The London School of Economics and Political Science

The long-term economic impact of migration and its significance for US prosperity

Viola Konstanze Sitta Freiin von Berlepsch

A thesis submitted to the Department of
Geography and Environment of the London School of Economics for the degree of
Doctor of Philosophy in Economic Geography, London, March 2018

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of 62.828 words.

Statement of conjoint work

I confirm that chapter 5 was jointly co-authored with Andrés Rodríguez-Pose (LSE, Geography & Environment) and Chapter 6 was jointly co-authored with Andrés Rodríguez-Pose and Neil Lee (LSE, Geography & Environment). In both these works I contributed a minimum of 50%.

Viola v. Berlepsch

Abstract

Does past migration matter for economic development in the long-term? Does an area's history in migration affect economic performance long after the initial migration shock has faded away? And – does it matter what type of immigrant settles in a territory for the economic impact of migration to persist in time? This dissertation examines the long-term economic impact of migration, connecting migrant settlement patterns at the turn of the 19th to the 20th century to present day levels of income per capita. It firstly estimates the effect of different compositional features of the historical migrant stock on long-term economic development levels in the United States (US), a country founded and essentially formed by migrants. Secondly, it tests whether there is a link between past European and recent Latin American migration to the US to identify whether one potential transmission mechanism could be at play in transferring the migrants' economic impact across time.

The results of the analyses conducted using a variety of methods – OLS, IV, and panel data estimation techniques – provide three novel insights. Firstly, historical migrant stock is one of the very few historical county features that still explain current levels of development. In contrast to other factors, such as past income and education levels or industry structure, the influence of past migration on economic development does not seem to fade over the very long-term.

Secondly, compositional aspects related to the historical migrant stock remain highly decisive for economic development outcomes more than 100 years later. The diversity of the migrant population, the gender balance, as well as the average distance travelled by the migrant stock over a century earlier still influence regional economic

development levels today. All three features have growth-enhancing implications over the short as well as over the long-term.

Lastly, past migration – irrespective of the presence of family connections, ethnic ties, or migration networks – shapes the geographical patterns of successive migration waves spanning multiple decades and even generations. An area's migration history acts as a crucial pull factor for future migrants and is at the root of the formation of migration-prone and migration-averse regions. Consequently, previous migration contributes to 'rework' the places of destination, making them more attractive for future generations of migrants.

All in all, the findings show that migration not only matters for economic development, but that its economic influence determines the success and prosperity of territories and the well-being of their inhabitants over the very long-term.

Acknowledgements

There are a range of people to whom I would like to express my deepest gratitude. It is them who made this PhD possible, who kept me going, challenged me, cheered me on, and supported wherever they could.

First and foremost, I wish to thank my parents and sister who have always been there for me, turned me into the person I am today and always supported my – sometimes maybe not always straight forward – choices in life. Thank you for all your encouragement, inspiration, your never-ending support and love. Mama, Papa and Pia – Danke für alles.

My second 'thank-you' is to my supervisor Andrés Rodríguez-Pose. Andrés, about eight years ago, you started to guide me with your professional, academic and personal advice. In all this time, you have become a mentor and great friend. Thank you for always challenging me and pushing me to my limits. This PhD would have been impossible without you.

I also want to thank my second supervisor, Neil Lee. Neil, I am very grateful for your support and valuable feedback over the years. Many thanks for always having your door open to answer questions, to challenge my thoughts and to encourage me.

I am also deeply grateful to my PhD colleagues and London friends who have turned these past years into such an enjoyable, rewarding and fun ride. No matter the day, hour or moment, you guys have always listened, laughed, celebrated, travelled or worked out with me, keeping me on track and, most of all, sane! You have become incredible friends. Thank you for everything – I will miss you!

Lastly, I would like to thank everyone in the department for their support, advice and feedback. The LSE has been a truly extraordinary place for my PhD. Thank you for making this such a special and wonderful experience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

Migration is increasingly becoming the centre of policy discussions, political discourse, and the scientific research agenda. From prehistoric times until today, humans have spread across the globe. Moving to a different region, territory or country has always been a natural consequence of factors such as climate, violence, poverty or population change. Yet, in recent decades, the size of the total worldwide migrant community has skyrocketed to levels unknown since the beginning of recorded global migration numbers. During the past 55 years, worldwide migrant stock has more than tripled from about 72 million in 1960 to 244 million in 2015. While the number of migrants grew at around one percent per year between 1960 and 1985, its growth rate picked up speed and increased to, on average, three to five percent per year between 1985 and 2015. A 71 million absolute increase in the size of global migrant stock was recorded in the 2000s alone (United Nations, 2012). Thus, the volume of international migration has grown faster than the world's population as a whole, resulting in migrant communities representing often more than 10 percent of the population in the receiving countries (United Nations, 2016). Consequential to these global population movements across space and – more importantly – across international borders, the question of how inward migration affects a country's economy has re-emerged at the top of the political agenda. Policy makers and scientific researchers alike have been prompted to react.

A vast amount of literature has been developed to assess the economic effect of migration on the receiving country's economy, responding to old-established fears that migration may lead to economic downturn and unemployment, and crowd-out the

native population from the labour market. Some voices do indeed consider migration as growth-deteriorating and a fiscal burden on the native population. They refer to increasing welfare costs, skill alterations on the labour market, lower native wages, rising inequality and crime (i.e. Card, 2001, 2009; Borjas, 2003; Storesletten, 2003; Dustman et al., 2004; Barrett and McCarthy, 2008; Alonso-Borrego, Garoupa and Vázquez, 2012).

The vast majority of studies, however, tends to adopt a contrasting view. Inward migration is generally believed to be growth-enhancing, boosting levels of innovation, technology, and productivity (i.e. Clemens, 2011; Kennan, 2013; Di Giovanni, Levchenko and Ortega, 2015). Over the years, inward migration has been connected to higher total GDP levels (Ortega and Peri, 2009), rising (or unchanged) native wages (Dustman et al., 2004; Card, 2005; Ottaviano and Peri, 2012; Dustmann, Frattini and Preston, 2013), higher productivity (Hirschman and Mogford, 2009; Hunt and Gauthier-Loiselle, 2010), rising employment (Card, 1990; Ortega and Peri, 2009; Peri and Sparber, 2009), increased innovation (Partridge and Furtan, 2008; Özgen, Nijkamp and Poot, 2012), accelerated technology formation (Alesina, Harnoss and Rapoport, 2016; Bove and Elia, 2017), rising efficiency levels (Kennan, 2013), and higher entrepreneurial activity (Wadhwa *et al.*, 2007; Clark and Drinkwater, 2010; Fairlie and Lofstrom, 2015). Restricting migration would therefore entail profound economic losses, significantly lowering the receiving region's economic growth prospects.

The aforementioned studies have one factor in common: The impact of migration on economic development outcomes is generally analysed over the short- to medium-term. The economic and social consequences of inward migration on the aggregate production function or redistribution effects within the host country or region tend to

be estimated just for a few election cycles or only one generation, respectively – usually a maximum timeframe of 25 years. The potential far-reaching implications of migration over longer periods of time after the initial arrival of the migrant, are either completely ignored or regarded as irrelevant.

The often-used metaphor of the melting pot in conjunction with immigration and assimilation in heterogeneous societies would support this lack of interest in the long-term consequences of migration. Borjas (1992: 123) explains the melting pot theory in the context of immigration to the United States (US): "Over time the children and grandchildren of immigrants moved out of ethnic enclaves, discarded their social and cultural background, and […] became indistinguishable from the native population". Thus, large numbers of inflowing migrants would lead to rising dynamism, increased economic activity, and ultimately higher growth rates only as long as they are significantly different from the native population. Once skill complementarities, different cultural perspectives, ideas and distinctive cultural identities fade away with time – or, in other words, once migrant groups 'melt' with their host society – migration would turn into a simple population redistribution mechanism, making the long-term analysis of migration meaningless.

Contenders of the melting pot idea, however, support the view of cultural pluralism emphasising a multicultural, pluralistic population where each ethnic or cultural group preserves their culture, tradition, and national heritage forming a society embedded in a particular cultural mosaic (i.e. Glazer, 1970, 2000; Alba, 1999). A growing number of scientific studies support the idea that cultural characteristics, ethnic capital, or a cultural institutional framework do not fade away with time but survive over decades and even centuries (i.e. Borjas, 1992; Acemoglu, Johnson and Robinson, 2001; Duranton, Rodríguez-Pose and Sandall, 2009; Algan and Cahuc, 2010; Tabellini,

2010). Borjas (1992: 124), for example, finds that "if the external effect of ethnicity is sufficiently strong, ethnic differences […] are likely to persist for many generations (and may never disappear)". He estimates that they survive at least four generations or about 100 years (Borjas, 1994). Acemoglu et al. (2001) and Duranton et al. (2009) extend this time frame. They show that cultural differences established in colonial times (the former) or even in medieval times (the latter) persist and influence socioeconomic outcomes to this very day. Thus, if indeed cultural capital, ethnic institutions, or ancestral characteristics survive over multiple generations affecting socioeconomic outcomes over very long timeframes, should migrants from a multitude of countries not generate economic consequences that last longer than 25 years, potentially affecting regional economic growth levels for decades to come?

Studies on the long-term impact of migration, however, are few and far between. Only recently, researchers have started to analyse the impact of migration extending the time dimension after the initial arrival of the migrant. Rodríguez-Pose and von Berlepsch (2014) were the first to evaluate questions such as: Does inward migration entail long-term implications for future generations? Do migrant stocks generated in historical times explain current disparities in economic development? Evaluating the United States at the turn of the 19th to the 20th century, they find that migrants leave an undeniable imprint on economic development for more than 100 years. Regions that received a large number of inward migrants 130 years ago are significantly more prosperous today than those which were largely bypassed by migration routes. These results are confirmed by Sequeira et al. (2017).

The strength of these findings triggers a number of additional questions: Why and how does migration leave such a long-lasting, growth enhancing legacy for regional economic development? Does the secret lie, potentially, in the composition of the

migrant stock? What are the mechanisms that transfer the undeniable imprint of historical migration across decades? How is it possible that the migrants' trace left in historical times still affects regional growth levels despite high internal migration rates and assimilation mechanisms within the new society? This work intends to answer these questions.

The first part of this dissertation treats the *composition* of the regional migrant community and its implications for economic development over long time periods. Evaluating different characteristics of migrant communities in historical times, I seek to ascertain if a specific set-up of the migrant stock is more beneficial in terms of regional economic development than others. In particular, I look at three different compositional features: diversity, gender, and distance.

The diversity of the migrant stock is highly likely to be a decisive element determining the impact of migration on economic development outcomes. A large amount of literature has been developed over the past decades, going back to the works by Jane Jacobs (1961, 1969), identifying diversity as central driver of creativity, and thus, innovation. With a diverse migrant stream, a multitude of ideas, abilities and perspectives convene, creating a fertile soil for creativity, boosting innovation levels and consequently, economic activity. The question that remains is if the economic benefits of this 'diversity buzz' apply only to the short-term or extend over longer time frames. Are high diversity levels, generated in a region's population more than 100 years ago, still relevant for present-day levels of economic development?

The second compositional feature studied in this thesis is gender. When analysing the economic impact of migration, most academic research pools female as well and male migrants to form one homogenous mass of migrants. Women are assumed to be 'tied

movers' or economic dependents following their husbands' migration decision. If gender enters the equation, its role is the one of a control variable, rather than a variable of interest (Pfeiffer *et al.*, 2008). The fact that female migrants could possibly have a different macroeconomic impact on their receiving region than their male counterparts, is – in most cases – simply ignored. This dissertation tries to fill this gap. Does gender affect the long-term impact of migration? Do female migrants leave a different trace for long-term economic development than the migrant in general? Does the gender composition of migrant stocks matter?

The last compositional aspect evaluated in this thesis is the distance travelled by migrants. Generally, migrant streams are made up of both foreign-born as well as native-born populations. While the former cross international borders, the latter cross provincial, state, or regional lines. As all originate from the same country, internal migrants, are predominantly regarded as one homogenous group. I question this fact. I seek to determine whether long distance internal migrants affect long-term economic development outcomes differently than those which travelled only over short distances. Does the migrants' covered distance between outset and destination matter for the impact of migration on economic development over long time frames?

In the evaluation of the different compositional features of migration and its long-term impact for economic development, I assume the transmission mechanism across time to be the very territory of the migrants' settlement. I presume circumspectly that the positive impact of inward migrants created via the dynamic, creativity-sparking, risk-seeking, and entrepreneurial characteristics generally associated to the migrant population have become embedded in the territory where they settled in large numbers. As they rework the territory itself, the growth stimulating features of the migrant population become part of the territory's local institutional framework. This means

that territories that receive large influxes of migrants would end up incorporating a 'buzz' of migration[1], while others, largely bypassed by migration over time, would lack this feature. A clear separation of counties into migration-prone and migration-averse regions would be the logical result. Consequently, migrant flows arriving to the US today would be more likely to be settling in migration-prone regions or, in other words, in the areas where huge numbers of migrants around the turn of the 19th to the 20th century took up residence. This implies a path dependency between old and contemporary migration. The quest for evidence supporting this theory provides the focus of the second part of this work.

