ABSTRACT

This chapter focuses on the economic analysis of what has been called the age of mass migration, 1850 to 1913, and its aftermath up to 1940. This has captured the interest of generations of economic historians and is still a highly active area of research. Here we concentrate on migration from Europe to the New World as this is where the bulk of the literature lies. We provide an overview of this literature focusing on key topics: the determinants of migration, the development of immigration policy, immigrant selection and assimilation, and the economic effects of mass migration as well as its legacy through to the present day. We explain how what were once orthodoxies have been revisited and revised, and how changes in our understanding have been influenced by advances in methodology, which in turn have been made possible by the availability of new and more comprehensive data. Despite these advances some issues remain contested or unresolved and, true to cliometric tradition, we conclude by calling for more research.

Keywords: Mass migration; the Atlantic economy; immigrants and emigrants.

JEL Codes: N31, N32, J61, F22.

Corresponding author:
Timothy J. Hatton
Department of Economics
University of Essex
Colchester CO4 3SQ, UK
Email: hatton@essex.ac.uk
Introduction

This paper provides a survey of cliometric research on international migration. We focus on studies of migration from Europe during the age of mass migration from the middle of the nineteenth century up to 1913, and during the interwar de-globalisation that followed. While there were other important migrations within Europe and Asia during this period, we focus on what might be called the “greater Atlantic economy” and especially on migration to the United States. This transformed the world economy and it is where the vast bulk of the cliometric literature has focused attention. The literature has evolved over recent years with old interpretations being challenged, as new insights have emerged, although some issues are still contested. The application of new approaches has strong parallels with the research of economists on more recent periods and this has been underpinned by the emergence of new and more comprehensive data.

We start with an overview of the determinants of aggregate migration streams, a well-established literature that has been less active in recent years. This includes time series analysis of the push-pull variety as well as studies that focus on cross sectional differences in emigration intensity between origin countries and regions. Migration flows were shaped by immigration policies that became increasingly restrictive in the early twentieth century. We review the literature on the formation of immigration policy, focusing on the political economy determinants. We then turn to the question of whether international migrants were positively or negatively selected from source country populations. This micro-level analysis, based on unit-record data, has been the focus of much of the recent literature and has challenged the conventional wisdom that migrants were positively selected. The self-selection of migrants has implications for the old debate about immigrant assimilation, and we review recent research that marks a return to the more pessimistic view of the older literature. We then turn to the economic effects of the immigrants in New World destinations as well as the effects of emigration on the Old World. The traditional concerns about the effects on wages and living standards have been examined using a variety of methods that yield a range of different results. Finally we briefly review a recent literature that seeks to explain economic outcomes in the present as the very long run consequences of the age of mass migration.

Determinants of international migration

From 1850 to 1940 an estimated 50 million migrants travelled from Europe to the New World (Figure 1), with about three fifths going to the United States and smaller streams heading for South America and the British Dominions. There is a long standing interest in what determined the year-to-year movements in migration, a literature that goes back to work by Jerome (1926) and Thomas (1941). This literature sought to explain cycles in transatlantic migration using time series models that focused on variations in economic conditions in source and destination countries. The “push” versus “pull” literature that flourished in the 1960s and 1970s was critically reviewed by Gould (1979). As he pointed out, when cyclical indicators such as industrial production are included for both source and destination countries, the latter typically
dominate in terms of the size and significance of the coefficients. On the whole, these studies suggested that short-term ups and downs in aggregate migration were dominated by pull rather than by push factors. And in their presence other variables representing the incentive to migrate, such as income or wages at home and abroad, often failed to achieve significance. Gould (1979, p. 668) concluded that “By and large, reaction to this literature is one of some disappointment, for not only has it failed to generate important new insights … it has had only limited success in confirming or denying old interpretations”.

It is difficult to believe that emigration decisions were determined exclusively by conditions in the destination or that they did not depend on the prospects of higher wages or incomes. Emigration must have been determined by some assessment of the expected gain and hence by some comparison, however approximate, between conditions at home and abroad. But the first generation of studies lacked a coherent underlying model of the emigration decision, which made the coefficients difficult to interpret. Using a basic economic framework, Hatton (1995) developed a model in which representative potential migrants base their decisions on the comparison of future expected incomes at home and abroad. In the manner of Todaro (1969) expected income in a location depends on both the wage rate and the probability of employment. The latter accounts for the effects of the business cycle and, if migrants are risk-averse and there is greater uncertainty about conditions abroad than at home, that would help to explain the dominance of cyclical conditions, especially at the destination. In addition, the time series dynamics that were often included to soak up serial correlation can be accounted for by adaptive expectation formation.

A number of studies have used this framework, or can be interpreted within it. Estimates for emigration from the UK from 1870 to 1913 show that the wage gap and employment rates at home and abroad all matter in a manner consistent with the model (Hatton, 1995). While short run fluctuations are largely accounted for by the business cycle, longer term trends can be accounted for by the slowly changing income gaps between origin and destination countries. For example, the 17 percent fall in the ratio of US to Irish wage rates between 1876-1880 and 1909-13 accounted for a long-run decline in emigration from Ireland of 4 per thousand of the population (Hatton and Williamson, 1998, p. 83). Another important variable that features in these models is the stock of previous migrants at the destination. This strong pull effect of emigrant networks or diasporas is often interpreted as reducing the costs and risks for new migrants. In the Irish case the declining ratio of emigrant stock to home population reduced emigration by another 4 per thousand. New migrants were often assisted through remittances or pre-paid tickets sent by previous emigrants. Consistent with this, Magee and Thompson (2006) find that the flow of money orders to the UK from 1880 to 1913 depended on both the migrant stock and the average income of the English-speaking destination countries.

The direct costs of passage should also matter but these have been absent from most econometric studies for lack of suitable data. The cost of sea voyage fell after 1850 with the shift from sail to steam and the effective price fell even more steeply with shorter passage times (Keeling, 1999; Cohn, 2005; Sánchez-Alonso, 2007), not to mention improvements in overland travel to and from seaports. On the North Atlantic routes sharp swings in ticket prices were
associated with changes in the effectiveness of shipping cartels. Using quarterly data for 1899-1913, Deltas et al. (2008) find that passenger numbers on routes to the US and Canada were 22 percent lower at times when cartels were keeping prices up. Other elements of costs also mattered. Assisted passages to Australia substantially reduced the costs for UK emigrants in 1911-13, leading to a surge of emigration (Pope, 1981). For Spanish emigrants, the depreciation of the Peseta between 1882 and 1905 increased the travel costs of emigration from Spain, which had the opposite effect. Sánchez-Alonso (2000a) finds that this reduced emigration during those years by as much as 30 percent.

Time series models account fairly well for the ups and downs of well-established migration streams, and perhaps this is one reason that they have not been the focus of much of the recent literature. A greater challenge is to account for the initial rise of emigration that got these streams started and the differences in emigration intensity between origin countries and localities. Among western European countries annual average gross emigration rates varied widely from 12 per thousand for Ireland to less than one per thousand for France. The trends also differed. The Irish emigration rate declined from the 1860s, and those of Germany, Norway and Sweden fell from the 1880s. From that time emigration from Italy and Spain and from some countries in Eastern Europe increased steeply up to the outbreak of war in 1914. One stylized fact is that, during the onset of modern economic growth, emigration rates often increased from low levels, reaching a peak after some decades and then gently declining. This long swing in migration, though often not complete, has been identified for at least some European countries (Akerman, 1976; Hatton and Williamson, 1998 Ch. 3).

