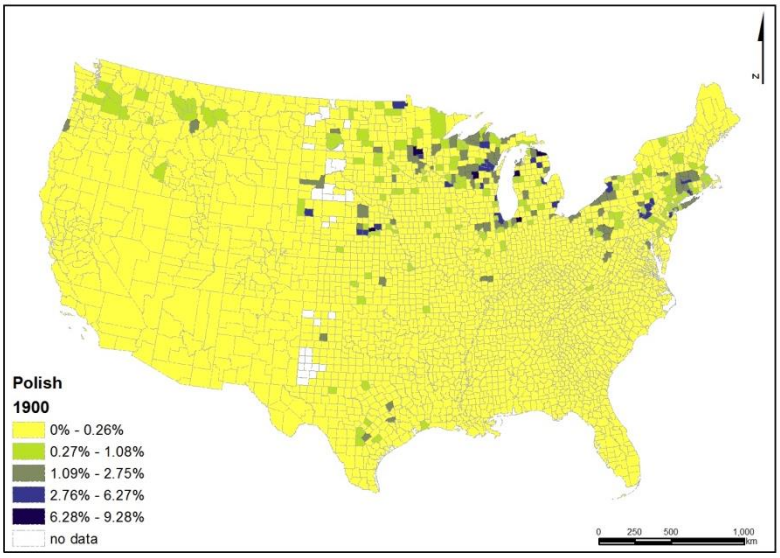


Source: Own elaboration

The second largest group of migrants during the second wave was made of Slavs, principally Poles and Russians. This group headed primarily for rural areas. The pioneers of the 2.5 million Poles who settled in the country established themselves in the rural regions of the

Midwest and Texas. Later arrivals preferred cities along the 'rust belt'. Chicago alone was home to about 400,000 Poles in 1920. New York (200,000), Pittsburgh (125,000) or Buffalo (100,000) also hosted considerable Polish contingents (Daniels, 1990; Alexander, 2007) (Figure 6). Early Russian migrants also settled mainly in farm areas in Iowa, Kansas, Minnesota, Nebraska, Oklahoma, Texas and Wisconsin. There was a particularly strong cluster in the Dakotas. Later on most of them flocked to the cities and by 1920 they were the most urban of all foreign groups (Ward, 1987). New York, Chicago, Cleveland or Saint Louis were their main destinations.

Figure 6: Percentage of Polish Migrants in the US, 1900



Source: Own elaboration

National enclaves, institutions and development

The dimension of migration in the late 19th and early 20th century US created, in combination with the path dependence which governed the settlement decisions of migrants, an institutional shock. It led to the emergence of national and ethnic enclaves and urban ghettos.

The armies of European immigrants not only supplied their skills and workforce, they also brought their cultures, beliefs, traditions, habits and social institutions with them. In these national enclaves and urban ghettos the presence of large percentages of migrants from the same origin facilitated reshaping local institutions mirroring those of their places of origin (Rice and Feldman, 1997; Joly, 2000). The habits brought by migrants permeated all aspects of daily life in these areas. It was not only language that was at first preserved.[2] The customs, behaviours and mentality of different European countries and even of specific regions became assimilated into the places where large numbers of migrants from a similar country of origin settled down (Gordon, 1964). "[Immigrants] came not to establish something new but to re-establish something old […]" (Daniels, 1991: 146).

The institutional frameworks developed by the migrants recreated the national constructs they had left at home. Migrants renovated entire local institutions according to their 'national blueprint' often establishing associations, political organizations, press, banks, businesses, and religious groups (Joly, 2000). Hence, any German would have felt almost a home in late 19th or early 20th century Wisconsin; Scandinavians modelled Minnesota in the shape of the countries they had left behind; Irish traditions and beliefs impregnate Boston culture to these days; and parts of New York and northern New Jersey were nothing more than southern Italian cities transferred to America.

The strength and quality of the institutional constructs migrants had left at home, however, varied enormously across national origins. As indicated by Tabellini (2010: 687), Germany, England and northern Italy have more positive cultural indicators which, in turn, determine their better economic trajectory and higher GDP per head. Places like southern Italy, Portugal and southern Spain fare, by contrast, far worse on both counts (Tabellini, 2010: 687). Consequently, whereas many migrants left countries with robust and relatively well-

[2] Although language was perhaps the first factor to fade away.

functioning institutions (e.g. those from England, Germany, the Netherlands or Scandinavia), others came from countries or territories where the institutional environment left a lot to be desired. That was the case of many second wave migrants from Greece, Poland, Russia or the Ukraine. Furthermore, a large percentage of Italian migrants came from southern Italy, a place renowned for a long history of institutional deficiencies. As cultural institutions leave a long-lasting imprint on economic development, importing Scandinavian- or German-type institutional settings, on the one hand, or Italian or Polish institutions, on the other, may have had implications for the ensuing economic trajectory of US counties (e.g. Duranton et al., 2009; Tabellini, 2010).

There is an abundant literature dealing with how institutional factors affect human behaviour and economic outcomes (i.e. North 1981, Putnam, 1993, Acemoglu et al 2001, 2005). According to this literature, institutions – or the social, economic, legal, cultural, or political organization of a society – are increasingly regarded as seminal factors moulding the economic performance of regions or societies (i.e. North, 1981). Prevailing social economic disparities have been ascribed to differences in institutional endowments across territories (i.e. Rodrik et al. 2004) and, local culture, defined as "a dynamic system of rules – explicit and implicit – established by groups in order to ensure their survival, including attitudes, values, beliefs, norms and behaviours […]" (Matsumoto, 2000: 24), is regarded as creating a path dependency over time determining the long-term economic performance of societies (Duranton et al., 2009; Tabellini, 2010).

Consequently, the institutions that migrants from a similar origin settling in a specific part of the US would have developed could have been passed down from generation to generation and/or become embedded in the territory continuing to affect current levels of economic development. There is, however, to the extent of our knowledge, no research which has addressed this question. A good approximation to the question of whether exogenous

institutions developed in history and passed down from generation to generation affect current economic development is provided by the work of Acemoglu et al. (2001). They analyse current cross-country differences in GDP per capita in former colonies in light of the origin of the colonising power, using mortality rates among European settlers as an instrument for current institutions. They argue that European colonial powers tended to stay in places where they faced few health risks and, therefore, had a greater incentive to replicate the solid institutions of their home countries. In these European settlement territories (e.g. Australia or the US), they transferred their habits, customs and traditions as well as their legal, political and economic institutions. In contrast, in colonies where the mortality rates of settlers were high, the incentive for the colonial power was to extract as much resources as possible without necessarily putting too much emphasis and/or interest in the development of European-style institutions. This was, for example, the case of the Spanish or Portuguese colonies in Latin America or of European colonies in sub-Saharan Africa. In their view, the differences in the colonial origin as a source of exogenous variation in institutions have left a long-lasting legacy which exerts a significant influence on a country's income per capita until today (Acemoglu et al., 2001). Hence, being colonised by the British resulted in a different legacy in terms of economic development than a colonisation by the Spanish, Portuguese, Belgians, or even the French. This line of argumentation is supported by La Porta et al. (1998, 1999), who emphasise the importance of the identity of the colonizer for the subsequent development of 'good' institutions. From this perspective, former British colonies had a head start relative to others. Likewise, Véliz (1994), Landes (1998) and North et al. (2000) claim that British colonies overtook former Spanish, Portuguese and French colonies in the Americas because of the superiority of their institutional heritage.