Seemingly predestined for the analysis of the long-term impact of migration is the US – a country essentially founded and formed by migrants. Built by pilgrims and settlers, colonists and slaves, Europeans, Asians, Latin Americans, and Africans, the US is the country which understands itself as a nation of migrants. Millions went to the United States in the past, trying to make a better life for themselves and their family whilst escaping from war, prosecution, draught, famine or political unrest, searching for freedom, democracy, economic opportunity, or religious liberty. Today, "more than 99 percent of the current U.S. population can at least theoretically trace its ancestry back to people who came […] from somewhere else" (Spickard, 2007: 4). Migration, thus, lies at the very core of the identity of the United States, shaping its past, present, and future.

This dissertation is therefore placed in the US context within the era of mass migration to the US (around the turn of the 19th to the 20th century). Connecting the migration settlement pattern generated in the 19th century to current levels of economic

[1] We use the word 'buzz' here in analogy to the path-breaking work by M. Storper and A. Venables (2004)

development, I seek to ascertain whether historical migration to the US in its different forms and compositions still impacts economic development of US counties today. Understanding the economic effect of large migration waves on regions 20 to 130 years after migrants first set foot on the new grounds will complement and advance the current body of research on the economics of migration and lead to improvements in our knowledge of how migration affects the long-term economic prospects of territories. “Ethnicity matters and it matters for a very long time” (Borjas, 1994: 572).

1.1 REFERENCES

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001): The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*(5), 1369–1401.

Alba, R. (1999): Immigration and the American realities of assimilation and multiculturalism. *Sociological Forum*, *14*(1), 3–25.

Alesina, A., Harnoss, J., and Rapoport, H. (2016): Birthplace diversity and economic prosperity. *Journal of Economic Growth*, *21*(2), 101–138.

Algan, Y., and Cahuc, P. (2010): Inherited trust and growth. *American Economic Review*, *100*(5), 2060–2092.

Alonso-Borrego, C., Garoupa, N., and Vázquez, P. (2012): Does immigration cause crime? Evidence from Spain. *American Law and Economics Review*, *14*(1), 165–191.

Barrett, A., and McCarthy, Y. (2008). Immigrants and welfare programmes: exploring the interactions between immigrant characteristics, immigrant welfare dependence, and welfare policy. *Oxford Review of Economic Policy*, *24*(3), 543–560.

Borjas, G. J. (1992): Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics*, *107*(1), 123–150.

Borjas, G. J. (1994): Long-run convergence of ethnic skill differentials: The children and grandchildren of the great migration. *Industrial & Labor Relations Review*, *47*(4), 553–573.

Borjas, G. J. (2003): The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, *118*(4), 1335–1374.

Bove, V., and Elia, L. (2017): Migration, diversity and economic growth. *World Development, 89*, 227–239.

Card, D. (1990): The impact of the Mariel Boatlift on the Miami labor market. *Industrial & Labor Relations Review*, *43*(2), 245–257.

Card, D. (2001): Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, *19*(1), 22–64.

Card, D. (2005): Is the new immigration really so bad? *Economic Journal*, *115*(507), F300–F323.

Card, D. (2009): Immigration and inequality. *The American Economic Review*, *99*(2), 1–21.

Clark, K., and Drinkwater, S. (2010): Patterns of ethnic self-employment in time and space: Evidence from British Census microdata. *Small Business Economics*, *34*(3), 323–338.

Clemens, M. (2011): Economics and Emigration: Trillion-Dollar Bills on the Sidewalk? *Journal of Economic Perspectives*, *25*(3), 83–106.

Di Giovanni, J., Levchenko, A., and Ortega, F. (2015): A global view of cross-border migration. *Journal of the European Economic Association*, *13*(1), 168-202.

Duranton, G., Rodríguez-Pose, A., and Sandall, R. (2009): Family types and the persistence of regional disparities in Europe. *Economic Geography*, *85*(1), 23–47.

Dustmann, C., Fabbri, F., Preston, I. and Wadsworth, J. (2004): The local labour market effects of immigration in the UK. *Home Office Online Report 06/03*, London, Home Office.

Dustmann, C., Frattini, T., and Preston, I. P. (2013): The effect of immigration along the distribution of wages. *The Review of Economic Studies*, *80*(1), 145–173.

Fairlie, R., and Lofstrom, M. (2015):. Chapter 17 - Immigration and entrepreneurship. In B. R. Chiswick and P. W. Miller (Eds.), *Handbook of the Economics of International*

Migration, Vol. 1, 877–911. North-Holland.

Glazer, N. (1970): *Beyond the melting pot: the Negroes, Puerto Ricans, Jews, Italians, and Irish of New York City*. D. P. Moynihan (Ed.), 2nd ed. Cambridge, Mass.: M.I.T. Press.

Glazer, N. (2000): On beyond the melting pot, 35 years after. *The International Migration Review*, *34*(1), 270–279.

Hirschman, C., and Mogford, E. (2009). Immigration and the American industrial revolution from 1880 to 1920. *Social Science Research*, *38*(4), 897–920.

Hunt, J., and Gauthier-Loiselle, M. (2010): How much does immigration boost innovation? *American Economic Journal: Macroeconomics*, *2*(2), 31–56.

Jacobs, J. (1961): *The death and life of great American cities*. New York, NY: Random House.

Jacobs, J. (1969): *The economy of cities*. New York, NY: Random House.

Kennan, J. (2013): Open borders. *Review of Economic Dynamics*, *16*(2), L1–L13.

Ortega, F., and Peri, G. (2009): The causes and effects of international migrations: Evidence from OECD countries 1980-2005. *NBER Working Paper*, No. 14833.

Ottaviano, G., and Peri, G. (2012): Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, *10*(1), 152–197.

Özgen, C., Nijkamp, P., and Poot, J. (2012): Immigration and innovation in European regions. In P. Nijkamp, J. Poot, and M. Sahin (Eds.), *Migration Impact Assessment: New Horizons*. Cheltenham, UK: Edward Elgar Publishing.

Partridge, J., and Furtan, H. (2008): *Increasing Canada's international competitiveness: Is there a link between skilled immigrants and innovation? American Agricultural Economics Association Annual Meeting*. Retrieved from https://core.ac.uk/download/pdf/6563849.pdf

Peri, G., and Sparber, C. (2009): Task specialization, immigration, and wages. *American*

Economic Journal: Applied Economics, *1*(3), 135–69.

Pfeiffer, L., Richter, S., Fletcher, P., and Taylor, J. E. (2008): Gender in economic reserach on international migration and its impacts: A critical review. In A. R. Morrison, M. Schiff, and M. Sjöblom (Eds.), *The International Migration of Women*. New York: Palgrave Macmillan.

Rodríguez-Pose, A., and von Berlepsch, V. (2014): When migrants rule: The legacy of mass migration on economic development in the United States. *Annals of the Association of American Geographers*, *104*(3), 628–651.

Sequeira, S., Nunn, N., and Qian, N. (2017): Migrants and the making of America: The short- and long-run effects of immigration during the Age of Mass Migration. *National Bureau of Economic Research Working Paper Series*, No. 23289.

Spickard, P. R. (2007): *Almost all aliens: immigration, race, and colonialism in American history and identity*. New York: Routledge.

Storesletten, K. (2003): Fiscal implications of immigration - A net present value calculation. *Scandinavian Journal of Economics*, *105*(3), 487–506.

Storper, M. and A. Venables (2004): Buzz: Face-to-face contact and the urban economy, *Journal of Economic Geography*, *4*(4), 351-370.

Tabellini, G. (2010): Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association*, *8*(4), 677–716.

United Nations. (2012): *Trends in Total Migrant Stock: 2012 Revision*. New York, NY.

United Nations. (2016): *International Migration Report 2015*. New York, NY.

Wadhwa, V., Saxenian, A., Rissing, B., and Gereff, G. (2007): *America's new immigrant entrepreneurs*. Durham, NC: Duke University.

2 STRUCTURE

The research is structured in 9 chapters. Chapter 3 and 4 lay out the foundations for the four main parts of the thesis. While chapter 3 provides a short review of the literature on the economic impact of migration in the receiving country, chapter 4 summarises the historical background of immigration to the United States covering the Age of Mass Migration around the turn of the 19th to the 20th century, as well as features of contemporary migration. The bulk of the empirical analysis is spread between chapters 5 and 8, covering different aspects of historical migration and its connection to long-term economic development. While chapter 5, 6, and 7 focus on different compositional elements of the migrant stock (diversity, gender, and distance travelled) and relate these to income per capita more than 100 years later, chapter 8 seeks to find evidence supporting the potential transmission channels of path dependency assumed to preserve the impact of mass migration of the 19th century across time.

Chapter 5, co-authored with Andrés Rodríguez-Pose, examines the economic impact of population diversity on US county wealth for a time frame covering 20 to 100 years. Both fractionalisation and polarisation indices are incorporated into the analysis to create a holistic measure of population diversity. Chapter 6 is a co-authored chapter together with Andrés Rodríguez-Pose and Neil Lee. It focuses on the gender element of migration flows and examines the short- and long-term economic impact of female migration on US county prosperity. Chapter 7, focuses on internal migrants and concludes the sequence of chapters treating different characteristics of the migrant population. It connects US born domestic migrants to long-term economic

development and evaluates if the distance travelled by the migrant is crucial for the determination of the aggregate economic returns of internal migration across a time-frame of more than a century. In the eighth chapter the focus shifts from the migrant stock composition to the validation of path dependency in migration flows. This chapter examines whether migrant waves from different backgrounds, origins and ethnicities, separated by several generations, settle in the same places thereby establishing a permanent separation between migration-prone and migration-averse areas. Chapter 9 concludes and provides policy implications.

2.1 POPULATION DIVERSITY AND ITS LONG-TERM IMPACT FOR ECONOMIC DEVELOPMENT

With growing migrant stocks and accelerating migrant flows around the globe, population diversity in territories attracting large numbers of migrants is increasing drastically. Formerly rather homogenous societies turn ‘multicultural’ in the space of a few years, being confronted with issues such as new language barriers, cultural and ethnic differences, or religious disparities. Transforming this diversity into economic activity has turned into a major task for national governments. Thus, population diversity and its link to economic development has become a prominent field in scientific research. Two opposing strands dominate the debate: On the one hand, diversity is regarded as central driver of innovation which in turn generates technological progress enhancing economic growth (i.e. Jacobs, 1961, 1969; Florida, 2002; Saxenian, 2006; Özgen, Nijkamp and Poot, 2012; Bove and Elia, 2017). On the other hand, diversity is regarded as destabilising factor within a society creating societal tensions leading to social unrest and conflict undermining economic growth

(i.e. Esteban and Ray, 1994, 2008; Easterly and Levine, 1997; Alesina et al., 2003; Montalvo and Reynal-Querol, 2005; Kemeny, 2012; Gören, 2014). Studies commonly refer to two measures of diversity: fractionalisation, placing the emphasis on the number of groups in a society, and polarisation, stressing the distance between them. The focus of these studies, however, has generally been short-term. We simply do not know if population diversity levels generated in historical times have an impact on economic outcomes over more than a 20-year time frame. Chapter 5 intends to fill this gap.

Two research questions are evaluated: Firstly, we examine if initial population diversity – measured as fractionalisation and polarisation – generated during the Age of Mass Migration across US counties, matters for economic development in the long-run. Secondly, we evaluate whether the influence of diversity on economic development has changed over time. Employing US census data from 1880, 1900, and 1910, the settlement pattern of migrants across the counties of the 48 continental states is mapped, building county-level indices of both population fractionalisation and polarisation. Employing OLS as well as IV methods, we regress current economic development, proxied by income per capita in 2010 on county level, on past diversity levels generated around the turn of the 19th to 20th century. Factors which may have determined the attractiveness of a county to migrants at the time as well as those that can impact economic development today, are controlled for.

The results of the analysis show that strong initial levels of fractionalisation, generated more than one hundred years ago, leave a highly significant and positive trace for economic development across time. Polarisation, on the other hand, is found to significantly deteriorate economic development. Hence, counties with a more pronounced population diversity more than one century ago are significantly richer

today, while counties which were strongly polarised during the Age of Mass Migration have endured persistent negative economic implications. Despite a significantly stronger impact on income levels over the short- rather than the long-term, the effect never fades away but remains measurable in terms of higher average income levels to this very day.

2.2 A WOMAN'S TOUCH? FEMALE MIGRATION AND LONG-TERM ECONOMIC DEVELOPMENT

Previous findings in chapter 5 on the long-term impact of population diversity trigger further questions related to the structure of the migrant stream and its connection to long-term economic development. Does the composition of the migrant stock in a county matter for the preservation of the short-term impact of inward migration over the long term? Previous studies have looked at only one of these compositional characteristics, namely national origin. They confirm that the nationality of the migrant is unrelated to variations in long-term economic outcomes (Rodríguez-Pose and von Berlepsch, 2015). The long-term effect of migration seems, thus, not be linked to the national origin composition of the migrant stock.

A different compositional characteristic potentially determining the economic legacy of mass migration to the US is the migrants' gender and the resulting gender balance within the county's migrant stock. Do migrant women trigger a different economic impact over the long-term than migrants in general? Gender, usually regarded purely as control variable, has, in recent decades, moved more closely to the core of academic research. The economics of migration are no exception in this regard. Studies show

that individual characteristics, settlement patterns, the effect on sending countries, the integration into the destination's labour market, or the reasons behind the migration decision distinctively differ from those of their male counterparts (i.e. Ravenstein, 1889; Pessar, 1986; Hondagneu-Sotelo, 1994; Massey;, 2001; Pessar and Mahler, 2003; Oishi, 2005; Andall, 2013). Due to a focus on the individual migrant, macroeconomic approaches on an aggregated level taking the gender component into consideration are rare to find. In the few cases where the economic development aspect of female migration was investigated, studies focused on the short-term impact rather than looking beyond a maximum 10-year timeframe (i.e. Blau, Kahn and Moriarty, 2003; Smith and Bailey, 2006; Riaño and Baghdadi, 2007; Collins and Low, 2010). Chapter 6 seeks to make up for these two shortcomings by investigating the impact of female migration on US economic development over both the short-term as well as a timeframe of more than 100 years.