Several studies suggest that demographic trends were important. The present value of migration would be greatest for those with the longest time horizons and the lowest costs: the young and single. Ellis Island records show that the average age of immigrant arrivals to the US in 1900-10 was 26, two thirds were male, and the majority were single (Bandiera et al. 2013). Using decade average emigration rates for 12 European countries, Hatton and Williamson (1998, Ch. 3) found that the birth rate lagged two decades had a large positive effect. Annual time series for Scandinavian countries also supports this view (Quigley, 1972; Larsen 1982; Hatton and Williamson 1998, Ch. 4). Interestingly, Greenwood (2007) finds that a higher current birth rate tended to reduce the emigration rate for those of parenting age, probably reflecting the higher costs of family migration. However, demographic effects were weaker for countries such as Italy and Spain (Sánchez Alonso, 2000b; Hatton and Williamson 1998, p. 113) because other constraints mattered more.

Emigration often gathered momentum just when the pace of development was quickening at home. In the early stages most migrants were too poor to migrate, even though the incentive was large, but economic development served to ease the “poverty constraint”. Thus on the upside of the long swing emigration grew as wages at home increased, while on the downside further increases in the home wage reduced the incentive to emigrate. Using annual time series from the 1870s to 1913, Faini and Venturini (1994) find for Italy, and Sánchez-Alonso (2000b) finds for Spain, that rising income per capita at home positively influenced emigration. The evidence also suggests that poverty constraints became less binding as the stock of previous
emigrants, who could provide assistance, increased (Hatton and Williamson 2005, p. 65). This interaction helps to explain why emigration from Ireland (with its large post-famine emigrant stock) declined while emigration from equally poor Italy (with a low initial stock) increased as development gathered pace. Bohlin and Eurenius (2010) find a positive effect of the interaction between poverty and the emigrant stock in a panel of emigration from Swedish counties 1881-2010. Poverty constraints are also observable at the micro-level. In the Hesse Cassel region of Germany from 1832 to 1857 Wegge (1998) found that networked emigrants carried less cash with them. By the 1920s these effects were weaker but Armstrong and Lewis (2017) provide evidence that the need to save served to delay emigration to Canada.

There were many political and economic events that led immigrants to seek refuge during the age of mass migration, including the Great Irish Famine in the late 1840s, failed political revolutions in 1848, the Swedish and Finnish famine of 1866 to 1868, the Mexican Revolution of the 1910s, and the religious persecution of Jews in the 1880s, the early 1900s, and the interwar period. Mokyr and Ó Gráda (1982) and Cohn (1995) examine immigrant flows during the Great Famine, and Boustan (2007) and Spitzer (2014) study Jewish emigration. While short-term shocks boosted emigration in the short-term and then subsided, their effects often persisted through the powerful effect of the stock of previous migrants on subsequent flows.

Less attention has been paid to how potential migrants chose among alternative destinations in the New World. Differences between destinations in the pre-existing migrant stock, colonial ties, common language and cultural affinity meant that they were often poor substitutes (Taylor, 1994). But there was greater scope for choice among English speaking countries, as between the US and Canada (Green et al., 2002) and also within South America. Balderas and Greenwood (2010) examine the effect of emigration to one destination on emigration to other destinations using annual time series for emigration from 12 source countries to three destinations. Using instrumental variables they find evidence of substitution between Argentina and Brazil, but not between either of these and the US. The political climate in host countries could also matter. Bertocchi and Strozzi (2008) examine the effects of political institutions on migration between 11 European countries and three New World Countries (Argentina, Canada and the US) on decade-average data for 1870-1910. In addition to economic and demographic variables they find that immigration was positively related to political participation (democracy and suffrage) as well as to rights for immigrants (access to citizenship, land and education).

Many migrants returned, often after just a few years. The conventional estimates indicate that the ratio of out-migrants to in-migrants was about 40 percent (Gould, 1980; Kuznets and Rubin, 1954). However, recent estimates by Bandiera et al. (2013) put the figure much higher. Using the Ellis Island records they calculate that gross immigration from 1900 to 1910 was about 20 percent higher than the official figures report. Comparing this inflow with the change in the stock of immigrants from the census implies that outflows were 60 percent of inflows, as compared with the previously accepted figure of 40 percent. The same calculation for the following decade, when the immigrant stock hardly changed, yields an outflow-inflow ratio of
75-80 percent. Some of these migrants may have been short-term visitors but, even so, high return rates call into question the treatment of migration as a once-and-for-all decision, suggesting instead a significant share of circular migration. Indeed, Ward (2016) shows that between 1897 and 1914 repeat entrants (foreign-born who had previously been in the US) accounted for 10-20 percent of all foreign-born arrivals. These were not mainly the golondrinas or seasonal workers from Southern Europe so often referred to in the literature: repeat entrants were instead often from Northern and Western Europe and were higher skilled, suggesting that the circular flow was not dominated by low-skilled laborers.

Immigration policy

The era before the First World War is seen as one of free migration in the greater Atlantic economy—but only for some migrants. The United States severely restricted immigration from China in 1882 and from Japan in 1907. Immigration from the poorest parts of the world was unrestricted only because there was no pressing ‘threat’ of immigration. In 1885 Canada introduced an entry tax on Chinese immigrants, which was raised in subsequent steps. Similar policies were adopted in the Australian colonies and in 1901 newly federated Australia followed the 1897 policy of Cape Colony and Natal by introducing a dictation test designed to keep out non-Europeans. This was adopted by New Zealand in 1907 and by Canada in 1910. Other more incremental changes included entry taxes and restrictions on criminals, lunatics, and those ‘liable to become a public charge’.1 And the positive inducements to immigration offered by some countries such as Argentina, Australia and Brazil were eventually reduced or withdrawn.

In the face of continuing anti-immigration pressure, more radical restrictions were introduced from the First World War onwards. In 1917 the US introduced a literacy test for all immigrants and a ban on immigration from the ‘Asiatic Barred Zone’. This was shortly followed by the first numerical quota in 1921, which became tighter in the Act of 1924 and its amendment in 1929. As illustrated in Figure 1, international migration was dramatically lower in the 1920s as compared with the decades before 1914, and much of this was due to policy. The decline in immigration was most dramatic in the US; for the source countries covered by the 1921 quota, immigration fell by two thirds in the following year. By linking them to the source composition of the foreign-born in the past, the quotas bore heavily on the so-called ‘new immigrant’ countries, as illustrated in Table 1. Although the sharp shift towards restriction in the US has received most of the attention, other countries followed. In 1923 Canada introduced a formal distinction between immigrants from preferred and non-preferred countries. Escalating restrictions were adopted in South Africa and Brazil, culminating in quota systems in 1930 and 1934 respectively. Even the British Dominions adopted severe restrictions limiting immigration from Britain: Australia in 1930, New Zealand in 1931 and Canada in 1932.