The mass migration of around 27 million Europeans to a country which in 1860 had barely 31 million inhabitants over a period of 60 years and the very distinct settlement patterns of

migrants according to national origin provide an excellent testing ground to analyse whether the different types of institutional constructs brought by the migrants to the places where they settled has made a difference for subsequent economic development. Did the nationality of migrants settling in any given space and the quality of the institutions of their country of origin affect economic development in different parts of the US? Was it better for territories to have been settled by Germans or Scandinavians than by Italians or Poles?

Model and Data

The model

In order to answer whether the different settlement patterns of migrants according to nationality impinged on the economic development of individual US counties, we propose the estimation of the following model:

$$y_{i,t} = \alpha + \beta Nat_{i,t_0} + \gamma X_{i,t-10} + \delta Z_{i,t_0} + \varepsilon_i \qquad (1)$$

where *y* depicts the income per capita of county *i* at time *t* (*t* = 2010); *Nat* is our independent variable of interest and denotes the percentage of inhabitants of county *i* at time t_0 (t_0 = 1880, 1910) originating from a given European country; *X* describes a vector of control variables which are believed to have an effect on the level of development in any given county at time *t-10* (*t-10* = 2000), while *Z* is a vector of control variables representing factors which may have had an impact on the development of county *i* and therefore its attractiveness to migrants at the time of migration (t_0). ε represents the error term. We assume that while the influence of the variables included in vector Z is likely to have disappeared over the more than 100 years covered in the analysis, the variables included in vector X will be key in explaining current differences in GDP per head. A detailed description of the variables is given in Annex II.

The data

Our dependent variable, *y*, consists of the income per capita at county level in 2010. These data are extracted from the data selection of the US Bureau of Economic Analysis (BEA).

The main independent variables of interest – the nationality variables linking the percentage of inhabitants originating from a specific European country to their county of residence in the US – were generated using the database of the IPUMS USA (Integrated Public Use Microdata Series). This data collection contains a large array of American population samples taken from the various US Censuses and American Community Surveys in the period between 1850 and 2010. In order to construct nationality variables the birthplaces of 5,791,531 individuals in 1880 and 923,153 individuals in 1910 were aggregated and allocated to the individuals' county of residence in the US. As the number and geography of counties in the US has been modified between 1880 and the present day (2,875 counties in 1880 and 3,123 in 1910), counties in 1880 and 1910 respectively were matched to counties in 2010 according to their geographical location, taking their current population and density of population into account. The reason for choosing 1880 and 1910 as our points of reference relates to the fact that they represent the peak of foreign presence after each migration wave: 1880 for the first migration wave, 1910 for the second migration wave.

In order to attain more parsimonious results, we focus on ten migrant nationalities. Five of them cover countries or groups of countries which played a particularly important role during the first migration wave of the late 19th century (England, Germany, Ireland, Scandinavia and France). Italians, Poles, Russians (or, more exactly, citizens of the former Soviet Union), Portuguese and Spanish are selected as representatives of the second migration wave.

As mentioned earlier, we introduce two sets of control variables to our model (vectors X and Z). The variables in vectors X and Z are, whenever possible, the same and include different measures of the size of the population, the percentage of black population, the level of education, the percentage of population employed in agriculture, the rate of female participation in the labour force, the rate of unemployment, the infant mortality rate and the average personal income of the respective county *i*.

The rationale for including vector Z is to consider the factors that would, in all probability, have influenced both individual migration decisions, as well as county development levels around the time of the two historical censuses included in the analysis (1880 and 1910). The introduction of variables in vector X (*t-10* dimension, year 2000) reflects the necessity to consider the factors that would determine current levels of development across counties in the US. The controls for the year 2000 are all extracted from the US census database and the US Bureau of Labour Statistics (BLS) and the US BEA. The only exception was the infant mortality data, which stem from the CDC (Centers for Disease Control and Prevention) database. A detailed description of the variables, their source and interdependencies is given in Annex II.

Despite the fact that our two nationality variables depict the situation more than 130 and more than 100 years ago respectively, it is well-known that analyses including migration and GDP per capita are prone to endogeneity. Wealthy territories attract migrants, while migrants contribute to make territories dynamic and rich, and vice versa. We address this potential issue of endogeneity not only by the use of seriously lagged independent variables, but also by resorting to instrumental variable regressions (2SLS). We use religion as our chosen instrument. In particular, we consider the percentage of Baptists and Catholics living in US counties in 1930 as a means to reveal the true underlying effect of migration background onto county development. The reason for choosing religion as an instrument is that it is highly

correlated with migration to the US during the migration boom of the late 19th and early 20th century, without displaying any correlation with the error term. Baptists make up the largest Protestant group in the US and Baptists were among the early settlers in the country, escaping religious persecution in Europe. Yet, Baptists represented a very small contingent in the migration waves considered in the analysis. Catholics, by contrast, formed the majority of migrants entering the US in the second migration wave. The large majority of southern Europeans (Italians, Portuguese and Spaniards) and of Poles were Catholics. We combine these two indicators into a single instrument. The validity of our chosen instrument is confirmed by the performance of the relevant tests. When testing for weak instruments using the Staiger and Stock (1997) tests and the Cragg and Donald minimum eigenvalue statistics and the Stock and Yogo (2002) critical values, our measurement of the presence of Baptists and Catholics in US counties in 1930 – obtained from the ARDA (Association of Religion Data Archives) data collection – is revealed to be a strong instrument of national migration background. The null hypothesis of weak instruments is rejected even if only a 5 percent relative bias is accepted.

Analysis of Results

We first estimate model (1) by means of OLS. Tables 1 and 2 report the results of the OLS estimation focusing on our main independent variable of interest *nationality*. *Nationality* represents the percentage of migrants originating from different European countries living in county *i* in 1880 (Table 1) and 1910 (Table 2) and its possible impact on income per capita in 2010. The nationalities which made the bulk of migrants in the first migration wave (English, German, Irish, Scandinavian and French) are presented in the five left-hand-side columns of

each table, while the coefficients of the second wave migrant origins (Italian, Portuguese, Spanish, Polish and Russian) are reported on the right-hand side columns.

The results of the OLS analysis indicate that, by and large, the connection between the diverse geographical settlement patterns of European migrants from different national origins and current levels of economic development is positive and strongly significant. This relationship holds for the peak years of both migration waves – 1880 (Table 1) and 1910 (Table 2) respectively. Counties where migrants from any given nationality settled in large numbers tend to be significantly richer today than counties which received much lower numbers of population inflows from specific countries, even after other factors which may have spurred migration at the time or may affect current levels of development are controlled for.

This result is consistent regardless of the national origins of the settlers. We find little evidence that the national origin of migrants and their diverse settlement patterns across the US have resulted in statistically different levels of territorial development today. While migrants from different national origins may have been able to instil their culture, habits, institutions, and, to a certain extent, identity into those places where they represented a large percentage of the population, there is no evidence whatsoever of a supposedly 'good' form of institutions associated with, say, German or Scandinavian migrants. Their presence has not led to better economic outcomes than that of what are often considered as 'inferior' institutional constructs linked to, say, large concentrations of Irish, Italian, or Polish settlers. Our results stress that having been settled by Germans or Italians, by Scandinavians or Russians, by French or Portuguese has not made a huge difference for subsequent economic development, as all nationalities – no matter whether first or second wave, or whether they stemmed from the North or the South, the East or the West of Europe – seem to have left a long lasting positive impact on the counties where they settled over 100 years ago.