Using US census data of 1880 and 1910, we map the settlement pattern of migrant women across the counties of the 48 continental US states and calculate female migrant concentration shares. We connect these to current levels of economic development, proxied by GDP per capita in 2010 at county level, while controlling for a multitude of factors which might have influenced both the economic attractiveness of the county at the time of migration as well as the economic development level today. In a second step, we examine the indirect effect of female migrants when analysing the first generation born on American soil and their imprint on long-term economic growth. Challenging the view of the migrant mother as 'cultural carrier' of the migrants' ethnic and institutional baggage, we seek to ascertain if the long-term economic impact of children born to migrant mothers differs from those born to migrant fathers or two American-born parents.

The results of the analysis, conducted using both ordinary least squares and instrumental variable estimations, underline that a) while large shares of migrant women in a county's population have led to significantly lower levels of economic development in US counties both in the short- and long-term, b) immigrant women have left a positive trace for local economic development via their children. Counties with a larger share of children born to migrant mothers have been more economically dynamic over the long term than those with a large share of children born to either foreign-born fathers or both American-born parents.

2.3 INTERNAL MIGRATION AND ITS LONG-TERM IMPACT FOR ECONOMIC DEVELOPMENT

Migration around the turn of the 19th to the 20th century within the United States did not solely consist of international, that is foreign-born migrants, but also American-born, domestic migrants. Internal migration within the US led to an even larger population redistribution phenomenon at the time than the foreign-born inflow from Europe. The analysis of the internal migrant and the distance covered between outset and destination lies at the heart of chapter 7, treating therefore yet another compositional characteristic of a county's migrant stock.

Even though the vast majority of global geographical mobility movements is located within and between regions of the country of birth, a large strand of literature argues internal migration to have received significantly less attention in scientific research than its international counterpart (Skeldon, 2006; Ellis, 2012; Bell et al., 2015). Studies evaluating the macroeconomic effect of internal migration are therefore few

and far between (White and Lindstrom, 2005). Moreover, the few contemporary studies delving into this link deliver inconsistent results (e.g. Yap, 1976; White and Lindstrom, 2005; Rodríguez Vignoli, 2008; Berker, 2011; Molloy, Smith and Wozniak, 2011; Kuhn, 2015).

Two dimensions that may significantly shape the returns of internal migration have, in particular, been overlooked: time and geographical distance. While internal migration research has tended to put the emphasis on the short- to medium-term (thereby rarely covering more than two decades), geographical distance, and its effect on the economic impact of internal migration does not seem to be covered by the social science literature at all. If anything, distance was studied in the context of long distance migration drawing international comparisons across countries (i.e. Long, Tucker and Urton, 1988) or evaluating dynamics, characteristics of migrants and causes of migration (i.e. Biagi, Faggian and McCann, 2011; Pendakur and Young, 2013; Niedomysl and Fransson, 2014).

This chapter intends to cover both of the aforementioned shortcomings in the literature by first, evaluating the effect of US internal migration on long-term economic development and second, examining whether the distance covered by American-born migrants of the late 19th and early 20th centuries matters for the long-term economic impact of domestic migration. Do large shares of internal migrants leave a long-lasting trace for economic development on the territory where they settle in large numbers? Are internal migrants from a faraway county economically more beneficial for long-term economic development than those from next door? Does the distance travelled between outset and destination matter for the impact of internal migrants on subsequent local economic development?

These research questions are analysed using US Census data from 1880 and 1910. The individual data are allocated to the county of residence of the migrant and subsequently aggregated to retrieve internal migrant shares at county level. The settlement pattern of domestic migrants across the 48 continental states is then linked to the average distance travelled by a given county's migrant stock and to current levels of county development proxied by per capita GDP at county level in 2010.

Both ordinary least squares and instrumental variable estimation techniques are employed in order to regress income per capita levels in 2010 on, firstly, the share of internal migrants and secondly, on the average distance travelled by the local migrant stock. Factors which significantly influenced a county's prosperity, both at the time of migration as well as today, are controlled for.

Internal migrants are found to have a highly significant, positive and long-lasting impact on economic development at county level over the very long time frame. Counties which received large numbers of internal migrants in historical times are significantly more prosperous today than those that were largely bypassed. Furthermore, distance is revealed as decisive element for the relevance of internal migration for long-term economic outcomes. The greater the average distance travelled by the migrant stock of a given county, the larger the influence on the long-term economic development outcomes.

2.4 MIGRATION-PRONE AND MIGRATION-AVERSE PLACES: PATH DEPENDENCE IN LONG-TERM MIGRATION

Chapters 5 to 7 show that historical migration in its different compositional dimensions leaves a long-lasting legacy for economic development which can still be traced more than 100 years later. The transmission mechanism assumed to be at work for transferring the positive impact of migration across decades is the territory itself, or to be more exact, the institutional constructs associated to past 'migration buzz' shocks. Chapter 8 seeks to find evidence supporting this potential transmission channel. Connecting recent to past migration stocks, it evaluates if both migrant waves have settled in the same places despite having different backgrounds, origins, traditions, and customs, and being separated in time by at least three to five generations. Thus, it seeks to prove the assumption of an institutional division, established over a century ago, into migration-prone and migration-averse areas.

Scientific research has identified a multitude of regional factors as decisive for a migrant's settlement decision. Next to regional characteristics such as employment opportunities, wages, social welfare spending, public goods endowment, the educational system, as well as urban and natural amenities (i.e. Ritsilä and Ovaskainen, 2001; Zimmermann, 2005; Rappaport, 2007; Partridge, 2010; Biagi, Faggian and McCann, 2011; Ketterer and Rodríguez-Pose, 2015), the existing stock of migrants is considered a crucial factor in determining the attractiveness of a region to incoming population (i.e. Daniels, 1990; Carrington, Detragiache and Vishwanath, 1996; McGovern, 2007; Radu, 2008; Jewell and Molina, 2009; Bodvarsson, Simpson and Sparber, 2014). However, previous and newly arriving migrants have often been connected in the literature via kinship, ethnicity, common background, or friendship

which, in turn, has significantly determined not only the direction of the migrant stream, but also its volume. A crucial factor in reaping the benefits of these migrant networks was sharing a common origin. As 19th century migrants (mostly Europeans) and current US migrant stock (mostly Latinos)[2] neither share a common background nor ethnicity, customs, or traditions, the current academic literature seems to contradict the assumption of path dependency across migration waves.

Analysing US Census data from 1880, 1910 and 1960-2010, aggregated at the county level, I regress current Latino migrant stock on European migrant shares within a county's population generated during the Age of Mass Migration. Controlling for push- and pull factors determining size and direction of past and current migration flows, I employ ordinary least squares, instrumental variable, and panel data estimation methods to ascertain if historical migration created a path-dependence determining the direction and size of current Latino population settlement patterns in the US.

The results of the analysis, underline the importance for historical migration for location-decisions of future migrants. Counties which attracted a large number of European migrants at the end of the 19th century are more appealing to migrants from Latin America 90 to 130 years later. Despite fundamental differences in background, ethnicity, origin and a separation in time of three to five generations, historical migration stocks act as magnet for current foreign-born population serving as an influential pull factor increasing a county's attractiveness. The results therefore provide evidence supporting the hypothesis of the presence of a mark transferred by historical migrants onto their receiving territories. Their legacy created a perpetuating

[2] I am aware of debates regarding the use of the term Latinos vs the term Hispanics (Taylor et al., 2012). In this thesis, I follow Sáenz and Morales (2015) and use both as synonyms. They are employed to refer to the population originating from Mexico, the Caribbean, and Central and South America.

path dependency, permanently differentiating regions into migration-prone and migration-averse areas. A division which seems to persist for centuries even after the ‘original’ migrants have long gone.

2.5 REFERENCES

Biagi, B., Faggian, A., and McCann, P. (2011): Long and short distance migration in Italy: The role of economic, social and environmental characteristics. *Spatial Economic Analysis*, *6*(1), 111–131.

Bodvarsson, Ö., Simpson, N., and Sparber, C. (2014): Migration theory. In B. Chiswick and P. Miller (Eds.), *Handbook of the Economics of International Migration,* 3–51. Burlington: Elsevier Science.

Carrington, W., Detragiache, E., and Vishwanath, T. (1996): Migration with endogenous moving costs. *American Economic Review*, *86*(4), 909–930.

Daniels, R. (1990): *Coming to America: A history of immigration and ethnicity in American life*. New York, NY: Harper Perennial.

Jewell, R., and Molina, D. (2009): Mexican migration to the US: A comparison of income and network effects. *Eastern Economic Journal*, *35*(2), 144–159.

Ketterer, T., and Rodríguez-Pose, A. (2015): Local quality of government and voting with one's feet. *The Annals of Regional Science*, *55*(2), 501–532.

McGovern, P. (2007): Immigration, labour markets and employment relations: Problems and prospects. *British Journal of Industrial Relations*, *45*(2), 217–235.

Partridge, M. (2010): The duelling models: NEG vs. amenity migration in explaining US engines of growth. *Papers in Regional Science*, *89*(3), 513–536.

Radu, D. (2008): Social interactions in economic models of migration: A review and appraisal. *Journal of Ethnic & Migration Studies*, *34*(4), 531–548.

Rappaport, J. (2007): Moving to nice weather. *Regional Science and Urban Economics*, *37*(3),

375–398.

Ravenstein, E. (1889): The Laws of Migration. *Journal of the Royal Statistical Society*, *52*(2), 241-305.

Ritsilä, J., and Ovaskainen, M. (2001): Migration and regional centralization of human capital. *Applied Economics*, *33*(3), 317–325.

Zimmermann, K. (2005): European Labour Mobility: Challenges and Potentials. *De Economist*, *153*(4), 425–450.

3 THE ECONOMIC IMPACT OF MIGRATION – A BRIEF SKETCH

The rising number of migrants and their increased visibility has triggered a shift in the scientific research agenda, bringing the analysis of the economic implications of migration into the fore. A vast amount of scientific research has since been published focusing on a wide variety of topics ranging from the determinants of the migration decision or the individual characteristics of those likely to move, to the evaluation of the economic effect on the host and sending regions as well as the push- and pull factors determining a region's attractiveness to migrants. Within the context of this work, I will be focusing primarily on the economic effects of migration on the host economy, that is, on the region receiving the migrant. In this field, a range of studies has emerged systematically assessing the implications of inward migration for economic growth. Research fields include, among others, the labour market (e.g. wages, jobs, employment), public finances (e.g. welfare services, social benefits), innovation and entrepreneurship (e.g. patents, innovation in products and processes, ethnic firms, competition), and population (e.g. diversity, gender). In general, immigration is confirmed as a growth-enhancing factor, generating substantial economic gains to the global economy in general, but particularly to the receiving country (Ortega and Peri, 2009; Clemens, 2011; Kennan, 2013; Di Giovanni, Levchenko and Ortega, 2015).

Probably the largest concern for policy makers and therefore potentially one of the most controversially discussed research fields, is the impact of migration on local incomes in the host country and the effect on the overall income distribution.

Triggered by the often-held opinion that immigration may reduce the wealth of local inhabitants, research in this area focuses on both the actual effect on native incomes via altering the level of labour supply and the compositional effect related to the skill-set of migrants changing the distribution of skills among a country's residents (LaLonde and Topel, 1991; Borjas *et al.*, 1997; Borjas, 2003; Card, 2005, 2009; Dustmann, Fabbri and Preston, 2005; Cohen-Goldner and Paserman, 2011; Glitz, 2012; Ottaviano and Peri, 2012; Dustmann, Frattini and Preston, 2013). While some find evidence of a lowering of wages and an increase in wage inequality linked to migrant inflows (Card, 2001, 2009; Borjas, 2003), the big majority of studies confirms that despite altering the skill-composition on the local labour market, neither the wages of local citizens nor the wage distribution are significantly negatively affected by these supply shocks. Whenever a negative effect was found, they were regarded as almost negligible (i.e. Friedberg and Hunt, 1995; Longhi, Nijkamp and Poot, 2005; Okkerse, 2008). In fact, a significantly positive effect on the wages of local citizens has been detected when using aggregate production function approaches, meta-analytic analysis, or evaluating cross-region/industry datasets (Card, 2005; Özgen, Nijkamp and Poot, 2010; Ottaviano and Peri, 2012; Dustmann, Frattini and Preston, 2013).[3]

Migration is also believed to increase both the labour supply of local women and of local high-skilled labour (Furtado and Hock, 2010; Cortés and Tessada, 2011). Hence, the impact of migrants on the local, native labour force is not considered negative. Migrants were found to displace native labour, if at all, only in the short-term (Glitz, 2012; Cattaneo, Fiorio and Peri, 2015) and do not seem to be connected to negative effects on the unemployment or employment rates of native labour (Altonji and Card, 1991; Longhi, Nijkamp and Poot, 2008). If anything, migrants often lead to local

[3] For an encompassing overview on this relationship, please refer to Blau and Kahn (2015).

employment booms (Card, 1990; Ortega and Peri, 2009; Peri and Sparber, 2009). "Even those natives who should be the closest substitutes with immigrant labour have not been found to suffer significantly as a result of increased immigration" (Friedberg and Hunt, 1995: 42). Growing unemployment or a crowding-out effect of natives due to increased migrant numbers has been ruled out by Ortega and Peri (2009), whose results show that migration causes a rise in GDP, without decreasing labour productivity.