---

1 For example, the US Immigration Act of 1882 introduced a head tax (extended in 1891); the Act of 1903 imposed restrictions on anarchists, beggars, those with epilepsy and importers of prostitutes.
The retreat from free migration is sometimes seen as a backlash against mounting numbers (Williamson 1998). But what were the political economy mechanisms? As Foreman-Peck (1992, p. 360) put it: “The two key questions of any political economy of international migration are: (1) who gains and who loses from migration? And (2) who is in a position to do something about it?” The answer to the first question depends on who was most likely to face direct competition from immigrants in the labor market. As immigration from Southern and Eastern Europe expanded, the immigrants were increasingly low-skilled workers. In the 1860s the average immigrant to the United States came from a country with GDP per capita 95 percent of the US. By the 1900s the ratio had fallen to 50 percent and similar declines occurred for immigration to Canada and Argentina (Hatton and Williamson 2007, p. 223). On the other hand the owners of capital and land, and perhaps the more skilled and educated workers, were less likely to favor restriction.

In order to test for these effects, Timmer and Williamson (1998) developed an annual index of immigration policies for five New World countries for 1870 to 1930. They found that for Argentina, Brazil, Canada and the United States more restrictive policies were associated with declining relative unskilled wages. In Canada and the United States, where immigrant origins were more diverse, it was the rise of immigration from low wage countries, differing in ethnicity and religion from earlier waves, which helped to close the door. For the US this is consistent with the imposition of quotas that radically reduced immigration from Southern and Eastern Europe. In Argentina and Brazil, where immigrant origins were less diverse, the rising share of all foreign-born also led to restriction.

While immigration policies typically get tighter in recessions (Shughart et al., 1986), recessions seem to have more decisive effects when they are preceded by a gradual accumulation of forces that shift opinion against immigration. This may help to explain the imposition of the emergency quota in the US, as the postwar resurgence of immigration combined with an increase in the unemployment rate from 5.2 percent in 1920 to 11.7 in 1921. While the US led in the early 1920s, the door was closed in other New World countries as unemployment increased during the Great Depression. But another factor may also have been important. The interwar period saw a dramatic decline in international capital mobility. So the effect of capital inflows in muting the wage effects of immigration would have been smaller than before 1914 (see further below). For this reason workers could have been more opposed to immigration than in the era when capital chased labor (Hatton and Williamson, 2007).

At first sight the answer to Foreman-Peck’s second question depends on who had the right to vote. As suffrage widened it typically percolated down the hierarchy of class and income, diluting the political weight of landowners and capitalists, and giving a stronger voice to urban blue-collar workers. At the turn of the century voting rates were about one third of the adult population in North America (much higher for males) but less than 10 percent in Latin America (Engerman and Sokoloff, 2005). Severe immigration restrictions came earlier in the United States than in the (comparably democratic) British Dominions, which chose to remain open, at least to immigrants from the UK. On the other hand, in independent Latin America, where the franchise was much narrower and excluded immigrants, and where the Latifundia retained their
grip on power, restriction came earlier than might have been expected. Sánchez Alonso (2013) analyzed a newly-created index of immigration policy for Argentina from 1870 to 1930. She found that some of the shift towards restriction was accounted for by rising immigration, declining migrant education relative to natives, and growing inequality. But restrictive immigration policies were instigated not through the ballot box, but as a result of strikes and protracted labor unrest.

The underlying political economy mechanisms have been explored in the greatest detail for the United States. Turning to the attitudes of workers, Richardson (2005) analyzed the opinions of 2,000 blue collar workers surveyed by the Kansas State Labor Bureau in 1895-1897. He found that the vast majority wanted immigration to be restricted or to be completely suppressed. Opposition to immigration was stronger among union members than among non-unionists; interestingly those with incomes in the middle of the range were more opposed than those with the lowest wages. Not surprisingly, the first- or second-generation foreign-born were the least likely to favor restriction. There was also some evidence that opposition to immigration was greater in counties with strong growth in the immigrant population.

How did this anti-immigrant sentiment get translated into restrictive immigration policy? In an influential paper Goldin (1994) examined voting patterns in the US Congress. Beginning in 1897 a series of bills incorporating a literacy test failed to pass into law until 1917 when the House of Representatives and the Senate both overrode President Wilson's veto. By 1906 representatives from the South had joined those from other rural areas in supporting restrictions and the key battleground was the cities of the East and the Midwest. Goldin found that a representative was more likely to vote for an override in 1915 the more rapid the growth in the foreign-born population and the slower the growth of wage rates in their city in the preceding decade. But the higher the immigrant share in the district (the level, not the change) the less likely was a vote for restriction, and especially when the share reached 30 percent. Looking across cities from 1910 to 1930 Tabellini (2017) finds that greater immigrant inflows, especially from ‘new’ source countries, reduced electoral support for the Democrats, led to the election of more conservative congressmen, and increased the likelihood that they would vote for the quota act of 1924.

Biavaschi and Facchini (2017) analyse the House votes on all 14 immigration bills from 1897 to 1924, estimating with fixed effects by session and by congressional district. They find that a representative was less likely to vote for restriction the greater the share of employment in manufacturing and the lower the share urban in the district. The characteristics of the Representative mattered too, with northern democrats and those with non-Ivy League backgrounds more likely to vote for restriction. Interestingly they also find that, while higher shares of foreign-born made a vote for restriction less likely, this was tempered where state residency laws made it more difficult for immigrants to naturalize and to gain the vote. This is consistent with the findings of Shertzer (2016) that the rate of naturalization was greater where immigrant interests aligned with other sizeable minorities to form a winning coalition.
Contemporary controversies over immigration policy have stimulated renewed interest in the shaping of immigration policies in the past. Recent cliometric studies have uncovered a range of influences that are broadly consistent with the political economy of immigration policy mapped out in more qualitative accounts. Working through interest group politics, immigration itself has mixed effects. While a sizeable and well-established stock of past immigrants tends to favor more liberal immigration policy, the recent inflow has the opposite effect. It also seems likely that under certain conditions shocks such as wars and recessions trigger an immigration backlash. As most of the evidence is for the United States it is less clear how these influences played out in countries with different political and institutional settings.

Immigrant selection

Who migrates and how they compare with those left behind is a long-standing question in the literature. The selection of immigrants is related to the fortunes of immigrants in the host country labor market, immigration’s impact on the wages and incomes of the native-born and to the longer run effects of immigration—topics explored further in later sections. The selection of immigrants can be measured in a variety of ways: one dimension is the origin-country composition, which, as we have seen, shifted towards the poorer parts of Europe, raising the concerns of policy-makers. Another is the demographic profile; immigration to the United States increasingly comprised young unattached males—those with low dependency burdens and high labor force participation rates. But the focus of much of the recent cliometric literature is on selection by skill or ability.