The only exception[3] is that of counties with the largest percentage of English migrants. The English are precisely the group of migrants who had the least problems in adapting to the new environment. In contrast to the great majority of migrants from other national origins, the barriers to assimilation to the new environment were low for the English immigrant population. There was no new language to learn and the culture, institutions and habits these migrants came across in the US were very similar to the ones they had left at home. Yet, US counties which attracted large numbers of English migrants in 1880 (Table 1) were poorer in 2010 than counties where this was not the case. By 1910, when the number of English living in the US had declined markedly, the relationship ceased to be significant.

Paradoxically, the ease of assimilation of English migrants in the new environment may be at the root of this either negative or insignificant connection. Ease of assimilation would have contributed to dilute the specific traits associated with the migrant character early on and thus undermined the impact left by the presence of large concentrations of English migrants. In contrast, the presence of strong cultural and/or linguistic barriers, as was the case for Germans, Scandinavians, Italians, Polish, Russians, or Portuguese, or of religious divides, as in the case of the Irish, may have contributed to forge and preserve for longer a more differential, migrant-imbued identity in those areas where large numbers of difficult-to-assimilate migrants lived. Hence, rather than current levels of development responding to different institutional formations associated with the type and quality of institutions the migrants left behind in their places of origin, a large influx of migrants may have implanted in US counties receiving large foreign populations a dominant institutional environment related to the character of the migrant. Migration may indeed have also contributed to the formation of different cultural and institutional forms depending on the national origin of migrants. But, regardless of their national origin, migrants were (and are) self-selecting individuals: they

[3] Another smaller exception is the coefficient for Italians in 1880, which is not significant. This may simply reflect that the number of Italians living in the US in 1880 was still very small, as seen in Figure 1 – the peak of Italian migration takes place only after 1900.

tend to be more entrepreneurial, dynamic and willing to take risks than those in similar conditions who decided to stay at home. And this character seems to have prevailed in defining the long-term economic performance of US counties above that of the specific institutions of the countries they left behind.

The coefficients of the two groups of control variables – those which may have affected the settlement pattern of migrants at the time of migration and those for 2000, which are bound to have an important influence on 2010 county levels of GDP per capita – have, in general, the expected results. The connection between most of the 1880 or 1910 controls and county-level economic development 2010 tends to be statistically irrelevant. Neither the mean income in the county, nor the rate of female participation in 1880 is significant. Agricultural employment is barely significant and has a negative sign, signalling that rural counties in 1880 have performed worse than urban ones. The only other factors, besides migration, which are positively and significantly associated with current levels of development, are the literacy rate and the presence of a large black population within the county (Table 1). For the 1910 controls, the presence of a large percentage of black population, female participation and agricultural employment are all insignificant, while literacy levels remain positive and significant (Table 2). The main change with respect to 1880 is that the income variable is now significant and positive.

As expected, the connection between the 2000 controls and GDP per head levels in 2010 is somewhat stronger, but only for a few specific variables. The educational attainment of the population and unemployment display the most significant coefficients. Counties with a better endowment of human capital in 2000 tend to be richer in 2010, whereas higher levels of unemployment affect GDP per head negatively (Tables 2 and 3). The other 2000 control variables – with some exceptions for black population in the 1880 regressions and female participation in the 1910 regressions – tend to be insignificant, implying that neither the

population size of a county, nor infant mortality, or employment in agriculture in 2000 have any correlation with GDP per capita in 2010.

Table 1: OLS-Regression – Immigrants in the United States 1880

	First Immigration Wave					Second Immigration Wave				
VARIABLES	English	German	Irish	Scandinavian	French	Italian	Portuguese	Spanish	Polish	Russian
Nationality_1880	-0.386**	0.867***	1.100***	0.380***	3.261***	0.779	3.993***	16.08***	2.021***	0.632***
	(0.151)	(0.0921)	(0.141)	(0.0576)	(0.727)	(0.842)	(0.908)	(3.781)	(0.560)	(0.181)
total population2000	-0.00267	-0.00169	-0.00136	0.000363	-0.00301	-0.00231	-0.00307	-0.00337	-0.00197	-0.00171
	(0.00247)	(0.00243)	(0.00244)	(0.00248)	(0.00246)	(0.00247)	(0.00247)	(0.00247)	(0.00246)	(0.00247)
black population 2000	-0.000917**	-0.00129***	-0.00117***	-0.00103***	-0.000987***	-0.000966**	-0.000955**	-0.000936**	-0.000971**	-0.000993***
	(0.000378)	(0.000374)	(0.000375)	(0.000375)	(0.000377)	(0.000378)	(0.000377)	(0.000377)	(0.000377)	(0.000377)
college education 2000	0.0140***	0.0142***	0.0137***	0.0139***	0.0140***	0.0139***	0.0138***	0.0139***	0.0140***	0.0139***
	(0.000475)	(0.000468)	(0.000470)	(0.000471)	(0.000473)	(0.000474)	(0.000474)	(0.000473)	(0.000473)	(0.000473)
female participation 2000	0.000399	-0.000713	0.000199	-0.000408	0.000236	0.000360	0.000434	0.000490	0.000228	0.000360
	(0.000631)	(0.000631)	(0.000625)	(0.000636)	(0.000629)	(0.000632)	(0.000630)	(0.000630)	(0.000630)	(0.000630)
unemployment 2000	-0.0218***	-0.0225***	-0.0231***	-0.0223***	-0.0225***	-0.0219***	-0.0225***	-0.0222***	-0.0221***	-0.0215***
	(0.00220)	(0.00217)	(0.00218)	(0.00219)	(0.00220)	(0.00220)	(0.00220)	(0.00220)	(0.00220)	(0.00220)
agriculture 2000	-0.00172	-0.00123	-0.00165	-0.00157	-0.00165	-0.00162	-0.00189	-0.00175	-0.00153	-0.00156
	(0.00153)	(0.00151)	(0.00152)	(0.00152)	(0.00153)	(0.00153)	(0.00153)	(0.00153)	(0.00153)	(0.00153)
infant mortality 2000	-0.000231	-0.000234	-0.000183	-0.000249	-0.000237	-0.000220	-0.000199	-0.000199	-0.000272	-0.000172
	(0.000398)	(0.000393)	(0.000395)	(0.000396)	(0.000397)	(0.000399)	(0.000397)	(0.000398)	(0.000398)	(0.000398)
black population 1880	0.138***	0.176***	0.181***	0.168***	0.146***	0.146***	0.145***	0.146***	0.148***	0.147***
	(0.0344)	(0.0339)	(0.0342)	(0.0342)	(0.0342)	(0.0343)	(0.0342)	(0.0342)	(0.0342)	(0.0342)
income 1880	0.00184	0.00114	-0.00128	0.00119	-0.000171	0.000905	0.000484	0.000477	0.00120	0.00123
	(0.00129)	(0.00123)	(0.00128)	(0.00124)	(0.00128)	(0.00127)	(0.00126)	(0.00126)	(0.00125)	(0.00125)
female participation 1880	0.0141	-0.000491	-0.0618	0.00230	0.0263	0.0138	0.0226	0.0159	0.0122	0.0156
	(0.0560)	(0.0552)	(0.0563)	(0.0557)	(0.0560)	(0.0561)	(0.0559)	(0.0559)	(0.0560)	(0.0560)
agriculture 1880	-0.0692**	-0.0297	0.00693	-0.0638**	-0.0397	-0.0561*	-0.0604*	-0.0478	-0.0576*	-0.0586*
	(0.0317)	(0.0311)	(0.0321)	(0.0312)	(0.0315)	(0.0314)	(0.0313)	(0.0314)	(0.0313)	(0.0313)
literacy 1880	0.248***	0.198***	0.194***	0.247***	0.239***	0.245***	0.243***	0.248***	0.243***	0.243***
	(0.0276)	(0.0276)	(0.0280)	(0.0273)	(0.0275)	(0.0276)	(0.0275)	(0.0275)	(0.0275)	(0.0275)
Observations	2,822	2,822	2,822	2,822	2,822	2,822	2,822	2,822	2,822	2,822
R-squared	0.519	0.533	0.528	0.525	0.521	0.518	0.521	0.521	0.520	0.520