Inward migration also affects public finances[4]. The migrants' usage of welfare services and other social benefits is a controversially discussed field within political science as well as economic research as it is often used as justification for tighter restrictions on immigration. Depending on the country, the migrant age group, their gender, and skill-level analysed, conclusions vary greatly (i.e. Baker and Benjamin, 1995; Hu, 1998; Gustman and Steinmeier, 2000; Crossley, McDonald and Worswick, 2001; Büchel and Frick, 2005; Blume and Verner, 2007; Barrett and McCarthy, 2008; Pedersen, 2013). Long-term immigrants, for example, were confirmed as highly beneficial as their life-time tax payment greatly outbalances their public sector cost (i.e. Ablett, 1999; Bonin, Raffelhüschen and Walliser, 2000; Moscarola, 2003). Older, female, and short-term immigrants, however, were shown to be a slight fiscal burden to their receiving country (i.e. Hu, 1998; Gustman and Steinmeier, 2000; Sinn and Werding, 2001). Storesletten (2000, 2003) provides an aggregate calculation of the total fiscal burden of a model economy. The costs of immigration are estimated to slightly outweigh its benefits. Similar results are reported for the Netherlands (Roodenburg, Euwals and TerRele, 2003) and Germany (Sinn and Werding, 2001).

[4] For an overview, see Kerr and Kerr (2011).

However, Rowthorn (2008) estimates the net fiscal impact (if positive or negative) to be no larger than 1 percent of the respective country's GDP.

Migrants arriving from diverse locations are depicted as an important input factor in the process of technological progress. Bringing their skillsets, ideas, experiences, and abilities to host regions, inward international migration is more and more regarded as knowledge-generating and -diffusing element, raising innovation at the regional level, and linking it directly to higher economic growth outcomes. Özgen et al. (2012) evaluate the innovativeness of the European receiving regions with respect to size, skills, and diversity of the regional migrant stock. More than the sheer size of the migrant inflow, they find especially the composition in skills and backgrounds to be the decisive element for innovation. Skilled immigrants can boost knowledge creation, efficiency levels, and, therefore, productivity (Hirschman and Mogford, 2009; Hunt and Gauthier-Loiselle, 2010). A more diverse migrant base is considered to spur new ideas and new technology, leading to innovation (Partridge and Furtan, 2008; Lee and Nathan, 2010; Özgen, Nijkamp and Poot, 2012, 2013; Alesina, Harnoss and Rapoport, 2016; Bove and Elia, 2017). Moreover, according to Jacobs (1961, 1969) and Florida (2002), diversity in the cultural composition of the population represents a fertile soil for new ideas, innovation, and economic growth. Further channels generating an effect of migration on innovation, summarized by Özgen et al. (2013), include the positive self-selection of migrants (being more risk-seeking, entrepreneurial and creative), their youthfulness (implying higher mobility, progressivity and creativity), their resilience (enhancing decision making) and their volume (allowing firm expansion, reducing shortages or vacancies of key personnel).

Closely linked to innovation is the migrants' impact on entrepreneurship. Since migrants are perceived as more risk-seeking and entrepreneurial than the native-born

population, they have been shown to start a large number of new companies boosting economic activity in their destination areas. This has been the case, among others, in the UK, the US or Australia. Migrant entrepreneurs in these countries have on average founded more firms than locals (i.e. Borjas, 1986; Lofstrom, 2002; Wadhwa *et al.*, 2007; Clark and Drinkwater, 2010; Fairlie, Zissimopoulos and Krashinsky, 2010). They are thus vital for both new job creation as well as for the emergence of business start-ups. They bring vibrancy, diversity, and economic dynamism, enriching the neighbourhoods and benefitting the local population (Sahin, Nijkamp and Rietdijk, 2009; Audretsch, Dohse and Niebuhr, 2010). Various empirical analyses support these views characterising the migrants as "a highly motivated and qualified entrepreneurial group" (Brunow, Nijkamp and Poot, 2015: 1065) who substantially contribute to knowledge formation, technological progress, business income, and employment (see Fairlie and Lofstrom, 2015).

A diverse compositional structure of the migrant inflow is not solely regarded as growth enhancing element. Diversity in the migrant composition can also be seen as a destabilising factor within a society enhancing the potential for polarisation in its population leading to social unrest and conflict. Migrant inflows can result, under certain circumstances, to the formation of cultural, religious, or language barriers. This can generate tension, communication problems, and lower trust, which, in turn, lead to decreasing productivity and lower efficiency (i.e. Easterly and Levine, 1997; Alesina *et al.*, 2003; Montalvo and Reynal-Querol, 2005; Özgen, Nijkamp and Poot, 2013; Churchill and Smyth, 2017).

Closely connected to migrant diversity is the gender composition of the migrant inflow – an aspect of the economic impact of migration on the host country which has been analysed to a far smaller extent than cultural or ethnic diversity. Past research has

shown that women migrants indeed differ from their male counterparts. Not only individual characteristics of those who decide to migrate differ along gender lines, the settlement patterns, the impact on sending regions as well as the effect on host labour markets was shown to be significantly different for both men and women (i.e. Pessar, 1986; Hondagneu-Sotelo, 1994; Massey;, 2001; Pessar and Mahler, 2003; Oishi, 2005; Andall, 2013).

Female migration may influence economic development in a number of (sometimes indirect) ways, such as increasing the country's labour force (i.e. Lechman and Kaur, 2015; Cuberes and Teignier, 2016), improving gender equality (i.e. Berik, Rodgers and Seguino, 2009; Klasen and Lamanna, 2009), or via the empowerment of women (i.e. Duflo, 2012). Nevertheless, we still know very little about the specific direct effects of large numbers of female migrants on regional growth in the receiving region. The few studies analysing the gender effects of migration reach diverging results: Some studies hint at a positive economic impact of a strong presence of women migrants in the labour force (see Blau, Kahn and Moriarty, 2003 for the case of the US) and highlight their contribution to entrepreneurial activity (see Collins and Low, 2010 for the case of Australia). Others, however, have pointed in the opposite direction. Female migration has been linked to negative economic outcomes when evaluating gender gaps in the labour market participation of natives and foreign-born population in the UK (Smith and Bailey, 2006). Riaño and Baghdadi (2007) link a potential negative economic contribution of female migrants to a poorer assimilation into local labour markets than men and provide evidence of their underused economic potential. According to their findings, most female migrants, when entering the receiving country's labour market, seem to end up in jobs well below their actual skill-level.

Lastly, when evaluating the composition of migrant stocks and its impact on economic growth in the host region, scientific research strongly differentiates migrants according to their birthplaces. The biggest division of inward migration by origin is between external (foreign-born) and internal (native-born) migrants. Whether the migrant crosses an international border or not puts her/him in one category or the other. Both groups, however, have not attracted the same attention. The analysis of international migrants has dominated the social sciences – in fact all of the aforementioned studies focus on the economic impact of this specific group. Studies on the economic implications of internal migration are, by contrast, few and far between (Skeldon, 2006; Ellis, 2012; Bell *et al.*, 2015). Thus, existing knowledge of the macroeconomic consequences of internal migration from one to another place within the boundaries of the same country are based on a handful of studies (White and Lindstrom, 2005). Some find that internal migration lifts incomes, fosters modernization, and decreases inequality, making it a growth-enhancing factor (e.g. Yap, 1976; Berker, 2011; Kuhn, 2015). Conversely, other research partly blames regional divergence, poverty rates, and declining quality and endowments of public goods on internal migration (i.e. Rodríguez Vignoli, 2008; Molloy, Smith and Wozniak, 2011). Some simply do not find any significant connection between internal migration and economic growth (i.e. White and Lindstrom, 2005).

In short, over the last decades, migration and its implications for economic growth in the host or receiving territory has grown to a wide and intensively analysed field in economic research. There is certainly no shortage of studies explaining its implications for the local labour market, public finances, innovation or entrepreneurial activity. Many insights have also been put forward about the economic effects of different compositional structures of the migrant population, be it diversity, gender, or

geographical origin. It is important, however, to stress, once again, that the overwhelming majority of studies only take into account the short- or medium-run consequences of migration. Beyond a timeframe of a maximum of 25 years, we simply know nothing about what long-term economic imprint, if at all, may be associated to past migration.

3.1 REFERENCES

Ablett, J. (1999): Generational accounting in Australia. In A. Auerbach (Ed.), *Generational Accounting around the World*, 141-160. Chicago: University of Chicago Press.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003): Fractionalization. *Journal of Economic Growth*, *8*(2), 155–194.

Alesina, A., Harnoss, J., and Rapoport, H. (2016): Birthplace diversity and economic prosperity. *Journal of Economic Growth*, *21*(2), 101–138.

Altonji, J. G., and Card, D. (1991). The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives. In J. M. Abowd and R. B. Freeman (Eds.), *Immigration, Trade, and the Labor Market*, 201–234. University of Chicago Press.

Andall, J. (2013): Gendered Mobilities and Work in Europe: An Introduction. *Journal of Ethnic and Migration Studies*, *39*(4), 525–534.

Audretsch, D., Dohse, D., and Niebuhr, A. (2010): Cultural diversity and entrepreneurship: a regional analysis for Germany. *The Annals of Regional Science*, *45*(1), 55–85.

Baker, M., and Benjamin, D. (1995): The Receipt of Transfer Payments by Immigrants to Canada. *Journal of Human Resources*, *30*(4), 650–676.

Barrett, A., and McCarthy, Y. (2008). Immigrants and welfare programmes: exploring the interactions between immigrant characteristics, immigrant welfare dependence, and welfare policy. *Oxford Review of Economic Policy*, *24*(3), 543–560.

Bell, M., Charles-Edwards, E., Kupiszewska, D., Kupiszewski, M., Stillwell, J., and Zhu, Y. (2015): Internal Migration Data Around the World : Assessing contemporary practice. *Population, Space and Place*, *21*(1), 1–17.

Berik, G., Rodgers, Y., and Seguino, S. (2009): Feminist Economics of Inequality, Development, and Growth. *Feminist Economics*, *15*(3), 1–33.

Berker, A. (2011): Labor-Market Consequences of Internal Migration in Turkey. *Econonomic Development and Cultural Change*, *60*(1), 197–239.

Blau, F., and Kahn, L. (2015): Chapter 15 - Immigration and the Distribution of Incomes. In B. R. Chiswick and P. W. Miller (Eds.), *Handbook of the Economics of International Migration*, Vol. 1, 793–843. North-Holland.

Blau, F., Kahn, L., and Moriarty, J. (2003): The role of the family in immigrants' labor-market activity: An evaluation of alternative explanations: Comment. *The American Economic Review*, *93*(1), 429–447.

Blume, K., and Verner, M. (2007): Welfare dependency among Danish immigrants. *European Journal of Political Economy*, *23*(2), 453–471.

Bonin, H., Raffelhüschen, B., and Walliser, J. (2000): Can immigration alleviate the demographic burden? *FinanzArchiv / Public Finance Analysis*, *57*(1), 1–21.

Borjas, G. J. (1986): The self-employment experience of immigrants. *Journal of Human Resources*, *21*(4), 485–506.

Borjas, G. J. (2003): The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, *118*(4), 1335–1374.

Borjas, G. J., Freeman, R. B., Katz, L. F., Dinardo, J., and Abowd, J. M. (1997): How much do immigration and trade affect labor market outcomes? *Brookings Papers on Economic Activity*, *1997*(1), 1–90.

Bove, V., and Elia, L. (2017): Migration, diversity and economic growth. *World Development*, *89*, 227–239.

Brunow, S., Nijkamp, P., and Poot, J. (2015): Chapter 19 - The Impact of International Migration on Economic Growth in the Global Economy. In B. R. Chiswick and P. W. Miller (Eds.), *Handbook of the Economics of International Migration*, Vol. 1, 1027–1075. North-Holland.

Büchel, F., and Frick, J. (2005): Immigrants' economic performance across Europe – does immigration policy matter? *Population Research and Policy Review*, *24*(2), 175–212.

Card, D. (1990): The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial & Labor Relations Review*, *43*(2), 245–257.

Card, D. (2001): Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, *19*(1), 22–64.

Card, D. (2005): Is the new immigration really so bad? *Economic Journal*, *115*(507), F300–F323.

Card, D. (2009): Immigration and inequality. *The American Economic Review*, *99*(2), 1–21.

Cattaneo, C., Fiorio, C. V, and Peri, G. (2015): What Happens to the Careers of European Workers When Immigrants "Take Their Jobs"? *Journal of Human Resources*, *50*(3), 655–693.

Cerrutti, M., and Massey, D. (2001): On the Auspices of Female Migration from Mexico to the United States. *Demography*, *38*(2), 187–200.

Churchill, S. A., and Smyth, R. (2017): Ethnic diversity and poverty. *World Development*, *95*, 285–302.

Clark, K., and Drinkwater, S. (2010): Patterns of ethnic self-employment in time and space: evidence from British Census microdata. *Small Business Economics*, *34*(3), 323–338.