The basic framework is the Roy model as adapted to migration by Borjas (1987). In this model the incentive to migrate depends on the individual’s skill level and the returns to skills in the home country compared with the destination country. If the return to skills is greater at the destination then the skilled will be more likely to emigrate; if the return to skills is higher at the source then the unskilled will be more likely to emigrate. But the costs of migration must also be considered; higher costs of migration will reduce the net present value of migrating more for poorer, less skilled workers than for those with higher skills and incomes. This framework is particularly applicable to the age of mass migration, an era before the advent of restrictive immigration policies, and a time when most migrants were moving for economic reasons rather than fleeing famines, wars or political persecution. If the post-bellum United States became more unequal than European sources then that would favor more positive selection from any given source country. On the other hand, declining transport costs and, above all, assistance from the growing stock of previous emigrants would work in the opposite direction.

Testing different theories on the selection of immigrants first requires measuring selection, which can be difficult for numerous reasons. First, measures of immigrant quality in historical data are often crude and rarely include key metrics of interest such as ability, entrepreneurial spirit or attitude towards risk. One way to get a summary measure of a migrant’s productivity is by using his wage, but since the researcher does not observe the immigrant’s wages in the
source country if he had stayed, we need to either measure selection on a pre-migration variable or estimate the counterfactual wage. Despite these methodological problems, there have been significant advances in the literature due to the recent digitization of census and immigrant arrival records. Arrival records document immigrants prior to any assimilation forces causing a change in occupation or other observables, which allows for a (mostly) straightforward comparison of immigrants relative to stayers as observed in source-country census data. However, the quality of information in ship records may be low if ship captains were careless filling them out, and variables in the arrival records may not always match with variables in census data. Instead, some researchers have linked censuses across countries, allowing us to observe the immigrant’s characteristics before and after migration.

The predominant finding of studies using arrival records or census data is that European immigrants to the United States during the age of mass migration were intermediately or negatively selected. Using individual census data on Norwegian immigrants to the United States in the late 19th century, Abramitzky et al. (2013) find that they were negatively selected as the fathers of movers had less wealth than the fathers of stayers. Negative selection has also been found for Irish emigrants prior to 1850, based on the age-heaping reported in arrival records and in the Irish census (Mokyr and Ó Gráda, 1982); Irish immigrants to the US were also negatively selected in the early 20th century based on linked census records (Connor, 2016). Italian emigrants in the early 20th century were shorter than the average Italian left behind, suggesting that they had a worse health environment than stayers (Spitzer and Zimran, 2017).

Most studies measure migrant selection on observable characteristics, but the selection of immigrants on unobservable characteristics is more difficult to gauge. Furthermore, it is unclear whether selection on observable and unobservable characteristics are correlated; it could be that a low-skilled immigrant from an observably poor family was also highly ambitious and entrepreneurial. Abramitzky et al. (2012) use an innovative approach to measure selection on unobservables: they compare the return to immigration measured with and without brothers fixed effects. The idea is that a naïve estimate of the return to immigration by comparing migrants’ occupations to stayers’ occupations is biased if immigration is correlated with a positive or negative omitted variable; however, one can gauge the direction of this bias by estimating the return to immigration within brothers to control for unobservables across households. Abramitzky et al. (2012) find that the naïve return to immigration is more negative than the estimate within brothers, and interpret the negative bias on the naïve estimate as reflective of negative selection into immigration. The negative selection on unobservable characteristics is consistent with other findings of negative selection on observable characteristics such as father’s wealth, perhaps suggesting that the observable measures of selection are a good proxy for unobservable selection.

Can these findings of mostly intermediate or negative selection be rationalized in the context of the Roy model? Stolz and Baten (2012) test the Roy model on data from the age of mass migration by comparing measures of inequality with measures of selection. They use measures of inequality and selection that are well known within the cliometric literature: for inequality,
the amount of variation in height data, and for selection, the amount of age-heaping on zeroes or fives for immigrants relative to those remaining in the source (Steckel, 1995; Steckel, 2008; Tollnek and Baten, 2016). The coefficient on relative inequality across sources offers strong support for the Roy model: migrants to a destination from more equal origins were more negatively selected on numeracy. They also find that sharing a common language is associated with positive selection while wars and upheavals at the origin lead to negative selectivity. This latter effect may also apply to famines, which would help explain the negative selection of emigrants from mid-century Ireland. However, neither the predicted positive effect of origin-country poverty, nor the negative friends and relatives effect could be consistently identified.

The effect of liquidity or wealth constraints on migrant selectivity is something of a puzzle, but it seems likely to have become less important over the nineteenth century, if for no other reason than assistance from the growing stock of previous migrants. For example, in 1908 92 percent of immigrant arrivals to the US were joining a relative or friend at arrival and about one-third had their tickets paid for by someone else. Therefore, while on the margin poverty constraints did limit some migrants, they were not binding for many. However, wealth constraints were likely stronger during the earlier stages of the mass migration, when networks were less developed and prior to cost reductions due to the diffusion of steam technology following the Civil War (Cohn, 2009). As noted earlier, Wegge (1998) shows that networked migrants from the Hesse-Cassel region in mid-century Germany required less cash to move, implying that they depended on help from previous migrants. And even though Irish emigrants were negatively selected overall, during the Great Famine the very poorest were less able to emigrate and were more likely to have starved (Mokyr and Ó Gráda, 1982; Cohn, 1995). There is other evidence that poverty constraints limited migration from the lower end of the skill distribution. While Italian immigrants were negatively selected overall, they were positively selected from the poorest provinces, suggesting that costs restricted migration of the poorest (Spitzer and Zimran, 2017).

Focusing on the drivers of immigrant skill levels, Covarrubias et al. (2015) use United States inflow data from 1899 to 1932 to determine the relative importance of source country income and costs of immigration on the skill composition of inflows. They find that higher origin-country income levels increased the total number of migrants but slightly decreased their quality, suggesting that a poverty constraint was alleviated. On the other hand transport costs had little effect on the skill level of the inflow, perhaps because costs are represented by freight rates rather than by passenger fares. The severe restriction on immigration imposed from the early 1920s caused the selection of immigrants to be more positive, at least in terms of occupation, as the implied policy costs fell more heavily on the lower skilled. Therefore, not only did the immigration quotas have a between-country effect as the quotas imposed different degrees of restriction across origin countries, but there was also a within-country effect as lower skilled individuals were more likely to be restricted (Covarrubias et al., 2015; Massey, 2016).

Perhaps what policy makers and economists are more interested in is the quality of the net flow rather than just the inflow; as we have seen, since millions of migrants returned to Europe, the quality of the net flow depends on the selectivity of return migration. The evidence suggests
that the average return migrant was lower-skilled than the average permanent immigrant, a pattern that holds both across countries and within country (Abramitzky et al., 2014; Ward, 2017a). The negative selection of return migrants presents something of a puzzle as the Roy model predicts that the selectivity of the return flow should be opposite to the selectivity of the inflow (Borjas and Bratsberg, 1996), which was also negatively selected. This may be because those who returned home were more likely to receive a negative shock in the United States labor market, perhaps due to occupational downgrading after arrival.