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 2: OLS-Regression: Immigrants in the United States 1910

	First Migration Wave					Second Migration Wave				
VARIABLES	English	German	Irish	Scandinavian	French	Italian	Portuguese	Spanish	Polish	Russian
Nationality_1910	-0.0244	1.030***	1.876***	0.479***	1.911***	0.864***	2.164***	3.947***	3.320**	1.183***
	(0.227)	(0.105)	(0.285)	(0.0646)	(0.654)	(0.166)	(0.536)	(1.305)	(1.543)	(0.114)
total population2000	-0.0116***	-0.0116***	-0.0122***	-0.00797***	-0.0119***	-0.0124***	-0.0121***	-0.0119***	-0.0119***	-0.0104***
	(0.00247)	(0.00243)	(0.00246)	(0.00250)	(0.00247)	(0.00247)	(0.00247)	(0.00247)	(0.00247)	(0.00243)
black population 2000	-2.74e-05	-0.000348	-0.000283	6.05e-07	-2.74e-05	-0.000108	-4.24e-05	-1.95e-06	-6.24e-05	-0.000336
	(0.000417)	(0.000412)	(0.000416)	(0.000413)	(0.000416)	(0.000415)	(0.000416)	(0.000416)	(0.000417)	(0.000411)
college education 2000	0.0145***	0.0147***	0.0142***	0.0143***	0.0144***	0.0144***	0.0144***	0.0145***	0.0145***	0.0142***
	(0.000472)	(0.000462)	(0.000467)	(0.000465)	(0.000469)	(0.000467)	(0.000468)	(0.000468)	(0.000468)	(0.000461)
female participation 2000	0.00131**	0.000149	0.00128**	0.000486	0.00128**	0.00135**	0.00130**	0.00135**	0.00133**	0.00162***
	(0.000599)	(0.000601)	(0.000594)	(0.000603)	(0.000598)	(0.000596)	(0.000597)	(0.000598)	(0.000598)	(0.000589)
unemployment 2000	-0.0247***	-0.0250***	-0.0248***	-0.0261***	-0.0250***	-0.0251***	-0.0251***	-0.0247***	-0.0246***	-0.0234***
	(0.00220)	(0.00216)	(0.00218)	(0.00219)	(0.00220)	(0.00219)	(0.00219)	(0.00219)	(0.00219)	(0.00216)
agriculture 2000	-0.00112	-0.00119	-0.00129	-0.00165	-0.00112	-0.00101	-0.00120	-0.00113	-0.00114	-0.000676
	(0.00150)	(0.00148)	(0.00149)	(0.00149)	(0.00150)	(0.00149)	(0.00150)	(0.00150)	(0.00150)	(0.00148)
infant mortality 2000	-0.000227	-0.000171	-0.000179	-0.000319	-0.000204	-0.000296	-0.000194	-0.000204	-0.000227	-0.000138
	(0.000373)	(0.000367)	(0.000370)	(0.000370)	(0.000372)	(0.000371)	(0.000372)	(0.000372)	(0.000372)	(0.000366)
black population 1910	-0.00743	0.0375	0.0239	0.00778	-0.00264	0.00574	-0.00297	-0.00691	-0.00328	0.0334
	(0.0329)	(0.0326)	(0.0329)	(0.0326)	(0.0328)	(0.0327)	(0.0327)	(0.0327)	(0.0328)	(0.0325)
income 1910	0.00680***	0.00723***	0.00438**	0.00531***	0.00604***	0.00432**	0.00637***	0.00621***	0.00665***	0.00560***
	(0.00186)	(0.00180)	(0.00185)	(0.00183)	(0.00185)	(0.00188)	(0.00183)	(0.00184)	(0.00183)	(0.00180)
female participation 1910	-0.0589	-0.0751*	-0.0823**	-0.0583	-0.0538	-0.0425	-0.0555	-0.0557	-0.0608	-0.0711*
	(0.0392)	(0.0386)	(0.0390)	(0.0388)	(0.0391)	(0.0391)	(0.0390)	(0.0391)	(0.0391)	(0.0385)
agriculture 1910	0.0210	0.0248	0.0336	0.0123	0.00962	0.0245	0.0122	0.0174	0.0233	0.0206
	(0.0336)	(0.0331)	(0.0334)	(0.0333)	(0.0338)	(0.0334)	(0.0336)	(0.0336)	(0.0336)	(0.0330)
literacy 1910	0.154***	0.112***	0.145***	0.147***	0.160***	0.180***	0.155***	0.159***	0.154***	0.174***
	(0.0360)	(0.0357)	(0.0358)	(0.0357)	(0.0360)	(0.0362)	(0.0359)	(0.0360)	(0.0360)	(0.0354)
Observations	3,072	3,072	3,072	3,072	3,072	3,072	3,072	3,072	3,072	3,072
R-squared	0.490	0.506	0.497	0.499	0.492	0.495	0.493	0.492	0.491	0.507

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

In order to assess whether the results of the OLS regressions are affected by problems of endogeneity, we reproduce the regression analysis using an instrumental variable (IV) approach. It can be envisaged that the OLS results may be misleading due to the fact that "immigrants were not distributed evenly but went, in the final analysis, where the jobs or other economic opportunities were and where they perceived economic growth to be the strongest […]" (Daniels, 1991: 126). Consequently, migrants may have been attracted by wealthy areas of the country – precisely those which offered them greatest opportunity. The influx of migrants, in turn, may have further contributed to keep those places ahead of the pack, by generating greater economic dynamism. In order to assess whether these problems of endogeneity cause biased results, we instrument, as indicated earlier, for religion and, in particular, for the presence of Baptists and Catholics in any given US county in 1930. Table 3 and 4 report the results.