Clemens, M. A. (2011): Economics and Emigration: Trillion-Dollar Bills on the Sidewalk? *Journal of Economic Perspectives*, *25*(3), 83–106.

Cohen-Goldner, S., and Paserman, M. D. (2011): The dynamic impact of immigration on natives' labor market outcomes: Evidence from Israel. *European Economic Review*, *55*(8), 1027–1045.

Collins, J., and Low, A. (2010): Asian female immigrant entrepreneurs in small and medium-sized businesses in Australia. *Entrepreneurship & Regional Development*, *22*(1), 97–111.

Cortés, P., and Tessada, J. (2011): Low-Skilled Immigration and the Labor Supply of Highly Skilled Women. *American Economic Journal: Applied Economics*, *3*(3), 88–123.

Crossley, T. F., McDonald, J. T., and Worswick, C. (2001): Immigrant Benefit Receipt Revisited. *Journal of Human Resources*, *36*(2), 379–397.

Cuberes, D., and Teignier, M. (2016): Aggregate Effects of Gender Gaps in the Labor Market: A Quantitative Estimate. *Journal of Human Capital*, *10*(1), 1–32.

Di Giovanni, J., Levchenko, A., and Ortega, F. (2015): A global view of cross-border migration. *Journal of the European Economic Association*, *13*(1), 168-202.

Duflo, E. (2012): Women Empowerment and Economic Development. *Journal of Economic Literature*, *50*(4), 1051-1079.

Dustmann, C., Fabbri, F., and Preston, I. (2005): The Impact of Immigration on the British Labour Market. *Economic Journal*, *115*(507), F324–F341.

Dustmann, C., Frattini, T., and Preston, I. P. (2013): The Effect of Immigration along the Distribution of Wages. *The Review of Economic Studies*, *80*(1), 145–173.

Easterly, W., and Levine, R. (1997):. Africa's Growth Tragedy: Policies and Ethnic Divisions. *The Quarterly Journal of Economics*, *112*(4), 1203–1250.

Ellis, M. (2012): Reinventing US Internal Migration Studies in the Age of International Migration. *Population, Space and Place*, *18*(2), 196–208.

Fairlie, R. W., and Lofstrom, M. (2015):. Chapter 17 - Immigration and Entrepreneurship. In B. R. Chiswick and P. W. Miller (Eds.), *Handbook of the Economics of International Migration*, Vol. 1, 877–911. North-Holland.

Fairlie, R. W., Zissimopoulos, J., and Krashinsky, H. (2010): The International Asian Business Success Story? A Comparison of Chinese, Indian and Other Asian Businesses in the United States, Canada and United Kingdom. In J. Lerner and A. Shoar (Eds.), *International Differences in Entrepreneurship*, 179–208. University of Chicago Press and National Bureau of Economic Research.

Florida, R. L. (2002): *The rise of the creative class: and how it's transforming work, leisure, community and everyday life*. New York, NY: Basic Books.

Friedberg, R., and Hunt, J. (1995): The Impact of Immigrants on Host Country Wages, Employment and Growth. *The Journal of Economic Perspectives*, *9*(2), 23–44.

Furtado, D., and Hock, H. (2010): Low Skilled Immigration and Work-Fertility Tradeoffs among High Skilled US Natives. *The American Economic Review*, *100*(2), 224–228.

Glitz, A. (2012): The Labor Market Impact of Immigration: A Quasi-Experiment Exploiting Immigrant Location Rules in Germany. *Journal of Labor Economics*, *30*(1), 175–213.

Gustman, A., and Steinmeier, T. (2000): Social security benefits of immigrants and the U.S. born. In G. Borjas (Ed.), *Issues in the Economics of Immigration*, 309–350. Chicago: University of Chicago Press.

Hirschman, C., and Mogford, E. (2009). Immigration and the American industrial revolution from 1880 to 1920. *Social Science Research*, *38*(4), 897–920.

Hondagneu-Sotelo, P. (1994): *Gendered Transitions: Mexican Experiences of Immigration*. Berkeley: University of California Press.

Hu, W.-Y. (1998): Elderly Immigrants on Welfare. *Journal of Human Resources*, *33*(3), 711–741.

Hunt, J., and Gauthier-Loiselle, M. (2010): How Much Does Immigration Boost Innovation? *American Economic Journal: Macroeconomics*, *2*(2), 31–56.

Jacobs, J. (1961): *The death and life of great American cities*. New York, NY: Random House.

Jacobs, J. (1969): *The economy of cities*. New York, NY: Random House.

Kennan, J. (2013): Open borders. *Review of Economic Dynamics*, *16*(2), L1–L13.

Kerr, S., and Kerr, W. R. (2011): Economic Impacts of Immigration: A Survey. *NBER Working Paper Series*, No. 16736, Cambridge, Mass., USA.

Klasen, S., and Lamanna, F. (2009): The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries. *Feminist Economics*, *15*(3), 91–132.

Kuhn, R. (2015): Internal Migration: Developing Countries. In *International Encyclopedia of the Social & Behavioral Sciences*, 433–442. Amsterdam, Netherlands: Elsevier.

LaLonde, R. J., and Topel, R. H. (1991): Labor market adjustments to increased immigration. In J. M. Abowd and R. B. Freeman (Eds.), *Immigration, Trade, and the Labor Market*, 167-199. Chicago: University of Chicago Press.

Lechman, E., and Kaur, H. (2015): Economic growth and female labour force participation – verifying the u-feminization hypothesis. New evidence for 162 countries over the period 1990-2012. *Economics and Sociology*, *8*(1), 246–257.

Lee, N., and Nathan, M. (2010): Knowledge workers, cultural diversity and innovation: evidence from London. *International Journal of Knowledge-Based Development*, *1*(1/2), 53–78.

Lofstrom, M. (2002): Labor market assimilation and the self-employment decision of immigrant entrepreneurs. *Journal of Population Economics*, *15*(1), 83–114.

Longhi, S., Nijkamp, P., and Poot, J. (2005): A Meta-Analytic Assessment of the Effect of

Immigration on Wages. *Journal of Economic Surveys*, *19*(3), 451–477.

Longhi, S., Nijkamp, P., and Poot, J. (2008): Meta-analysis of empirical evidence on the labour market impacts of immigration. *Région et Developpement*, *27*, 161–191.

Molloy, R., Smith, C. L., and Wozniak, A. (2011): Internal Migration in the United States. *Journal of Economic Perspectives*, *25*(3), 173–196.

Montalvo, J., and Reynal-Querol, M. (2005): Ethnic Polarization, Potential Conflict, and Civil Wars. *The American Economic Review*, *95*(3), 796–816.

Moscarola, F. (2003): Immigration flows and the sustainability of the Italian welfare state. *Politica Economica*, *1*, 63–90.

Oishi, N. (2005): *Women in Motion*. Stanford: Stanford University Press.

Okkerse, L. (2008): How to measure labour market effects of immigration: A review. *Journal of Economic Surveys*, *22*(1), 1–30.

Ortega, F., and Peri, G. (2009): The Causes and Effects of International Migrations: Evidence From OECD Countries 1980-2005. *NBER Working Paper*, No. 14833.

Ottaviano, G. I. P., and Peri, G. (2012): Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, *10*(1), 152–197.

Özgen, C., Nijkamp, P., and Poot, J. (2010): The effect of migration on income growth and convergence: Meta-analytic evidence. *Papers in Regional Science*, *89*(3), 537–561.

Özgen, C., Nijkamp, P., and Poot, J. (2012): Immigration and innovation in European regions. In P. Nijkamp, J. Poot, and M. Sahin (Eds.), *Migration Impact Assessment: New Horizons*. Cheltenham, UK: Edward Elgar Publishing.

Özgen, C., Nijkamp, P., and Poot, J. (2013): The impact of cultural diversity on firm innovation: evidence from Dutch micro-data. *IZA Journal of Migration*, *2*(1), 1–24.Partridge, J., & Furtan, H. (2008). *Increasing Canada's International Competitiveness: Is There a Link between Skilled Immigrants and Innovation? American Agricultural Economics Association Annual Meeting*. Retrieved from https://core.ac.uk/download/pdf/6563849.pdf

Pedersen, P. J. (2013): Immigration and welfare state cash benefits: the Danish case. *International Journal of Manpower*, *34*(2), 113–125.

Peri, G., and Sparber, C. (2009): Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, *1*(3), 135–69.

Pessar, P. R. (1986): The Role of Gender in Dominican Settlement in the United States. In J. Nash and H. Safa (Eds.), *Women and Change in Latin America,* 273–294. South Hadley, MA: Bergin and Garvey.

Pessar, P. R., and Mahler, S. J. (2003): Transnational Migration: Bringing Gender In. *International Migration Review*, *37*(3), 812–846.

Riaño, Y., and Baghdadi, N. (2007): Understanding the Labour Market Participation of Skilled Immigrant Women in Switzerland: The Interplay of Class, Ethnicity, and Gender. *Journal of International Migration and Integration / Revue de l'Integration et de La Migration Internationale*, *8*(2), 163–183.

Rodríguez Vignoli, J. (2008): Spatial distribution, internal migration and development in Latin America and the Caribbean. *Cepal Review*, 96, 137–157.

Roodenburg, H., Euwals, R., and TerRele, H. (2003): *Immigration and the Dutch economy*. CPB Netherlands Bureau for Economic Policy Analysis, The Hague: Netherlands.

Rowthorn, R. (2008): The fiscal impact of immigration on the advanced economies. *Oxford Review of Economic Policy*, *24*(3), 561–581.

Sahin, M., Nijkamp, P., and Rietdijk, M. (2009): Cultural diversity and urban innovativeness: personal and business characteristics of urban migrant entrepreneurs. *Innovation: The European Journal of Social Science Research*, *22*(3), 251–281.

Sinn, H., and Werding, M. (2001). *Îmmigration following the EU eastern enlargement*. CESifo Forum, *2*(2), 40-47.

Skeldon, R. (2006): Interlinkages between internal and international migration and development in the Asian region. *Population, Space and Place*, *12*(1), 15–30.

Smith, D. P., and Bailey, A. J. (2006): International Family Migration and Differential Labour- Market Participation in Great Britain: Is There a "Gender Gap"? *Environment and Planning A*, *38*(7), 1327–1343.

Storesletten, K. (2000): Sustaining Fiscal Policy through Immigration. *Journal of Political Economy*, *108*(2), 300–323.

Storesletten, K. (2003): Fiscal Implications of Immigration - A Net Present Value Calculation. *Scandinavian Journal of Economics*, *105*(3), 487–506.

Wadhwa, V., Saxenian, A., Rissing, B., and Gereff, G. (2007). *America's new immigrant entrepreneurs*. Durham, NC: Duke University.

White, M. J., and Lindstrom, D. P. (2005): Chapter 11: Internal Migration. In *Handbook of Population*, 311–346. Springer Science and Business Media B.V. / Books.

Yap, L. (1976): Internal Migration and Economic Development in Brazil. *The Quarterly Journal of Economics*, *90*(1), 119–137.

4 MIGRATION TO THE US

"Give me your tired, your poor, Your huddled masses yearning to breathe free, The wretched refuse of your teeming shore. Send these, the nameless, tempest-tost to me, I lift my lamp beside the golden door!"

~ Emma Lazarus, 1883

These words, written by Emma Lazarus are engraved at the foot of the Statue of Liberty in the harbour of New York City. No words could better describe this decisive element of American identity – the United States as a nation founded and formed by immigrants. The contribution of migrants is implicitly reflected in the nations motto: '*e pluribus unum'* (= 'from many to one'). Migration is thus at the very core of America's values, shaping large parts of its history (Martin and Midgley, 1999).

Since colonial times, the US has been the prime destination for people all over the world wanting to build a better future for themselves and their families, have a second chance, or simply live 'the American dream'. Back in the 17th and 18th century, thousands of migrants of mostly European origin, largely from the British Isles, Germany, France, the Netherlands, or Scandinavia, ventured across the Atlantic. Today it is millions of mostly Latin Americans and Asians who cross the border in search of a new, different, and better life. Since the day of the first settlements to the present, the US have always remained the number one destination of international migrant flows (Migration Policy Institute, 2015).

The United States' attractiveness to migrants is reflected in the sheer size of its foreign-born population. Over the past 160 years, the absolute number of US residents with a non-US birthplace increased from two to around 40 million in 2010 transforming the country into a 'nation of immigrants' (see Figure 4-1). Two main

eras of immigration can be distinguished in US history: The century of mass migration to the US with a peak in immigration numbers around the turn of the 19th to the 20th century and contemporary migration starting from 1970 onwards. In both these periods of time, absolute migration numbers increased year on year, growth rates of the immigrant population reached unprecedented levels and the size of the foreign-born share of the total US population reached close to 15 percent. For these reasons, both of these eras will be in the focus of this work.