The recent literature has challenged the view often presented in qualitative accounts that immigrants into the United States were, on the whole, positively selected. Instead the evidence suggests that immigrants overall were somewhat negatively selected, but with variations by country and region of origin. So far there is little comparable evidence for other New World destinations during the age of mass migration, but there is reason to think that the results could be different. For example, Italian migrants to Argentina had relatively high literacy rates, especially those from Spain and Portugal (Sánchez Alonso 2007). And the occupational composition of emigrants to Australia 1877-1913 was more skilled than that of Canada and the United States (Pope and Withers, 1994).

**Immigrant assimilation**

These new findings on immigrant selectivity raise questions about other widely debated issues. Following (Borjas 1987), the selectivity of immigrants has been intimately connected with their assimilation into the US labor market. The assimilation of immigrants tackles a similar question of the selection literature but from a different reference point. Instead of comparing immigrants to non-migrants in the source country, the assimilation literature compares immigrants to the native-born in the destination country. This comparison is most often for wages or occupational status, but has also been made along different dimensions such as geographical location, fertility, or marriage rates. In this section we will primarily focus on standard economic outcomes such as occupation, wages and employment.

The classic finding in the literature is that for the post-1950 decades, immigrants initially received lower wages than natives and then converged to native wages in the decades after arrival (Chiswick 1978). This finding is often interpreted in a human capital framework: at arrival, immigrants lose some value of their human capital acquired at home since it does not transfer across borders perfectly, but then they gain host-country-specific human capital, such as language fluency, after arrival. Another reason for convergence between immigrants and natives (which links to the selection issue) is that immigrants are thought to have different abilities or ambition and drive than the native-born, which implies that they would upgrade income throughout the life cycle faster than the native-born. However this optimistic view of

---

2 Abramitzky et al. (2016b) provide new microdata on returners to Norway and show that despite return migrants being negatively selected relative to both movers and non-movers, they achieved higher outcomes upon return to Norway, reflecting a positive return to temporary migration.
immigrant assimilation has been downgraded in the subsequent literature that takes account of cohort effects and selective return migration (Borjas 1985; Lubotsky 2007).

Immigrant assimilation in the late nineteenth century US labor market has been explored using data on the earnings of immigrants and native-born from a range of reports from state labor bureaus. The initial results seemed to support the revisionist view that the age-earning profiles of immigrants were no steeper than those of the native-born and that, over the life cycle, immigrants may even have fallen further behind their native-born peers (see for example Hanes 1996). By contrast, Hatton (1997) showed that the results turn partly on the shape of the age-earning profile combined with the fact that the immigrants in these samples are typically older than the native-born. Taking this into account, immigrants that arrived before the age of sixteen had age-earning profiles very similar to the native-born. Those that arrived as adults had earnings that were initially lower, and that increased with age, but did not overtake the native-born. However these results stem from cross-sectional analysis of subsets of blue-collar workers and they do not account for cohort effects or for selective return migration. 3

This view has now been subjected to further revision. Abramitzky et al. (2014) find that the average immigrant arrived similarly skilled and upgraded at the same rate as natives. The difference in results is due to data structure: both cross sections and repeated cross sections, as were used in the earlier literature, overstate immigrants’ assimilation rate because of declining cohort quality and negative selection of return migrants. Abramitzky et al. (2014) improve on repeated cross sections with new panel data by linking immigrants across the 1900 to 1920 United States Censuses. Since the immigrants in the panel data are the same ones throughout time – as opposed to pseudo-cohorts used in the repeated cross-section methodology – selective return migration does not bias the estimates. Abramitzky et al. (2014) recreate the finding of positive assimilation using repeated cross sections, but then demonstrate that this pattern disappears once using the panel data, which implies that the positive rate of assimilation in earlier studies was primarily driven by negatively selected return migrants. 4

One limitation of these results is that the 1900 to 1920 Censuses do not record wages, but only occupational status; it is possible that immigrant assimilation profiles were steeper for earnings. But a lack of relative occupational upgrading after arrival is still surprising given the human capital model of assimilation, suggesting either that post-migration human capital was not valuable or that immigrants were negatively selected on ambition or drive relative to native-born Americans. Ward (2018) explores the importance of one piece of post-migration human capital that is highly valuable today – English fluency – and finds that acquisition of English skills was associated with only a small upgrade in occupation. While the reason for a low occupational return to English skill is unclear, it may be related to the structure of the early

---

3 However, analysis of grouped data by origin country from the reports of the US Immigration Commission and of cohort occupational progression in the 1900 and 1910 censuses reach similar conclusions (Hatton 2000; Minns 2000). For a more detailed review of this generation of studies, see Hatton (2011).

4 Abramitzky et al. (2014) focus on European arrivals given their importance in the flow. For non-European sources, Kosack and Ward (2018) estimate the assimilation rates of Mexican immigrants, and Hilger (2016) estimates outcomes for descendants of Asian immigrants.
20th century economy: the skill premium was relatively low and jobs which required communication skills were not as prevalent.

Despite a low occupational return to English fluency, people at the time were convinced of its importance (Jenks and Lauck, 1926). This was a contributing factor to the “Americanization” movement of the 1910s and 1920s, which aimed to assimilate immigrants into American society typically through education programs (US Immigration Commission, 1911; Bandiera et al. 2018). However, it is unclear how effective the Americanization movement was: Lleras-Muney and Shertzer (2015) argue that compulsory schooling and English-only laws were ineffective for raising immigrants’ income in 1940. Moreover, these laws also led to a backlash from the targeted population. For example, German Americans in states that banned German instruction were more likely to give their children German names and less likely to volunteer in World War II (Fouka, 2016). While programs aiming to assimilate immigrants may not have been effective, there is evidence that the appearance of being American was economically important: Biavaschi et al. (2017) show that adopting a more American-sounding name yielded a large occupational-based return, and Abramitzky et al. (2016a) show that the second generation with less foreign-sounding names had higher incomes relative to their brothers with more foreign-sounding names (see also Goldstein and Stecklov (2016) and Carneiro et al. (2015)); moreover, becoming a citizen was associated with improved occupational status (Catron, 2017). This is consistent with Alexander and Ward (2018) showing that younger-arriving brothers were both more socially assimilated and had a smaller wage gap with natives compared with their older-arriving brothers.

The assimilation literature for the United States is limited in that the United States Census does not record incomes until 1940; therefore, while immigrants had similarly paid occupations as white natives, they may not have had similar incomes. Fortunately, the 1911 to 1931 Canadian censuses did measure income, which Inwood et al. (2016) use to show that immigrants upgraded at faster rates than natives through the life cycle – a result that differs from the lack of occupational-based assimilation in the United States. The Canadian result is consistent with a positive return to Canadian-specific human capital and a positive return to English skills. Yet this result must be interpreted with caution because the Canadian random samples form a repeated cross section, making it unclear whether their assimilation result is due to true upgrading or selective return migration.

In contrast to the lack of upward occupational mobility for immigrants relative to natives in the United States, there was more success for Europeans who headed to South America. Using data that links immigrants and natives across censuses, Pérez (2017) shows that immigrants to Argentina upgraded their occupations at faster rates than natives, ending up higher skilled than natives after decades of stay. Importantly, this pattern exists when tracking the same individual over time, which avoids the problem that selective return migration may bias rates of upward

---

5 This behaviour of resisting assimilation is also explored for French Americans in New England by MacKinnon and Parent (2012).
mobility. This relative advantage for the first generation also persisted to the second generation of Europeans, who were more likely to be literate and own property than the sons of natives.