The IV analysis underlines the robustness of the results of the OLS analysis. The presence of large numbers of migrants of different national origins in US counties in either 1880 or 1910 has had a long-lasting effect on the levels of economic development of US counties. In virtually all cases, the coefficient of the percentage of migrants of any given nationality living in a particular county in 1880 or 1910 does not change relative to the OLS analysis and remains positive and significant (Tables 3 and 4). Counties which attracted large numbers of migrants at the peak of the big migration waves of the late 19th and early 20th centuries have tended to perform better than counties that did not, regardless of the nationality of the migrants they attracted. It does not matter whether migrants were primarily from Germany, Ireland, Italy or Poland. The mere presence of migrants trumps the potential influence linked to the institutional constructs they may have erected in the US on the basis of those of their countries of origin. What mattered is that they were migrants and that the institutions they contributed to build have left an imprint which has been translated into higher levels of

development in the counties where they settled. The main exception is, once again, related to English migrants in 1880 (Table 3). Their greater capacity to blend into the new environment may have limited their long-term economic legacy. It must be noted, however, that in the 1910 regressions, the coefficient for English migrants is, as in the case of all other nationalities considered in the analysis, positive and significant (Table 4).

In contrast to the OLS results, the IV analysis allows making an interesting distinction between the first and the second wave migration. This is particularly visible in the 1880 estimation results (Table 3). Whereas for the first migration wave countries (the English excluded) the coefficient is always positive and significant, meaning that counties which in 1880 were settled in large numbers by Germans, Irish, Scandinavians, or French have performed well over the last 130 years, the results for the second migration wave countries are considerably weaker. While the concentration of high percentages of Poles, Russians, or Spaniards in a small number of US counties by 1880 may have helped those counties to become more dynamic, for Italians and Portuguese this was not the case. These results may be explained by the fact that the numbers of migrants from second wave countries living in the US in 1880 was so small, that it would only have made a difference for the tiny fraction of counties which had, for whatever reason, high percentages of migrants from any of these origins. By 1910, when migration from Italy, Poland, Russia or Portugal had already taken off, the coefficient is always positive and strongly significant, indicating that the positive trace of migration can only be measured once migrants have reached sufficiently large numbers in order to be able to impose their institutions and/or the migrant character on the territories in which they settled.

The results for the controls, despite some variation in the different country regressions, confirm those reported in the OLS analysis. In general, the regressions in Table 3 stress once again the relevance of education and the impact of a large black population in 1880 for county

levels of GDP per head in 2010. These coefficients remain significant in the 1910 regressions (Table 4). The coefficients for female participation and employment in agriculture in 1880 and 1910 are in most cases insignificant or vary according to the specific country in question. However, perhaps the most interesting change relative to the OLS results is the altered coefficient of county income in 1910. Whereas, on the surface, average county income in 1910 seems to be a strong predictor of income on county level today, once we instrument for religion and therefore correct for possible endogeneity issues, the coefficient becomes in most cases strongly negative and significant. Income in 1910 which could have strongly influenced the original decision to migrate to any given county is not a good predictor of income 100 years later (Table 4). Rich counties then are not necessarily rich today, once migration and other factors are controlled for. Migration, in general, and the concentration of migrants of specific national origins in particular counties, by contrast, has generated a strong path dependency which is a good predictor of current levels of GDP per head.

The coefficients for the 2000 control variables vary according to the nationality considered. In particular, the significance level of the county's population size and of female participation rate changes according to the nationality in question. In any case the main results of the OLS analysis are confirmed. Unemployment and education in 2000 – with different signs – remain the main predictors of county GDP per head in 2010 (Tables 3 and 4).

Table 3: IV-Regression: Immigrants in the United States 1880

	First Migration Wave					Second Migration Wave				
VARIABLES	English	German	Irish	Scandinavian	French	Italian	Portuguese	Spanish	Polish	Russian
Nationality_1880	1.003	3.791***	5.037***	2.840***	42.57***	324.5	531.6	477.1***	41.90***	27.65***
	(1.686)	(0.436)	(0.712)	(0.642)	(6.492)	(242.4)	(691.9)	(126.1)	(7.491)	(7.990)
total population2000	-0.000922	-1.47e-05	0.00157	0.0169***	-0.0130***	-0.0550	-0.120	-0.0373***	0.00234	0.0187**
	(0.00327)	(0.00284)	(0.00280)	(0.00533)	(0.00387)	(0.0434)	(0.156)	(0.0111)	(0.00419)	(0.00949)
black population 2000	-0.00107**	-0.00239***	-0.00194***	-0.00145***	-0.00130**	-0.00336	-0.000213	-0.000248	-0.00118*	-0.00239**
	(0.000426)	(0.000463)	(0.000444)	(0.000493)	(0.000540)	(0.00329)	(0.00425)	(0.000962)	(0.000632)	(0.00120)
college education 2000	0.0137***	0.0151***	0.0128***	0.0137***	0.0146***	0.0106**	-0.00747	0.0135***	0.0142***	0.0125***
	(0.000613)	(0.000560)	(0.000554)	(0.000605)	(0.000682)	(0.00425)	(0.0283)	(0.00119)	(0.000793)	(0.00146)
female participation 2000	0.000162	-0.00424***	-0.000279	-0.00521***	-0.000929	0.0116	0.0138	0.00498**	-0.00184	0.00153
	(0.000700)	(0.000894)	(0.000710)	(0.00149)	(0.000917)	(0.00960)	(0.0188)	(0.00200)	(0.00112)	(0.00191)
unemployment 2000	-0.0222***	-0.0245***	-0.0276***	-0.0248***	-0.0299***	-0.0206	-0.0974	-0.0309***	-0.0259***	-0.00650
	(0.00228)	(0.00254)	(0.00259)	(0.00288)	(0.00336)	(0.0161)	(0.101)	(0.00598)	(0.00374)	(0.00791)
agriculture 2000	-0.00133	7.00e-05	-0.00177	-0.00125	-0.00211	-0.00293	-0.0387	-0.00556	0.000182	0.000884
	(0.00162)	(0.00176)	(0.00171)	(0.00195)	(0.00218)	(0.0112)	(0.0510)	(0.00396)	(0.00257)	(0.00460)
infant mortality 2000	-0.000195	-0.000277	-4.67e-05	-0.000430	-0.000431	0.000390	0.00273	0.000429	-0.00128*	0.00195
	(0.000406)	(0.000457)	(0.000446)	(0.000509)	(0.000567)	(0.00295)	(0.00581)	(0.00101)	(0.000691)	(0.00134)
black population 1880	0.164***	0.280***	0.310***	0.312***	0.152***	0.315	0.0893	0.154*	0.207***	0.227**
	(0.0470)	(0.0422)	(0.0448)	(0.0576)	(0.0487)	(0.281)	(0.382)	(0.0856)	(0.0583)	(0.105)
income 1880	-0.000863	0.00134	-0.00975***	0.00185	-0.0153***	-0.0744	-0.0791	-0.0170***	0.00355*	0.00725*
	(0.00351)	(0.00144)	(0.00207)	(0.00160)	(0.00307)	(0.0571)	(0.105)	(0.00571)	(0.00214)	(0.00413)
female participation 1880	0.00897	-0.0449	-0.328***	-0.0649	0.191**	0.493	1.340	0.110	0.00337	0.139
	(0.0571)	(0.0646)	(0.0790)	(0.0734)	(0.0842)	(0.544)	(1.833)	(0.142)	(0.0935)	(0.170)
agriculture 1880	-0.0272	0.0641*	0.238***	-0.104**	0.175***	0.512	-0.437	0.230**	-0.0589	-0.104
	(0.0601)	(0.0386)	(0.0545)	(0.0413)	(0.0570)	(0.483)	(0.601)	(0.109)	(0.0524)	(0.0942)
literacy 1880	0.235***	0.0396	0.0129	0.259***	0.175***	0.346	0.0473	0.333***	0.201***	0.167**
	(0.0321)	(0.0394)	(0.0449)	(0.0352)	(0.0406)	(0.215)	(0.396)	(0.0726)	(0.0466)	(0.0848)
Observations	2,822	2,822	2,822	2,822	2,822	2,822	2,822	2,822	2,822	2,822
R-squared	0.504	0.365	0.396	0.217	0.022					