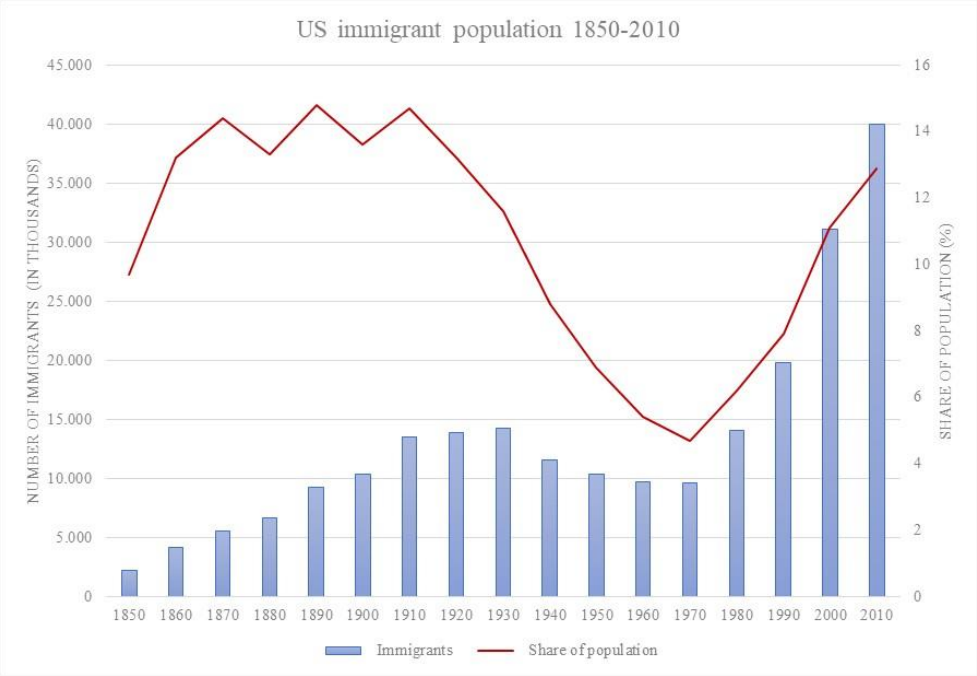


FIGURE 4-1 US MIGRANT POPULATION (ABSOLUTE AND SHARE OF TOTAL POPULATION), 1850-2010

Source: Migration Policy Institute (2017)

4.1 THE AGE OF MASS MIGRATION TO THE US

In colonial times – comprising the period between the beginning of the European settlement on the American continent around 1600 and the declaration of independence in 1776 – only a comparatively small absolute number of migrants made their way across the Atlantic Ocean. Starting from 1820 onwards, the "century of immigration" (Daniels, 1990: 117) began. In times of revolutionary movements, religious persecution, war, famine, draught, and overpopulation on the European continent, millions of people left their homes in search for a better life elsewhere. Leaving behind low living standards, low wages, and huge competition for work, they hoped for economic opportunity as well as for religious and political freedom in the US (Hatton and Williamson, 1994; Alexander, 2007).

Between 1850 and WWI alone, more than 40 million people from Europe moved to the 'new world', 60 percent of whom settled in the US (Bertocchi and Strozzi, 2006). Table 4-1 displays the entry figures of immigrants on a decade-by-decade basis from 1820 to 1924. Entry figures accelerated quickly from the beginning of the century until 1841 and remained in the millions thereafter. Two peaks can be detected: 1881-1890 and 1901-1910. Each of these eras marks a height of the two main migration waves during that period.

TABLE 4-1 MIGRATION TO THE US, 1820-1924

Decade	Number of immigrants entering the US	Number of Immigrants per thousand population
1820-30	151.824	1,2
1831-40	599.125	3,9
1841-50	1.713.251	8,4
1851-60	2.598.214	9,3
1861-70	2.314.824	6,4
1871-80	2.812.191	6,2
1881-90	5.246.613	9,2
1891-1900	3.687.564	5,3
1901-10	8.795.386	10,4
1911-20	5.735.811	5,7
1921-24	2.344.599	5,3
Total	35.999.402	

Source: Daniels (1990)

The first wave of migrants, often referred to as "the pioneers of the century of immigration" (Daniels, 1990: 121) is dated between the pre-Civil War years and 1890. During these decades, main source countries of incoming migrants were in North-Western Europe (the United Kingdom, Ireland, Scandinavia, or the German states) (Hatton and Williamson, 1994). 85 percent of the foreign-born population listed in the US Census of 1870, stated these countries as their places of origin (Alexander, 2007). Immigrants from the Mediterranean as well as territories within former Austria-Hungary, Poland, and Russia dominated the second wave between 1890 and 1920. Rising economic development levels in combination with decelerating population growth in North-West Europe had altered incentive structures resulting in lower emigration rates from these regions. The South and East of Europe, however, now experienced the "disruptive impact of industrialization on rural societies, which had afflicted North-Western Europe earlier in the century" (Ward, 1990: 308). Thus, shortly before the turn of the century in 1896, Italians, Poles, Russians, Greeks, and

Portuguese already outnumbered first-wave migrants (Rodríguez-Pose and von Berlepsch, 2015).[5] By 1924, the share of second-wave migrants had reached nearly 80 percent of the total migrant inflow (Ward, 1990).

Migrants from Europe usually entered the US via the ports of the East coast.[6] New York – and, in particular, Ellis Island – was by far the most common entry point for both waves of migration. Other coastal cities also became major gateways for incoming population. Baltimore (especially for German migrants), Boston (dominated by the Irish), Philadelphia (used by British, Germans, Poles, and Russians), or New Orleans (mainly Italians and Greeks) served as important entry sites (Alexander, 2007). After setting foot on US soil and having successfully passed through immigration processes, migrants either stayed in the port cities or moved farther inland. Main means of transportation were the railroad, waterways, carriages or, for the large poor parts of the population, their own feet. Migrants would travel hundreds of kilometres to join their families or relatives who had previously migrated and had already made a life for themselves. A distinct settlement pattern evolved. Figure 4-2 displays the migrants' settlement across the 48 continental states at the peak of the first wave of migration in 1880.

[5] Already during these years, migration from Latin America and Asia had commenced, settling predominantly in the west and southwest of the country. As both of these migrant streams taken together did not surpass 10 percent of total migration during these times (Ward, 1990), I refrain from discussing these population flows in this section.

[6] Asian and Latin American migrants disembarked at western ports (such as San Francisco) or entered the country via the southern border with Mexico (Alexander, 2007).

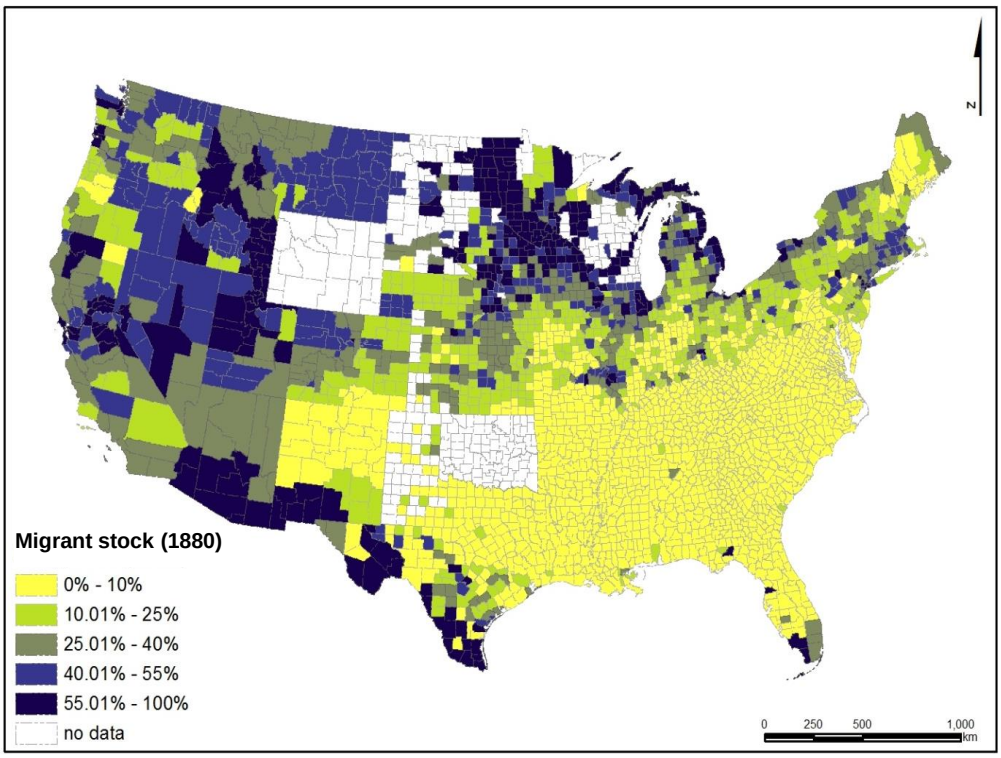


FIGURE 4-2 MIGRANT STOCK BY COUNTY, 1880

Source: Rodríguez-Pose and von Berlepsch (2014)

Migrants settled predominantly in the North and West of the country (Figure 4-2). Large groups remained close to their ports of entry and settled in New York, New Jersey, and other parts of New England. Others followed old migrant routes farther inland. Foreign-born population flocked to the Great Lake states, such as Michigan, Illinois, or Wisconsin, or put down roots in rural counties in Iowa, Minnesota, and the Dakotas. Missouri, Kansas, and Nebraska were also popular destinations, as was the case of the mountain states farther to the West, including Montana, Idaho, Colorado or Utah and of south-western states, such as Arizona, Nevada, and parts of California. The Old South, by contrast, remained almost untouched by the foreign-born population. With the exception of the southern tip of Florida and parts of Texas, the South with its "commercial agricultural system largely based upon the intensive use

of black labour and a decidedly slow rate of industrialization offered few attractions to foreign immigrants" (Ward, 1990: 301).

The big majority of the first wave migrant groups settled in rural areas. Early German migrants, for example, wanted to toil the land and clustered in rural areas of New York, the Great Lakes, and the Midwest, where land was still available at the time. Swedes, Danes, and Norwegians had similar aspirations and colonised the upper wheat belt – Minnesota, the Dakotas, Wisconsin and Nebraska. The Irish, by contrast, predominantly stayed in the port cities where they had previously arrived. More than 20 percent of the urban population in New England, New Jersey, and New York state declared Irish roots at the time (Daniels, 1990; Ward, 1990; Rodríguez-Pose and von Berlepsch, 2015).

The settlement pattern associated with the second wave of migration reproduces, to a large extent, the map of 1880 (Figure 4-3). Once again, the main destination states were in New England, the Mid-Atlantic, around the Great Lakes, the northern Mid-West, Mountain, and Western states. With the exceptions of Texas and the southern tip of Florida, Southern states, yet again, remained out of bounds for incoming migrants. Despite different national origins, the second wave thus followed in the footsteps of previous migrants, using established migrant routes.

However, differences in the settlement pattern between first and second wave can be detected on a more granular level. Migrant groups from Southern and Eastern European countries tended to settle in more highly urbanized areas than northern Europeans did decades earlier. 88 percent of all second-wave migrants at the time lived in cities rather than in the countryside (Daniels, 1990). New York, Boston, Philadelphia, and Providence had massive Italian communities. The Greeks stayed in

the cities of New England and in areas surrounding Chicago, while the Portuguese and Spaniards settled in urban areas of California and Nevada. The majority of Poles and Russians tended to move to cities in the rust belt such as Chicago, Pittsburgh, or Buffalo, while significant eastern European communities also concentrated in rural areas of the Midwest, Texas, Iowa, and the Great Lakes regions (Rodríguez-Pose and von Berlepsch, 2015).

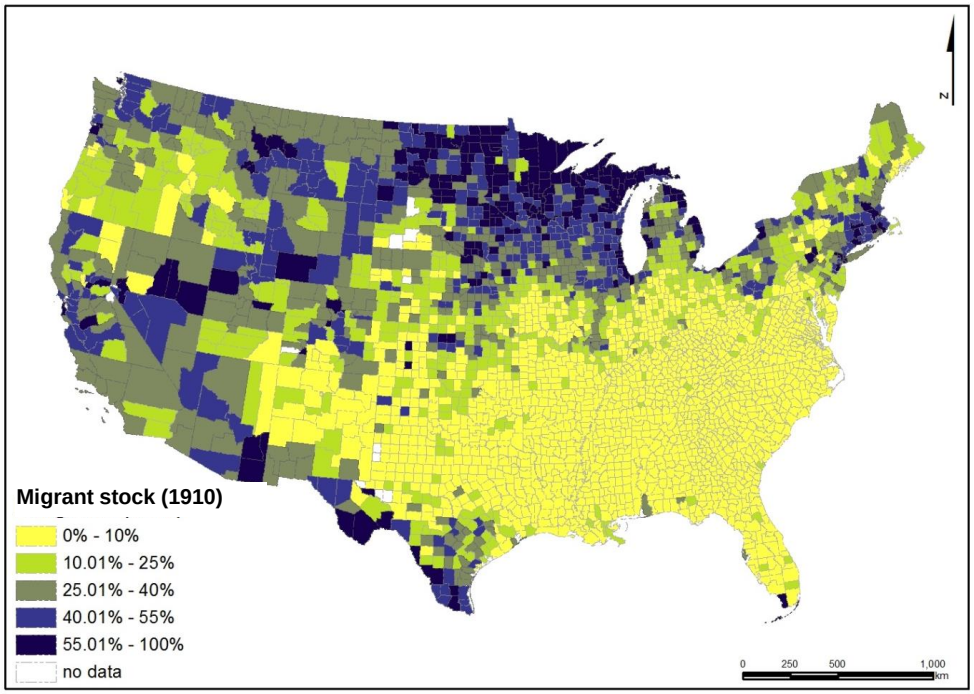


FIGURE 4-3 MIGRANT STOCK BY COUNTY, 1910

Source: Rodríguez-Pose and von Berlepsch (2014)

The typical European migrant around the turn of the 19th to the 20th century was, independent of nationality, a young male adult (see for example Fitzpatrick, 1984 for an analysis of the Irish; Bodnar, 1992 for Swedish migrants). Alexander (2007) reports that 80 percent of the foreign-born population was between fourteen and forty-four,

with a majority in their twenties and thirties. Families and married men travelling on their own made up only small parts of the migrant streams at the time. The large bulk of newcomers were single men. On average, only one in three immigrants entering the US during these decades was a woman, although a gradually shifting gender balance across time can be observed. Around the turn of the century, roughly 30 percent of European migrants were female. By 1915 the share of women had increased to almost 43 percent. Despite the changing gender pattern over time, the ratio of male to female incoming migrants highly depended on the source country. Men clearly dominated the migrant contingents from Italy, Greece, Croatia, Poland, and other Eastern European countries. More than 70 percent of these migrant groups were men. Irish and Swedish entrants were more gender-balanced and oftentimes dominated by female migrants. There were more women than men entering the US in 13 years between 1880 and 1920, in the case of Irish migrants, and in two years, in the case of Swedish migrants (Alexander, 2007).