This persistence of occupational outcomes across generations of immigrants has been shown not only in Argentina, but also in the United States (Abramitzky et al., 2014); moreover, it contrasts with the “melting pot” analogy which suggests that parents’ or grandparents’ country of birth does not strongly predict outcomes. Ward (2017b) extends the intergenerational literature beyond the second generation to the third generation and shows that immigrants’ grandchildren in 1940 tended to be in the same occupational categories as the first generation in 1880. Mean skill level by country of origin persisted more strongly than would be predicted from a standard multi-generational model, suggesting that it took more than three generations for occupational differences across source countries to disappear. Therefore, while the melting pot may have been hot in terms of social assimilation, strong persistence in occupational outcomes meant that it was cooler for economic assimilation (Abramitzky et al., 2016a).

The effects of migration

One of the most contentious issues in the immigration debate has been the effect of immigration on the economy in general, and on the wages and incomes of non-immigrants in particular. A variety of methods have been used to estimate the impact of migration on average wages in origin and destination countries during the age of mass migration. Taylor and Williamson (1997) estimate the effects of international migration from 1870 to 2010 for 5 New World countries and 12 Old World countries. They first calculate the cumulative effects of immigration or emigration on the country’s labor force and then use labor demand elasticities to estimate the long-term effects on real wages. For the New World they find that, in the absence of immigration after 1870, real wages would have been higher in 1910 by 9 percent in the United States, 17 percent in Australia and 27 percent in Argentina. Emigration from European countries increased real wages relative to the counterfactual but by smaller amounts, corresponding to the more modest effect of emigration on labor force growth. According to these estimates the real wage gap between the New and Old Worlds fell by 11 percent, whereas under the no-migration counterfactual it would have increased by 10 percent.

Some studies have allowed for a richer set of adjustments by using multi-sector computable general equilibrium models that include three factors of production (land, labor and capital) and also account for international trade. To give one example, O’Rourke et al. (1994) estimated that, in the absence of international migration after 1851, urban wages would have been higher in 1910 by as much as 34.0 percent in the United States and lower by 12.2 percent in Britain. To the extent that factor intensities become more similar, migration tends to be a substitute for trade and so trade grows less rapidly than it would have in the absence of migration. But more important is the assumption that is made about international capital mobility. With perfect capital mobility rather than totally inelastic capital (as assumed above), the effects of

---

6 Other studies in this vein are reviewed in more detail by Hatton (2011).
immigration on wages are much smaller as capital chases labor to exploit the incipient increase in its returns. Thus with full capital mobility the no-migration counterfactual US wage in 1910 would have been just 9.2 percent higher (as less capital flowed in) and the UK wage would have been 6.6 percent lower than the actual (as capital stayed at home).

Exercises such as these hinge on assumptions about the structure of the economy, in particular diminishing returns to labor and therefore downward sloping labor demand curves. In the economics literature, attempts to test these effects have met with mixed results and have led to substantial disagreement about how to model and estimate the labor market impacts of inflows within an economy. Various approaches have been taken, including exploiting the variation in skill mix of inflows at a national level (Borjas, 2003), the variation in the number of immigrants across geographical regions (Altonji and Card, 1991), or a combination of skill mix and region (Card, 2001). More recently Ottaviano and Peri (2012) assume that immigrants and natives are imperfect substitutes even within skill groups, which tends to reduce the direct negative effect of immigrants on the wages of comparable natives as estimated by Borjas (2003).

Rather than directly estimating the effect of immigration on native outcomes, several studies explore the effect of immigration on average earnings overall (including both natives and immigrants). Goldin (1994) uses average wages by occupation and city to demonstrate that a 1 percentage point increase in the foreign-born share led to a 1 to 1.5 decrease in wages for unskilled workers – an effect that is larger than one would expect from a compositional shift to more immigrants. Examining state-level time series from 1929 to 1957, Biavaschi (2013) finds that immigrant inflows had a negative effect on earnings while outflows had a positive effect. Exploiting the reduction of inflows from the 1920s immigration quotas, Xie (2017) finds that a one percentage-point decline in the foreign-born share at the county level increases manufacturing wages by 2 percent – once again, an effect that cannot be explained by compositional changes. Studies of national time series have estimated that the effect of mass migration was to widen skill premiums in the United States and narrow them in Europe (Anderson 2001; Betrán and Pons, 2004), results that are consistent with the effects on inequality found in general equilibrium models (Betrán et al. 2010).

Some studies have followed the modern-day literature in seeking to identify the effect of immigration on the earnings of natives (rather than on average earnings) using micro data. Green and Green (2016) use micro data on earnings by occupation from the Canadian censuses of 1911 to 1931 to estimate the impact of immigration on natives, following the methodology of Ottaviano and Peri (2012). Their estimates show that immigration did not have a large impact on native earnings and thus did not cause the increase in Canadian inequality during the early 20th century. This contrasts with the negative effects of immigration found for the United

---

7 Ottaviano and Peri’s (2012) method is to specify a structural model of the economy where newly arrived immigrants, long established immigrants and natives are different types of labor. Further, labor is differentiated by skill in terms of education and labor market experience. The elasticity of substitution between labor types is measured with the data, and then the structural model is estimated. Dustmann et al. (2016) critique this methodology because, due to occupational downgrading at arrival, immigrants’ effective skill level is lower than their observed skill-level, which implies a misspecification when allocating immigrants to education-experience cells.
States (using different approaches). But given the similarity in the immigration inflow between Canada and the United States it seems likely that the same method applied to United States data would yield the same null effect. Unfortunately there are no data on earnings in the US census until 1940, so any analysis of immigration’s effect on wages must use wage data at a level higher than the individual.

Studies that can separate native and immigrant outcomes with micro data in the United States must resort to occupational status, and these tend to find a small negative effect or even a positive effect of immigration on native outcomes. For example, Ferrie (1999) uses 1850 and 1860 United States linked data and estimates that immigration caused unskilled natives to move faster up the occupational ladder, while immigration had a negative effect on skilled natives. Ager and Hansen (2017) show that a reduction of inflows due to the 1920s quotas caused natives to enter lower-skilled jobs, suggesting that more immigration would have positively affected native jobs (as defined by occupational score), a finding which Tabellini confirms (2017). In addition to the measurement issue of lacking micro-data on native wages, these studies all use different empirical methodologies, leaving the effect of immigration on natives unclear, similar to the modern-day literature (Dustmann, Schönberg and Stuhler, 2016).