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4: IV-Regression: Immigrants in the United States in 1910

	First Migration Wave					Second Migration Wave				
VARIABLES	English	German	Irish	Scandinavian	French	Italian	Portuguese	Spanish	Polish	Russian
Nationality_1910	36.22***	5.116***	15.41***	5.748***	144.3***	22.23***	121.0***	412.0**	264.0***	11.14***
	(10.65)	(0.492)	(2.091)	(1.045)	(46.88)	(4.543)	(45.53)	(187.7)	(71.54)	(1.668)
total population2000	0.0101	-0.0116***	-0.0172***	0.0315***	-0.0367***	-0.0329***	-0.0426***	-0.0513**	-0.0387***	-0.000491
	(0.00987)	(0.00297)	(0.00332)	(0.00896)	(0.0129)	(0.00761)	(0.0155)	(0.0230)	(0.0108)	(0.00481)
black population 2000	-0.00158	-0.00162***	-0.00212***	0.000319	5.04e-05	-0.00207*	-0.000808	0.00273	-0.00273*	-0.00292***
	(0.00135)	(0.000524)	(0.000614)	(0.000737)	(0.00169)	(0.00113)	(0.00174)	(0.00270)	(0.00152)	(0.000877)
college education 2000	0.00592**	0.0157***	0.0123***	0.0128***	0.00821***	0.0112***	0.0108***	0.0148***	0.0157***	0.0123***
	(0.00290)	(0.000577)	(0.000678)	(0.000880)	(0.00278)	(0.00136)	(0.00237)	(0.00269)	(0.00154)	(0.000920)
female participation 2000	-0.00325	-0.00446***	0.00108	-0.00858***	-0.00119	0.00233	0.000891	0.00583	0.00309	0.00419***
	(0.00227)	(0.000909)	(0.000782)	(0.00208)	(0.00255)	(0.00152)	(0.00246)	(0.00399)	(0.00198)	(0.00118)
unemployment 2000	-0.0378***	-0.0263***	-0.0256***	-0.0422***	-0.0465***	-0.0353***	-0.0490***	-0.0301**	-0.0187***	-0.0123***
	(0.00773)	(0.00265)	(0.00287)	(0.00501)	(0.0114)	(0.00594)	(0.0128)	(0.0128)	(0.00722)	(0.00442)
agriculture 2000	-0.00205	-0.00147	-0.00252	-0.00751***	-0.00120	0.00182	-0.00564	-0.00250	-0.00235	0.00307
	(0.00459)	(0.00181)	(0.00197)	(0.00289)	(0.00608)	(0.00383)	(0.00640)	(0.00861)	(0.00482)	(0.00282)
infant mortality 2000	0.000704	4.95e-05	0.000160	-0.00134*	0.00142	-0.00202**	0.00159	0.00214	-0.000307	0.000606
	(0.00117)	(0.000449)	(0.000490)	(0.000687)	(0.00160)	(0.00101)	(0.00168)	(0.00239)	(0.00119)	(0.000693)
black population 1910	0.300**	0.215***	0.248***	0.173***	0.338*	0.326***	0.230	0.0249	0.306**	0.375***
	(0.135)	(0.0448)	(0.0551)	(0.0664)	(0.174)	(0.107)	(0.162)	(0.188)	(0.135)	(0.0829)
income 1910	-0.0444***	0.00908***	-0.0129***	-0.0107**	-0.0480**	-0.0561***	-0.0152	-0.0510*	-0.00289	-0.00425
	(0.0160)	(0.00221)	(0.00358)	(0.00452)	(0.0193)	(0.0137)	(0.0112)	(0.0283)	(0.00642)	(0.00373)
female participation 1910	0.145	-0.140***	-0.252***	-0.0531	0.313	0.357***	0.119	0.261	-0.221*	-0.176**
	(0.134)	(0.0477)	(0.0575)	(0.0690)	(0.199)	(0.130)	(0.174)	(0.267)	(0.133)	(0.0737)
agriculture 1910	0.277**	0.0392	0.123***	-0.0851	-0.848***	0.109	-0.477**	-0.370	0.194*	0.0158
	(0.127)	(0.0404)	(0.0460)	(0.0622)	(0.313)	(0.0865)	(0.233)	(0.262)	(0.117)	(0.0614)
literacy 1910	0.229**	-0.0570	0.0796*	0.0705	0.601***	0.809***	0.201	0.660**	0.157	0.339***
	(0.112)	(0.0478)	(0.0481)	(0.0652)	(0.206)	(0.162)	(0.149)	(0.309)	(0.115)	(0.0714)
Observations	3,072	3,072	3,072	3,072	3,072	3,072	3,072	3,072	3,072	3,072
R-squared		0.259	0.126							

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Conclusions

The approximately 27 million Europeans (Ward, 1987) who made the long voyage across the Atlantic in the late 19th and early 20th century did not arrive in the US with an innate capacity to immediately blend into the new environment. They came with baggage. A baggage determined by the language, culture, traditions, institutions and identity they left at the other side of the ocean. However, the path dependency of migration trends allowed them to preserve at least part of that baggage. The geography of migration in the US became cut along lines related national and/or ethnic origins. Newly arrived immigrants flocked to places where relatives, friends and co-nationals had already settled, leading to the formation of urban ghettos and of rural enclaves along national and/or ethnic lines. In these urban ghettos and rural enclaves migrants were relatively free to import their traditions, habits, customs, language and institutions, leaving an identity trace which, in some cases, remains to this day. As Daniels (1991) underlines immigrants, rather than blending in, re-established their home traditions in their new American environment.

However, the culture, traditions and institutions imported by migrants were far from uniform. They varied significantly from one European country to another. There is firm evidence that these institutions, linked to local cultures and identities, have shaped and continue to shape the economic trajectories of nations in Europe (e.g. Tabellini, 2010). Hence, it could be safely assumed that US counties settled by migrants coming from areas of Europe with more efficient institutions (say England, Germany or Scandinavia) would have tended to function better and, consequently, to have superior economic trajectories than those counties where the majority of settlers came from European countries traditionally regarded as having weaker or less efficient institutions (e.g. Italy, Poland, Russia or Portugal). This paper has addressed precisely this question; has the institutional heritage left behind by Scandinavian and German migrants led to better economic outcomes than that of Italians, Poles or Russians?

The results of the analysis stress that this is far from being the case. Counties which received a large intake of migrants in the late 19th and early 20th centuries have tended to perform significantly better than those which were bypassed by the great migration waves. Attracting large numbers of immigrants has been a key factor behind the economic development of US counties over the long-term. The national origin of migrants, however, did not matter. Counties where Germans, Irish, Scandinavians or French settled in large numbers have performed equally well than those which attracted masses of Italians, Portuguese, Spanish, Poles or Russians, and vice versa. In spite of stereotypes, an Italian or a Polish heritage has resulted in similar levels of economic growth as a Scandinavian heritage. And leaving some sort of European imprint has been more favourable for economic development than not managing to attract large numbers of migrants during the big migration waves in the first place.