Generally, outward migration from Europe affected every profession and all classes of people. However, concentration was highest in the middle and lower-middle ranks of society (Baines, 1994). The large majority of migrants was unskilled, showing – if at all – only rudiments of an education. Migrants were also generally unfamiliar with the English language apart from those originating from the British Isles (Daniels, 1990). Hard physical work as common labourers or servants in the cities, as agricultural hands on the fields or as workers in the factories became their day to day, placing them at the bottom of the economic structure. Some immigrant groups, however, brought substantial skills and therefore managed to quickly make their way up the social ladder. Skilled labourers were much more common among the north-western Europeans (especially the Germans and Swedes) than among the south-

eastern groups. Many northern and western Europeans ended up working in skilled trades, including large numbers of bakers, butchers, distillers, machinists, technicians, tailors, or carpenters (Daniels, 1990). The English in particular did not only have usually higher skill levels than the average migrant, but spoke the language, knew local institutions, and were accustomed to the system, paving their way towards the higher ranks of society (Abramitzky, Boustan and Eriksson, 2014).

An overwhelming majority of migrants originated from the rural and proto-industrial hinterlands of Europe (Kamphoefner, 1976). In rural regions of Ireland, Sweden, Poland, or Germany, for example, high birth rates coupled with crop failures, exhaustion of the soil and, in many cases, a system of partible inheritance, created a young generation of landless farm labourers who were forced to seek employment elsewhere (Fitzpatrick, 1984; Guinnane, 1992; Hatton and Williamson, 1993). In southern Europe, industrialisation hit later, causing regional divergence and out-migration. In Italy, for example, the north had started to industrialise undergoing a wave of mechanisation of its agricultural sector. Rising productivity levels in the North aggravated competition for the premodern peasant agriculture of the southernmost provinces. Coupled with taxation policies favouring the industrial north, the south fell further behind and masses of Southern Italians steered their fate towards the Americas (Spickard, 2007).

Many followed in the footsteps of predecessors such as family, friends, or acquaintances from back home who had previously undertaken the journey and were already settled in the US. Oftentimes, pre-paid tickets were purchased by relatives and sent home for the next of kin to join them in the ‘new world’. Once arrived, newcomers followed their country folk relying heavily on their assistance in the housing and job search as well as the assimilation process in the new community (Hatton and

Williamson, 1994). This path-dependence in migrant settlement coupled with a consistent migrant inflow across decades ensured the creation of significant migrant communities across the United States often originating from similar ethnic, national or even local origins (Vedder and Gallaway, 1972; Levy and Wadycki, 1973; Dunlevy and Gemery, 1977; Rodríguez-Pose and von Berlepsch, 2015). "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). As a result, between autumn 1907 and summer 1910, for example, nearly 94% of all those arriving at US ports claimed to join either family or friends (Alexander, 2007).

In these ethnic or national enclaves across the country, migrants often rebuilt the societal structures they were familiar with back home (Ward, 1990). Bringing their institutional baggage in the form of culture, traditions, and customs along with them, migrants "reconstitute[d] entire communities with their institutions and associations […], kinship networks, religious organizations, pressure groups, political organs, […] ethnic press, ethnic banks and businesses" (Joly, 2000: 33). Migrants "came not to establish something new but to re-establish something old" (Daniels, 1990: 146). Thus, a heavy inflow of Germans in Wisconsin led to the creation of German newspapers, schools, breweries, and traditional music clubs; New York and New Jersey turned into southern Italian cities nearly over night, and Boston became the stronghold of Irish catholic communities (Daniels, 1990).

Ever growing masses of migrants arriving predominantly in the ports of the east coast provoked not only the rise of population density in eastern states, but went hand in hand with shortages in employment and housing, increasing crime rates, and deteriorating economic prospects for residents (Merk, 1978). These developments

pushed not only the settlement of the foreign-born population further inland but triggered mass-movements of the native-born[7]. News of discoveries of mineral resources, vast amounts of animal stock, cheap land prices, highly fertile soils, and gentler climates out west further incentivised large parts of the domestic population all across the country to pack up their belongings and seek their fortunes elsewhere. In search to raise their economic opportunities, acquire cheap, good-quality land, and earn higher wages in better jobs, Americans "from worn-out lands in the East [were drawn to] virgin lands in the West" (Merk, 1978: 229). The rapid redistribution of the domestic population transformed vast amounts of US territory. Wheat-farmers and corn-producers settled predominantly in the Mid-west turning it into the 'corn-kingdom' and 'bread-basket' of the US. Pork and cattle breeders settled in the southern plains, similarly to wool producers and meat packers. The mountain states attracted cattle farmers, trappers, hunters, and fur traders. Californian ports appealed to tradesmen, while the soil and climate of the state lured lumber and wool industries. In the late 19th and early 20th centuries almost two thirds of the male US-born population above the age of 30 had crossed county- or state lines in order to build a better life away from their birthplace (Ferrie, 2005). By 1880, more than 30 percent of the entire US population lived outside their state of origin (Ruggles *et al.*, 2015, own calculations).

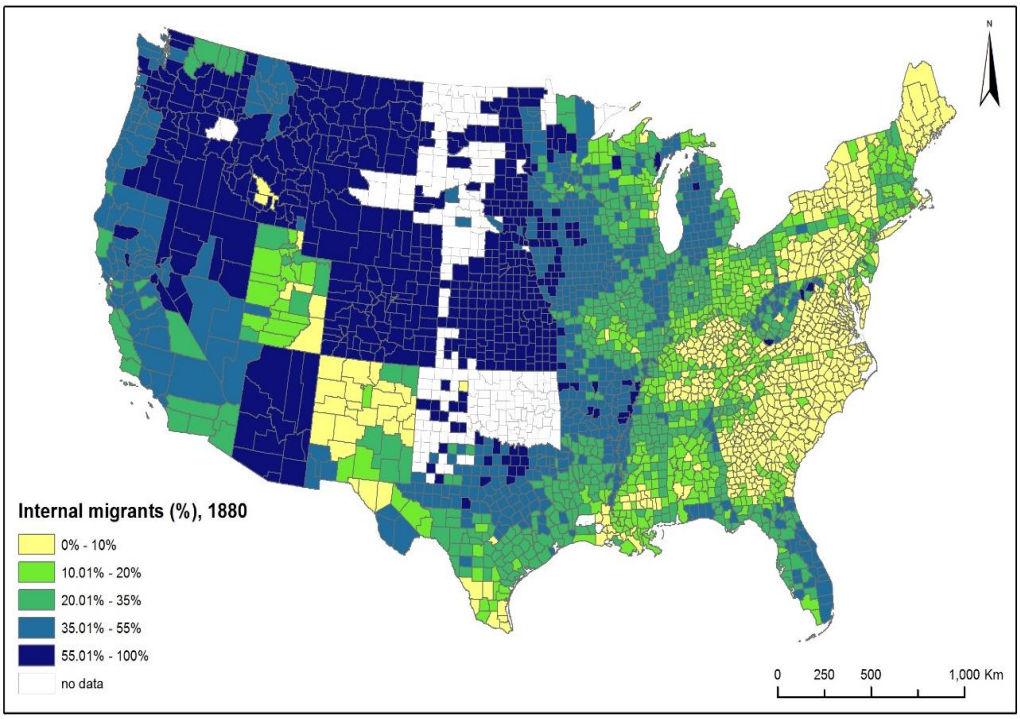
Figure 4-4 displays the settlement pattern of domestic migrants across the continental US territory, measured as share of a given county's total American-born population – (a) displays 1880, b) displays 1910. The drive to the West is clear: the lowest numbers of internal migrants were found mainly in the original thirteen states, with migrant

[7] The term native-born refers here to individuals born on American soil, not to the tribes of native populations.

shares rising along with distance to the Atlantic coast line. In 1880 (Figure 4-4a), more than 55% of the native-born migrants concentrated in states of the Mid-West and western Mountain regions.

In 1910 (Figure 4-4b), the drift towards the West becomes even more pronounced. The territory east of the Mississippi River hosted few internal migrants. The regions to the west of the river, however, displayed high shares within their population compositions. The average internal migrant in 19th century America often undertook multiple short distance moves from one state to the next, rarely staying in one location for long. The probability of an additional move increased manifold, once the uncertainty of the few first ones had been successfully mastered. In 1870 Trempleau County, Wisconsin, for instance, 68% of all farmers recorded in 1860 had left; 10 years later, only less than 20% remained (Atack, Bateman and Parker, 2000). This population movement is confirmed in Figure 4-4, which depicts the settlement pattern of internal migrants in both years. The "American population was a restless one, continually uprooting and moving to a new location […] 'every day was moving day'" (Atack, Jeremy and Passell, 1994: 237).

a)



b)

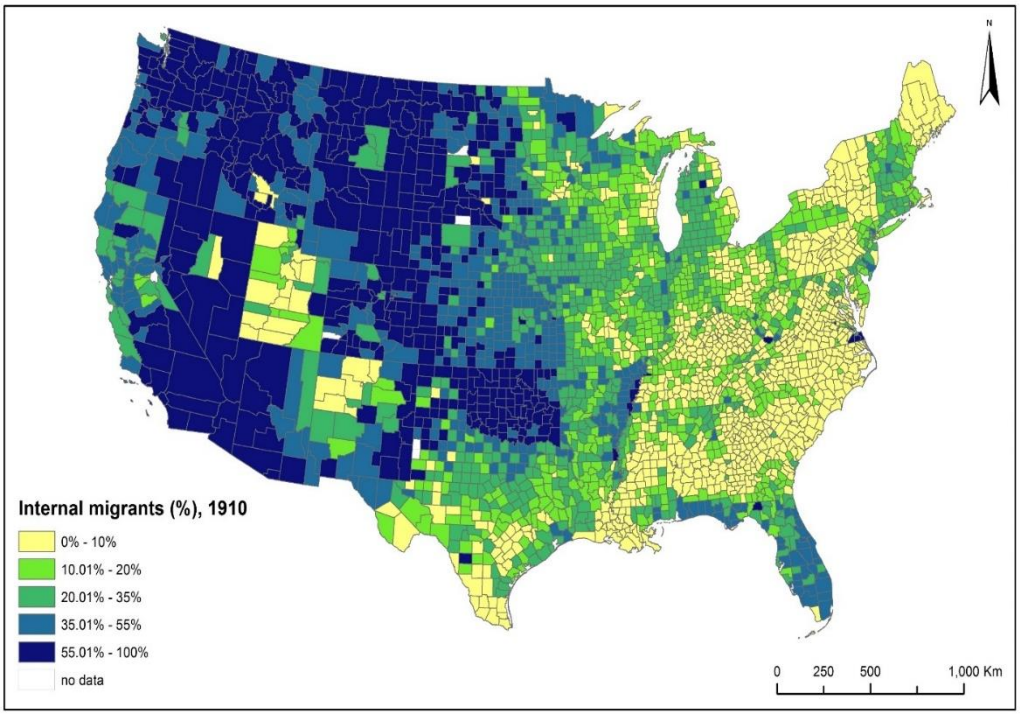


FIGURE 4-4 AMERICAN-BORN INTERNAL MIGRANTS BY COUNTY, 1880 AND 1910

(% OF TOTAL AMERICAN-BORN POPULATION)