There are several other margins besides wages and occupation on which immigration affects the economy, such as outflows from high-immigration areas, changes in the output mix, and technological adoption or invention. Hatton and Williamson (1998, Ch. 8) estimate that immigration to the north-eastern states displaced around 40 natives to other states for every 100 of the migrant inflow. And Collins (1997) demonstrates that the fall of inflows due to World War I and the immigration quotas contributed to the Great Migration of African-Americans from the South to the North. While immigration is often associated with urban areas in the early 20th century, immigration also had important effects in rural areas. For example, a reduction of inflows to a county between 1910 and 1940 shifted the crop mix towards less labor-intensive and more capital-intensive crops, although it had no effect on the adoption of labor-saving technologies (Lafortune et al., 2015). Besides agriculture, there were important effects on innovation: Moser et al. (2014) show that the inflow of Jewish scientists escaping from Germany to America during the 1930s led to a large increase in chemical patents, with no evidence of negative or positive spillovers on incumbent scientists.

The large volume of inflows into the New World implies that there were also important impacts on the sending countries. For Ireland, Hatton and Williamson (1998, Ch. 9) find that from the 1860s onwards the population loss induced by emigration led to more rapid wage growth. Rising labor costs also led to a shift within agriculture away from tillage, and fostered mechanization (O’Rourke, 1991). And in Sweden, higher out-migration rates were associated with more inventions per capita (Andersson et al. 2016). The effect of outflows from the source country may not simply be due to a loss of labor; those who left could still have important effects in the source due to remittances or eventually returning home with savings acquired abroad. As noted earlier, remittances were often used to finance further emigration but the flow of remittances also enhanced financial development in countries of the European periphery (Esteves and Khoudour-Castéras 2011). Qualitative evidence also suggests that return
migration to Italy increased farm prices due to savings brought back home (Wyman, 1993; Cinel, 1991), and linked census data from Norway shows that return migrants improved their occupational standing due to their temporary trip abroad, likely because of bringing back savings (Abramitzky et al., 2016b). Thus while the effects of immigration on the working population in the New World seem to have been rather mixed, for those that remained in the Old World the economic effects were largely positive.\(^8\)

The legacy of historical immigration

The immigration literature is primarily concerned with estimating the immediate consequences of inflows, but it is possible that historical migrations have important long-run effects, over 100 years later. It is often asserted that immigrants leave a distinctly positive legacy to subsequent generations. One recent book carried the title: “Exceptional people. How migration shaped our world and will define our future” (Goldin 2011). While such descriptive accounts tend to accentuate the positive, the economics literature has offered a more nuanced interpretation of the long-term effects of immigration on productivity and economic growth, as well as on inequality and on social cohesiveness.

The literature stresses two enduring effects. On one hand ethnic and birthplace diversity brings skills and knowledge that are complementary in production and may lead to higher incomes. For instance, in cross-country analysis Alesina et al. (2016) find that, in the presence of a range of controls, greater birthplace diversity, particularly within the immigrant population, leads to higher income per capita and more patenting activity. On the other hand, diversity could reduce inter-group communication and overall levels of trust, and it may also imply divergence in attitudes towards inequality and in preferences over welfare policy and public goods. Alesina et al. (1999) find that, across US cities and counties, ethnic fractionalization is associated with higher public expenditure but a lower share of expenditure on public goods. Similarly, for the age of mass migration Ager and Brückner (2013) estimate a positive effect of fractionalization on income at the county level, consistent with gains from specialization. On the other hand, polarization tended to decrease income; this may be because cities that received more immigrants, who were more culturally distant from the native population, also had less public spending (Tabellini, 2017).\(^9\)

\(^8\) On the other hand the contemporary social effects are less clear. For the extreme case of Sweden, which lost nearly a quarter of its population to the New World, Karadjia and Prawitz (2016) show that areas with higher outflows had more strike participation, welfare expenditures and support for left-wing parties, which they argue is due to greater political power for citizens.

\(^9\) Output is measured as the sum of manufacturing value added plus agricultural value added. Fractionalization measures the likelihood that two randomly selected individuals are from different backgrounds, while polarization measures the difference between the distributions of ethnicity from a bi-modal distribution. See Ager and Brückner (2013) for a further discussion. Rodríguez-Pose and von Berlepsch (2017) extend the analysis to 2000 and argue that fractionalization in the late 19th century is positively correlated with county income in 2000, while polarization is negatively correlated.
To the extent that culture, human capital or other traits are correlated across generations, there may be long-run effects stemming from the ancestries of today’s population and ultimately from the countries that their immigrant forebears came from. Although disentangling the effects of past immigration from other mechanisms is difficult, a recent literature investigates the effects inherited from immigration over the very long run. This reflects the wider trend in economics towards estimating the causal effects of events and characteristics from deep in the past that shape culture and institutions, and influence economic outcomes up to the present. As Spolaore and Wacziarg (2013) note, this literature has increasingly focused on the qualities that migrants brought with them; thus, what matters for economic development is not a country’s geographic location or resources but the characteristics of its population.

For a cross section of countries Putterman and Weil (2010) examine the effect of origin-country political organization and economic development since 1500 using as weights the ancestral origins of the current population. They find that GDP per capita is higher in countries where the ancestral composition is associated with more advanced state organization but that the greater the degree of fractionalization in ancestry the higher the inequality. Interestingly, they find that the effects of ancestry are stronger than those stemming from the current foreign-born mix. This is consistent with a range of studies finding that past levels of human capital positively influence institutional quality and predict living standards in the present (Glaeser et. al. 2004; Easterly and Levine 2016; Chanda et al. 2014; and Spolaore and Wazciarg 2013). Human capital persistence is one mechanism through which immigration shocks have long lasting effects, but cultural traits also matter. Algan and Cahuc (2010) show that trust levels persist across generations; for example, Swedish immigrants had high levels of trust in the past, which they transmitted to their descendants today.

Much of the mixing of ancestries occurred during the age of mass migration and, even within countries, this could have differential effects in the present that reflect the settlement patterns of the past. Fulford et al. (2015) explore these effects by estimating the ancestries of the population of the United States for each county in each census between 1850 and 2010. Analyzing the data as a panel, they find that incomes are higher in counties which received immigrants from origins with high levels of income, education and trust. And the effects of ancestry persist beyond the age of mass migration up to the present, with little change in intensity. Consistent with other studies they find that while the degree of ancestral fractionalization alone has a positive effect on GDP per capita, ancestry, when weighted by attitudes towards trust, thrift and cooperation, has a negative effect. This is consistent with the notion that while diversity itself has a positive effect, diversity in attitudes does not.

One channel for a long-run impact of immigration is stronger economic connections abroad. Buchardi et al. (2017) explore the correlation between a county’s population ancestry and foreign direct investment to or from the ancestor’s origin country. Using an instrumental variables strategy to predict ancestries across counties, they show that doubling the number of descendants from a foreign country increases the likelihood of present-day foreign direct
investment with that country by 4 percent.\footnote{Burchardi et al. (2017) instrument for the number of foreign born in a county using a mix of push factors, as proxied by the total outflow from the country to the United States, and pull factors to the specific county, which is proxied by the number of foreign born entering the county from other continents.} Provocatively, they argue that FDI with China would be much larger if not for the Chinese Exclusion Act of 1882 and the Asiatic Barred Zone in 1917. This finding parallels the shorter run work on international trade of Dunlevy and Hutchinson (1999), who show that between 1870 and 1910 more immigration from a country was associated with more international trade. One reason why FDI is correlated with the number of descendants from a country is due to a reduction of information frictions across countries (Burchardi et al., 2017).