The only exception is related to English migration, precisely the group which a priori should have had the greatest ease in settling and prospering in the new environment. However, the counties receiving the largest influx of English migrants have performed decidedly worse than those having large contingents of migrants from origins which faced higher barriers to adaptation.

It may therefore be argued that the economic imprint of mass migration into the US is not linked to the national origin of the migrants arriving into the country but to two specific factors: the distinct and self-selective character of the migrant and how barriers to assimilation helped preserve this character over time. Regardless of whether they came from Sweden, Poland, Italy or France, migrants tended to be more entrepreneurial and economic, more willing to take risks, and more eager to succeed than the folk they left behind in their places of origin. This would have generated a dynamic character associated to the traits of the migrant in places where migrants from any particular nationality settled in large numbers. Such

dynamic migrant institutional character would have superseded any of the potential advantages or flaws linked to the institutional constructs inherited from their places of origin. Greater barriers to integration would have helped preserve these traits for longer and to generate an institutional environment which favoured continued economic development in the long-run. Easy assimilation, by contrast, would have diluted the migrant character much earlier, helping to explain the relatively poor performance of those counties which were mainly settled by English migrants.

Although it is always difficult to establish parallels across historical time periods, the results of the analysis raise interesting questions about the dilemmas linked to current migration challenges in the US and elsewhere in the developed world. More than a century ago hordes of young, poor, uneducated, men with relatively little skills made a perilous voyage across the sea into an unknown territory in the hope of finding a better life. Upon arrival they settled where their country folk had set up house, leading to the formation of extensive urban ghettos and of rural enclaves. Huddling together, in combination with limited work and language skills and some discrimination by the locals, hindered their process of assimilation. Yet, all these conditions which are often regarded as negative from an economic point of view, have been at the very heart of the economic dynamism of the parts of the US were these migrants settled for more than a century. Lack of migration and/or an ease of assimilation of migrants may have created fewer problems in the short-term, but has had serious negative economic consequences in the long-run. These results provide some interesting food for thought about the long-term economic implications of current migration policies for those countries which have been more vocal in curbing it. Overall, the results emphasise the need to look at migration from a perspective that not only takes the short-term into consideration. Otherwise, it may not be just migrants who risk missing the boat, but the countries and territories that may have benefited from hosting them.

Acknowledgements

We are grateful to Alejandra Castrodad-Rodríguez for her help in building the maps. This paper has benefited from the comments and suggestions by participants at conferences in Kiel, Bratislava and Los Angeles. The financial support of the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement nº 269868 is gratefully acknowledged.

Reference list

Acemoglu, D., S. Johnson and J. A. Robinson (2001): The Colonial Origins of Comparative Development: An Empirical Investigation, *The American Economic Review*, Vol. 91, No. 5, pp. 1369-1401.

Acemoglu, D. and S. Johnson (2005): Unbundling Institutions. *Journal of Political Economy*, Vol. 113, No. 5, pp. 949-995

Alexander, J.G. (2007): *Daily Life in Immigrant America: 1870-1920*, Greenwood Press, London.

Bertocchi, G. and C. Strozzi (2006): The Age of Mass Migration: Economic and Institutional Determinants, Discussion paper 2499, IZA Institute for the Study of Labour, Bonn.

Bodnar, J. (1992): Remaking America: Public Memory, Commemoration and Patriotism in the Twentieth Century, Princeton University Press, Princeton, N.J.

Carlsson, S. (1976): Chronology and composition of Swedish emigration to America, in H. Runblom and H. Norman (eds.): *From Sweden to America*, pp. 114-148, University of Minnesota Press, Minneapolis.

Daniels, R. (1990): Coming to America. A History of Immigration and Ethnicity in American Life, HarperPerennial, New York.

Dunlevy, J. and H. Gemery (1977): The Role of Migrant Stock and Lagged Migration in the Settlement Patterns of Nineteenth Century Immigrants, *The Review of Economics and Statistics*, Vol. 59, No. 2, pp. 137-144.

Duranton, G., A. Rodríguez-Pose and R. Sandall (2009): Family structures and the persistence of regional disparities in Europe, *Economic Geography*, Vol. 85, No. 1, pp. 23-47.

Erickson, C. (1972): Who were the English and Scottish emigrants in the 1880s?, in D.V. Glass and R. Revelle (eds.): *Population and Social Change*, Arnold, London.

Fitzpatrick, D. (1983): Irish farming families before the First World War, *Comparative Studies in Society and History*, Vol. 25, pp. 339-374.

Gibson, C. and K. Jung (2006): Historical Census Statistics on the Foreign-Born Population of the United States: 1850-2000, Working Paper no. 81, Population Division, US Bureau of the Census, Washington D.C.

Gordon, M. (1964): Assimilation and American Life, New York: Oxford University Press.

Hatton, T.J. and J.G. Williamson (1994): What drove the mass migrations from Europe in the late nineteenth century?, *Population and Development Review*, Vol. 20, pp. 533-559.

Joly, D. (2000): Some structural effects of migration on receiving and sending countries, *International Migration*, Vol. 38, No. 5, pp. 25-40.

Kamphoefner, W. (1976): At the crossroads of economic development: Background factors affecting emigration from nineteenth century Germany, in I. A. Glazier and L. de Rosa (eds.): *Migration across Time and Nations*, pp. 174-201, Holmes and Meier, New York.

Landes, D. (1998): The wealth and poverty of nations. Why some are so rich and some so poor, W.W. Norton & Co, New York.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R.W. Vishny (1998): Law and Finance, *Journal of Political Economy*, Vol. 106, No. 6, pp. 1113-55.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R. Vishny (1999): The Quality of Government, *Journal of Law, Economics, and Organization*, Vol. 15, No. 1, pp. 222-79.

Levy, M.B. and W.J. Wadycki (1973): The Influence of Family and Friends on Geographic Labor Mobility: An International Comparison, *This Review*, pp. 198 – 203.

Matsumoto, D. (2000): Culture and Psychology, 2nd ed., Brooks Cole, Pacific Grove.

North, D. (1981): Structure and Change in Economic History, Norton, New York, Norton.

North, D., W. Sommerhill and B. Weingast (2000): Order, Disorder and Economic Change: Latin America vs. North America, in: B. Bueno de Mesquita and H. Root (eds.): *Governing for Prosperity*, pp.17-58, Yale University Press.

Putnam, R. (1993): Making democracy work. Princeton University Press, Princeton, N.J.

Rice, T. and J. Feldman (1997): Civic Culture and Democracy from Europe to America, *Journal of Politics*, Vol. 59, pp. 1143-72.

Rodrik, D., A. Subramanian and F. Trebbi (2004): Institutions rule: The Primacy of Institutions over Geography and Integration in Economic Development, *Journal of Economic Growth*, Vol. 9, pp. 131-165.

Staiger, D. and J.H. Stock (1997): Instrumental variables regression with weak instruments, *Econometrica*, Vol. 65, No. 3, pp. 557-86.

Stock, J.H. and M. Yogo (2002): Testing for Weak Instruments in Linear IV Regression. National Bureau of Economic Research, *NBER Technical Working Paper*, No. 0284.

Tabellini, G. (2010): Culture and Institutions: Economic Development in the Regions of Europe, *Journal of the European Economic Association*, Vol. 8, No. 4, pp. 677-716.

Vedder, R., and L. Gallaway (1972): The Geographical Distribution of British and Irish Emigrants to the United States After 1800, *Scottish Journal of Political Economy*, Vol. 19, pp. 19-35.

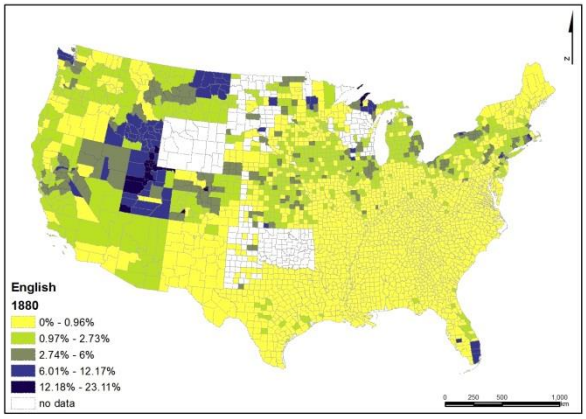
Véliz, C. (1994): The New World of the Gothic Fox: Culture and Economy in English and Spanish America, University of California Press, Berkeley.

Ward, D. (1987): Population growth, migration and urbanization, 1860–1920, in: R.D. Mitchell and P.A. Groves, eds. *North America: The Historical Geography of a Changing Continent*, pp. 299–320. Rowman and Littlefield, Totowa, New Jersey, USA.

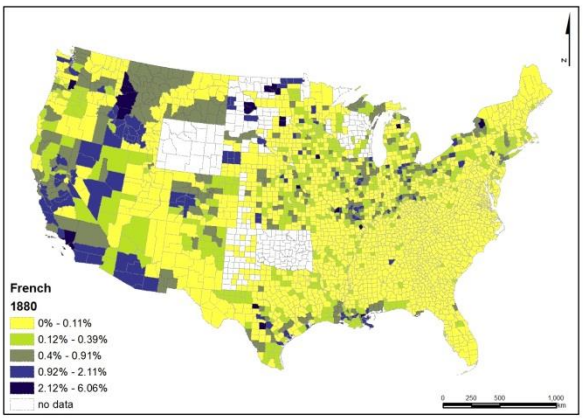
Annex I

Percentage of migrants from diverse of European countries at their peak of migration

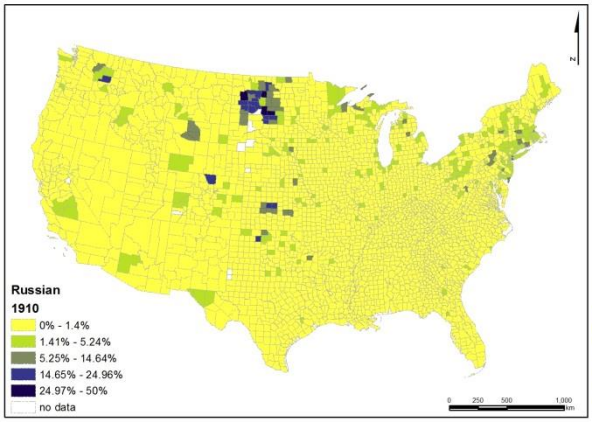
England (1880)



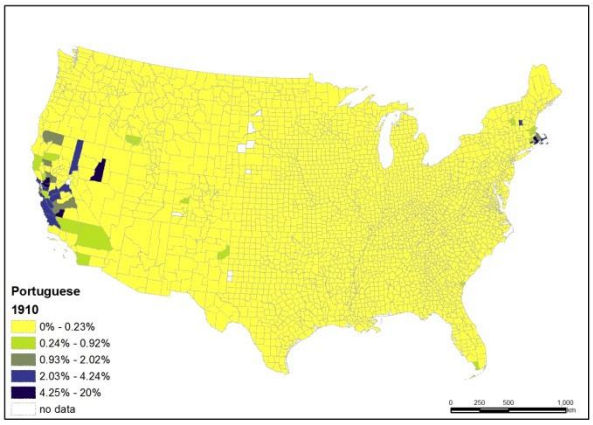
France (1880)



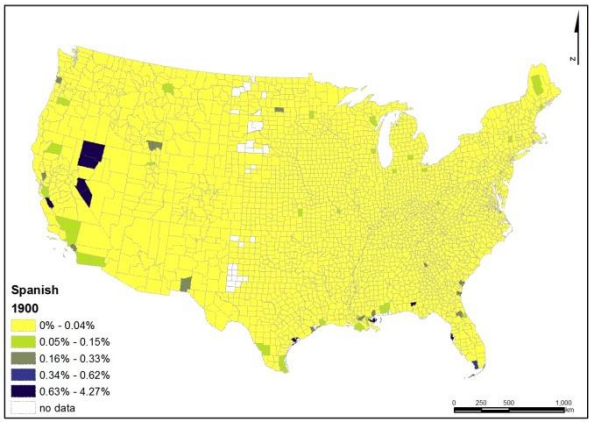
Russia (1910)



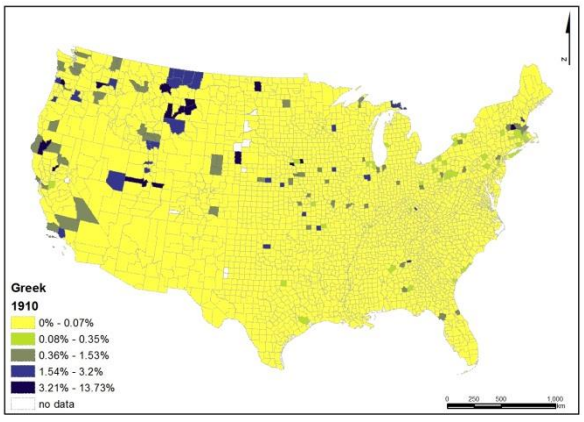
Portugal (1910)



Spain (1900)



Greece (1910)



Annex II

Variables used in the analysis and their sources

Variable	Description	Source
Dependent Variable *y*		
Income	Income per capita in county *i* in year 2010	US BEA
Main independent variables:		
*Nationality**	percentage of inhabitants of county *i* originating from country *j* in year 1880, 1910 respectively	IPUMS USA
Instruments:		
Catholics & Baptists	percentage of Catholics and Baptists in population of county *i*	ARDA
Controls:		
*population**	Log of total population in county *i* in year *	ICPSR , US census
*black population**	Percentage of black population of county *i* in year *	ICPSR, US census
*education**	Percentage of population of county *i* with college degree in year *	US census
*female participation**	Female participation rate in the labour force in county *i* in year *	ICPSR , US census
*unemployment**	unemployment rate in county *i* in year *	US census and US BLS
*agriculture**	percentage of the labour force employed in agriculture in county *i* in year *	US census
*infant mortality**	infant mortality rate in county *i* in year *	CDC
*income**	mean income of population in county *i* in year *	US BEA
*literacy**	literacy rate in county *i* in year *	ICPSR

* indicates a certain year from 1880 to 2000