Rather than connecting historical migrations with modern outcomes through the ancestral channel, Sequeira et al. (2017) gauge the relationship between a county’s foreign-born share (no matter the source) between 1860 and 1920 with county-level outcomes in 2010.\footnote{See Rodríguez-Pose and von Berlepsch (2014) on a similar long-run relationship, where instead of instrumenting immigrants with an interaction with the railroad and decade of arrival, they test the relationship with a variety of different instruments, such as distance from New York and the standard shift-share instrument. They also find that immigration had a long-run positive effect on modern-day outcomes, which they explore further by examining the effect by immigrant country of origin in a separate paper (2015).} Using an instrumental variables strategy to predict the foreign born share across counties, the authors find that counties with a higher foreign-born share between 1860 and 1920 had higher levels of income and education in 2010, an outcome that is not simply due to reallocation across counties.\footnote{To separate the effect of immigration from other factors on long-run growth, they interact the volatile time series of immigration with variation in when a county was connected to the railroad network; they argue that some counties received more or less immigrants not due to county-specific factors, but because counties happened to be connected to the railroad during a boom decade of immigration.} This effect did not appear without precedent in 2010, but can be traced through intervening decades. Counties with more immigrants had higher manufacturing output per capita, greater farm value per acre, and more patents also in 1930. Akcigit et al. (2017) confirm this positive effect of immigration on inventions by showing that areas with more foreign-born expertise between 1880 and 1940 had more patenting and citations between 1940 and 2000.

Some of the persistence in the effects of immigration during the age of mass migration may be due to the persistence of immigration itself. Long established traditions of migration from a particular origin, by eroding the communication gap with the host population, allow subsequent immigrants from the same source to assimilate more easily (Hatton and Leigh 2011). But perhaps more significant is the effect of immigration on education and human capital. As noted earlier, Bandiera et al. (2018) find that the introduction of compulsory state schooling took place earlier in states with higher proportions of immigrants, and especially those from origins with no experience of compulsory schooling. To the extent that immigration stimulated such nation building projects, these could have persistent effects.

It is possible that long term effects might be different for countries where immigrants met rather different host-country conditions than in the United States. However, there is evidence that immigration from Europe to Brazil and Argentina led to higher incomes in the long run. Rocha et al. (2017) examine long term outcomes of European immigration across municipalities in
Brazil, focusing on the establishment from the 1870s of state sponsored settlements, designed to attract immigrants. These municipalities experienced faster transitions to manufacturing and subsequently to services due to human capital accumulation. By 2000 they had income per capita 15 percent higher than municipalities without settlements. For Argentina, the economic effects seem to have been even larger. Looking across counties, and using frontier military campaigns as an instrument for immigration to localities, Droller (2017) finds that immigration from 1895 to 1914 was associated with more than double the level of income per capita in the present. Again, persistence in human capital was the mechanism, as reflected in the present by more years of education and a higher proportion of skilled workers.

Immigration may have effects on a range of social outcomes, transmitted down the generations by the persistence of cultural norms (Nunn, 2012). One example is the effects of immigration on sex composition. For the US from 1920 to 1940, Angrist (2002) finds that where inflows were heavily skewed towards men, second generation females had lower labor participation rates, were more likely to be married and to have a more-skilled spouse (Angrist 2002). And Lafortune (2013) shows that males responded to fewer females in the marriage market with more investment in education in order to improve their attractiveness in the marriage market. These imbalances tended to dissipate over time but their effects lingered. Grosjean and Khattar (2015) study the legacy of the high male to female ratio in mid-nineteenth century Australia, owing originally to convict transportation. They find that across counties high historic sex ratios are associated in the present with lower female participation rates, fewer women in professional occupations, and more conservative attitudes towards women.

Overall the literature is almost unanimous in finding that immigration during the age of mass migration generated positive economic outcomes both within and between countries. This is all the more remarkable as immigrant arrivals during the age of mass migration do not appear to be exceptional in terms of their selection and assimilation. The main transmission mechanism was by stimulating human capital formation, not only among immigrants and their descendants but also as an externality to the wider population. We may speculate that human capital diffusion also helps to explain why the negative effects of fractionalization, which are sometimes observed as contemporary effects, become muted in the longer term.

Conclusion

In this paper we have examined a wide range of studies focusing on the age of mass migration and its later demise. The literature on the time-series determinants of migrant streams is relatively settled and has not been the focus of the bulk of recent research. On the other hand, the cliometric literature on the political economy of restriction is still comparatively small and deserves further attention. The most recent wave of research on immigrant selection and assimilation has revised pre-existing views, downgrading the view of immigrants as somehow exceptional. And while their effects on host country labor markets are still contested, there is at least some evidence that they had characteristics broadly comparable with the native-born and competed on relatively equal terms. In that light, it is all the more remarkable that the long-
term legacy of the age of mass migration should have been so positive, especially for the United States.

The literature has produced many advances, not only in methodology and data but also in the questions asked and the subtlety of the answers given. Clearly there is room for further research to deepen these insights. But the focus has been heavily skewed towards migrants to, their experience in, and their effects on, the United States, something that is strongly reflected in the studies cited here. The largest dividends for future research probably lie in other New World countries on which less attention has been lavished. And our account completely omits migrations in Asia, which are quantitatively significant and have left enduring legacies but have not received comparable attention. Perhaps a future Cliometric Handbook chapter could be devoted to that, but if it were written now it would be very short.
References


**Figure 1: Emigration from Europe 1846-1939 (five-year averages)**


**Table 1: Pre-war immigration and post-war quotas.**

<table>
<thead>
<tr>
<th>Year</th>
<th>Annual immigration (000s)</th>
<th>Quota allocation (000s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1908-14</td>
<td>1921</td>
<td>1924</td>
</tr>
<tr>
<td>Belgium</td>
<td>5186</td>
<td>1563</td>
</tr>
<tr>
<td>Denmark</td>
<td>6117</td>
<td>5694</td>
</tr>
<tr>
<td>France</td>
<td>8209</td>
<td>5729</td>
</tr>
<tr>
<td>Germany</td>
<td>31,292</td>
<td>68,059</td>
</tr>
<tr>
<td>Ireland</td>
<td>27,571</td>
<td>–</td>
</tr>
<tr>
<td>Italy</td>
<td>202,222</td>
<td>42,057</td>
</tr>
<tr>
<td>Netherlands</td>
<td>6625</td>
<td>3607</td>
</tr>
<tr>
<td>Norway</td>
<td>11,874</td>
<td>12,202</td>
</tr>
<tr>
<td>Portugal</td>
<td>9166</td>
<td>2520</td>
</tr>
<tr>
<td>Spain</td>
<td>5021</td>
<td>912</td>
</tr>
<tr>
<td>Sweden</td>
<td>16,642</td>
<td>20,042</td>
</tr>
<tr>
<td>UK (includes Ireland in 1921)</td>
<td>42,658</td>
<td>77,342</td>
</tr>
</tbody>
</table>