Source: Ruggles et al., (2015), own elaboration

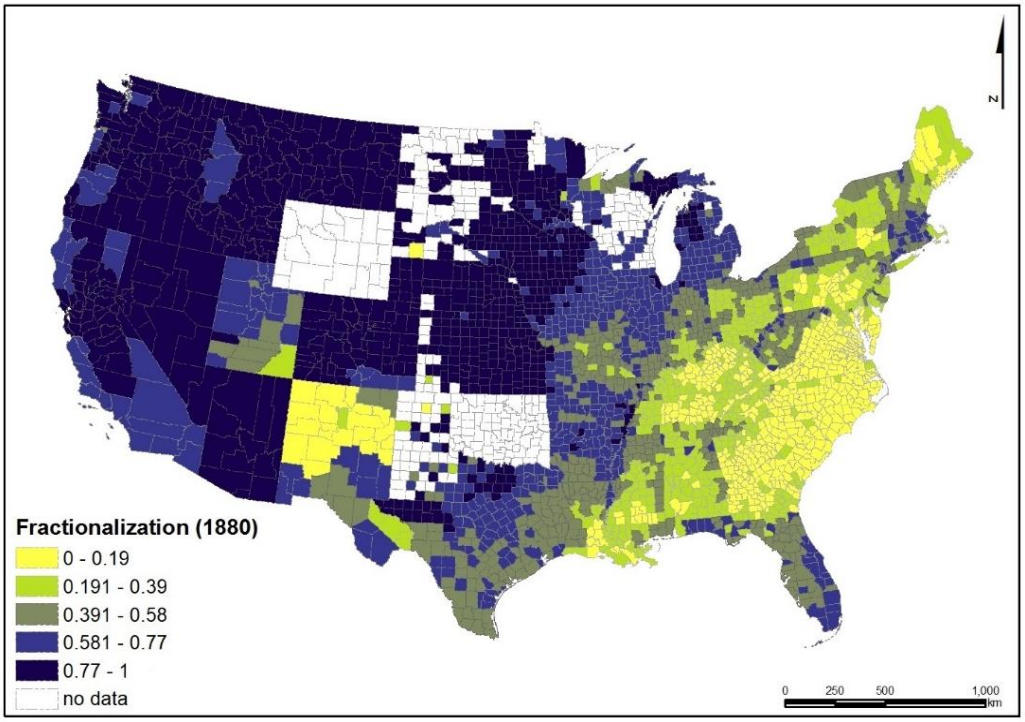
Both trends – the mobility of American-born US residents and the consistently high numbers of international migrants arriving in US ports – produced considerable changes in the local population composition across most US counties. Only the states of the 'Old South' remained an exception to this rule. Some areas were mainly settled by migrants of a particular national origin. The 'German Triangle' for example, between Saint Louis, Cincinnati, and Milwaukee evolved from having an almost entirely American-born population to an absolute majority of German-born residents in the space of just a few years. A similar situation could be found for areas in Minnesota, where nearly one fifth of all Swedish migrants settled forming "a nation within a nation" (Daniels, 1990: 173). Swedish schools, newspapers and community centres were built, giving the region a distinctively Swedish character. Parts of Iowa, turned Norwegian establishing ethnic press as the *Decorah-Posten* with more than forty thousand subscribers at the time. However, many counties in the US attracted a large variety of different nationalities as well as American-born population from all over the country. New York City, for example, lured a multitude of different nationalities from all over Europe, including Scandinavians, French, Germans, Irish, Italians, Hungarians, Russians, and Greeks. With the largest minorities even establishing ethnic neighbourhoods such as *Kleindeutschland* (little Germany) or Little Italy, New York City – as other major cities like Chicago, Boston, New Orleans, San Francisco, Philadelphia, or Baltimore – became a hub of population diversity (Daniels, 1990; Spickard, 2007).

Figure 4-5 displays the level of population diversity at county level across the 48 continental US states for the peaks of the two waves of migration, 1880 (a) and 1910

(b) respectively.[8] Despite lower diversity levels in large parts of the Midwest in 1910, the situation in both years is similar: Apart from New Mexico, parts of Utah, and Texas, high diversity levels dominate especially the west of the country. Long-established routes of migration in combination with the shift of the Western border resulted in high levels of population redistribution across the US. The East, in comparison, had a far greater population homogeneity, with the exception of the urban centres in the north and Florida in the south. While counties along the shores of the Great Lakes or within the corn and wheat belt were largely dominated by only very few population groupings, urban areas, such as New York City, Boston or Miami, were vibrant and buzzing migrant concentrations, attracting a large variety of nationalities. The states of the 'Old South', as expected, had the lowest diversity levels.

[8] In order to depict diversity, a widely-used fractionalisation index (Alesina *et al.*, 2003) is employed, placing the focus on the number of groups in one county. The higher the index, the higher the fractionalisation, or diversity, within the population of the respective county.

a)



b)

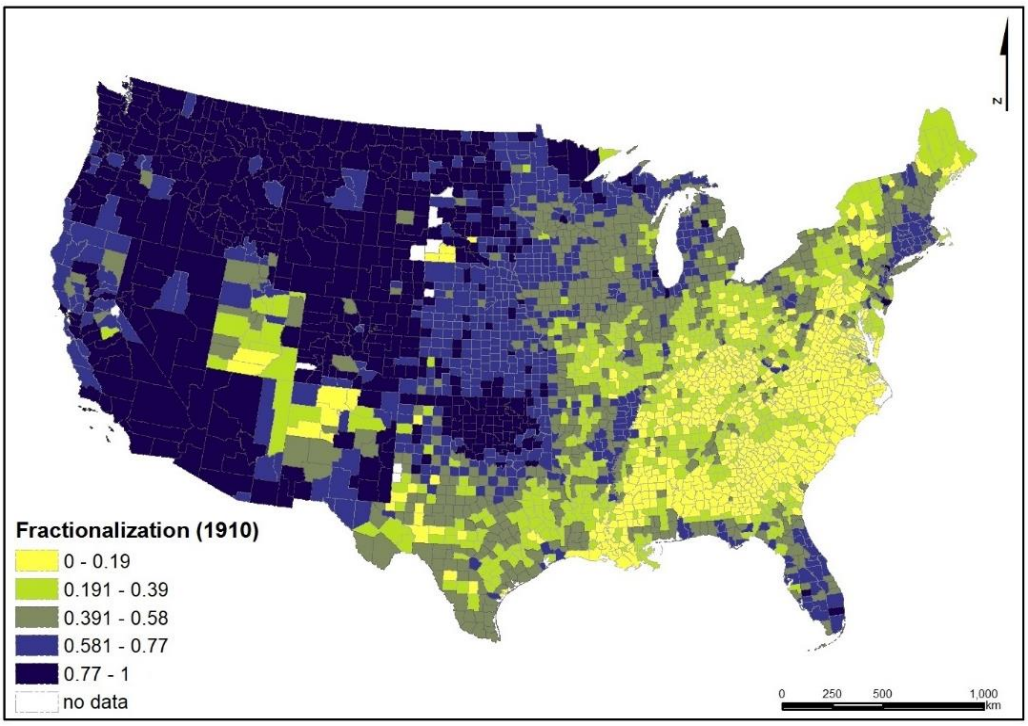


FIGURE 4-5 POPULATION DIVERSITY BY COUNTY, 1880 AND 1910[9]

Source: Ruggles et al., (2015), own elaboration

[9] The calculation of the fractionalisation index will be explained in detail in section 5.4.

With increasing migrant numbers and population diversity in the US, public opinion towards immigration started to shift. The *laissez-faire* policy towards immigration of the first century after independence had allowed nearly all migrants, regardless of nationality to enter the country and roam freely in search of a place to settle down. However, towards the end of the 19th century, voices of fear and concern started to break out among the American population (Carter and Sutch, 2006). While Scandinavians, Germans, Irish, and other western Europeans of the first wave seemingly 'blended in' with the white population of British, Welsh, and Scottish origin, "now there came multitudes of men of lowest class from the south of Italy and men of the meanest sort out of Hungary and Poland, men out of the ranks where there was neither skill nor energy nor any initiative of quick intelligence; and they came in numbers which increased from year to year, as if the countries of the south of Europe were disburdening themselves of the more sordid and hapless elements of their population" (Wilson, 1901). Greater diversity and, specifically, being seen as different from the 'pioneers of the century of immigration' became problematic.

The first to suffer from racial prejudice and economic concerns by – in this case – the predominantly white working-class population were the Chinese. In 1882, Congress passed the Chinese Exclusion Act, which became "the hinge on which American immigration policy turned" (Daniels, 1990: 271). It was the first immigration law passed by congress excluding a single ethnic group from immigration to the US, thereby actively regulating immigration.

Further concerns about the standard of living of American workers in combination with theories postulating superior innate characteristics associated with Anglo-Saxon and Nordic peoples (largely proclaimed by interest groups such as the Immigration Restriction League) led to growing support of further restrictions to the immigration

flow. The first significant general restriction was passed in 1917, when a literacy test for all immigrants arriving on the shores of the US was declared mandatory. Targeted to limit migrant numbers from especially the Southern and Eastern parts of Europe, this law was ultimately designed to hand-pick those who arrived and select based on skill-levels (Daniels, 1990).

The final closure of the doors to US immigration came into effect in 1924, when a quota system based on the National Origin Formula was introduced. A maximum quota was set for each sending country equivalent to 2 percent of the country's population living in the US at the time of the 1890 census (Snyderman and Herrnstein, 1983; Daniels, 1990). With the goal of proactively lowering immigration numbers while simultaneously "preserving racial homogeneity" (Snyderman and Herrnstein, 1983: 993), this law effectively restricted immigration by Southern and Eastern Europeans, while *de facto* banning immigration from other parts of the world. With the Immigration Act of 1924, the century of immigration had come to an end. The formerly wide open doors were now firmly shut.

4.2 CONTEMPORARY MIGRATION - THE LATINOS' CASE

The Great Depression and World War II in combination with the restrictive immigration legislation of the 1920s resulted in a significant drop of annual immigration numbers to the US. It was not until the 1950s when immigration picked up speed again. This time however, migration came from 'new world' destinations, to a large extent incentivised by the recruitment of guest workers which had already started during war times (Daniels, 1990).

With the US intervention in WWII, thousands of men headed off to Europe, leaving women to join the labour force, filling in for their husbands, brothers, uncles, or cousins. However, labour shortages, especially in the agricultural sector, were so extreme that business leaders mounted efforts to recruit guest workers from elsewhere. Mexico became the main country for additional labour supply. In 1942, for example, the *Bracero* Program was launched, allowing US employers to hire Mexican labourers to work in the US for a designated time, at a fixed minimum wage and decent living conditions (Sáenz and Morales, 2015). Over the program's 22 years of existence, more than 5 million Mexicans received employment contracts in more than 20 US states (Calavita, 1992; Spickard, 2007). By 1960, around 30 percent of the seasonal agricultural workers in the US were Mexican *braceros* (Daniels, 1990). Simultaneously, the number of Mexican immigrants receiving permanent resident status increased. The more information on the US way of life, wages, employment, and economic conditions *braceros* brought home, the more individuals felt the urge to move north of the Mexican border. While "nearly 60.000 came in the 1940s, almost 275.000 in the 1950s, more than 440.000 in the 1960s and almost 640.000 [entered the United States as resident aliens with permanent status] in the 1970s" (Daniels, 1990: 311). Unintentionally, the programme also attracted large numbers of illegal or 'undocumented' immigrants during its lifetime, enlarging the Latino, and especially Mexican, community in the US (Gann, 1986).

Further guest workers, albeit recruited to a far smaller extent than their Mexican counterparts, were Filipinos and Puerto Ricans. Working predominantly in agriculture in the East and Midwest or filling positions in the factories or service industry in large cities, the share of population from these 'new world' countries increased significantly (Spickard, 2007).

However, it was not until the 1960s that annual immigration numbers to the US grew exponentially. The trigger for this mass movement of Latinos and, to a lesser extent, Asians was a change in the immigration legislature. In 1965, the implementation of the Immigration and Nationality Act abolished the national quota system of the 1920s putting an end to favouritism of northern and western Europeans and repealed limits and entry barriers for other nationalities (e.g. southern, eastern Europeans and Asians). A complex regulatory system replaced the old restrictions, setting upper ceilings for immigration numbers to "170.000 for eastern hemisphere with a maximum of 20.000 for a given country […] and 120.000 for western hemisphere with no limit imposed on any country" (Sáenz and Morales, 2015: 37). Three groups, however, were exempt from these restrictions and given preferential entry to the country: (1) Foreigners who had immediate relatives in the US, (2) refugees and (3) individuals who had special skills and were needed to fill essential job vacancies (Martin and Midgley, 1999). Neither Asians nor Latinos, were believed to benefit greatly from this policy, primarily intended to boost immigration numbers from western economies. Mexican immigration was thought to be subject to the cap for the western hemisphere and Asians did not constitute large parts of the population at the time (Sáenz and Morales, 2015). Yet, what was intended as accelerator of European immigration, turned out, rather unexpectedly, to change the entire composition of the US immigrant community.

While historically Europeans had been the largest group arriving in the US, after 1965, the main sending regions shifted to Asia and especially to Latin America (Martin and Midgley, 1999). The newly implemented policy's focus on family reunifications benefited Mexicans with permanent residence in the US. Political unrest after Castro's revolution or the Mariel Crisis, granted Cubans refugee status which allowed

preferential entry. By 1980, almost 10 percent of Cuba's total population lived in the US (Gann, 1986; Bergad and Klein, 2010). Asians entered either as skilled workers or, in the case of the Vietnamese, Cambodians, or Lao, as refugees from the Vietnam war (Sáenz and Morales, 2015).

Figure 4-6 displays the shift in major sending regions covering more than a century of US immigration history. While the first two columns portray the Age of Mass Migration with its two peaks in 1880 (first-wave) and 1910 (second-wave), the latter two highlight the changing demographic makeup of the immigrant population in the second half of the 20th century. As expected, Europeans constituted around 80 to 90 percent of individuals obtaining permanent resident status in the US in historical times. Incoming population from Latin America, Asia, Africa, Oceania, or other North American regions represented mere minorities within the stock of immigrants at the time. By 2000, however, the European share of migrants had decreased drastically making room for two 'new' major sending regions: Asia and, fundamentally, Latin America. Today, migrants from Mexico, the Caribbean, and Central and South America make up more than 50 percent of all individuals who have been granted permanent residence in the US. Despite a significant growth of Asian immigration over time, the analysis in this dissertation will focus solely on the foreign-born communities from Latin America, the largest sending region for US immigration.

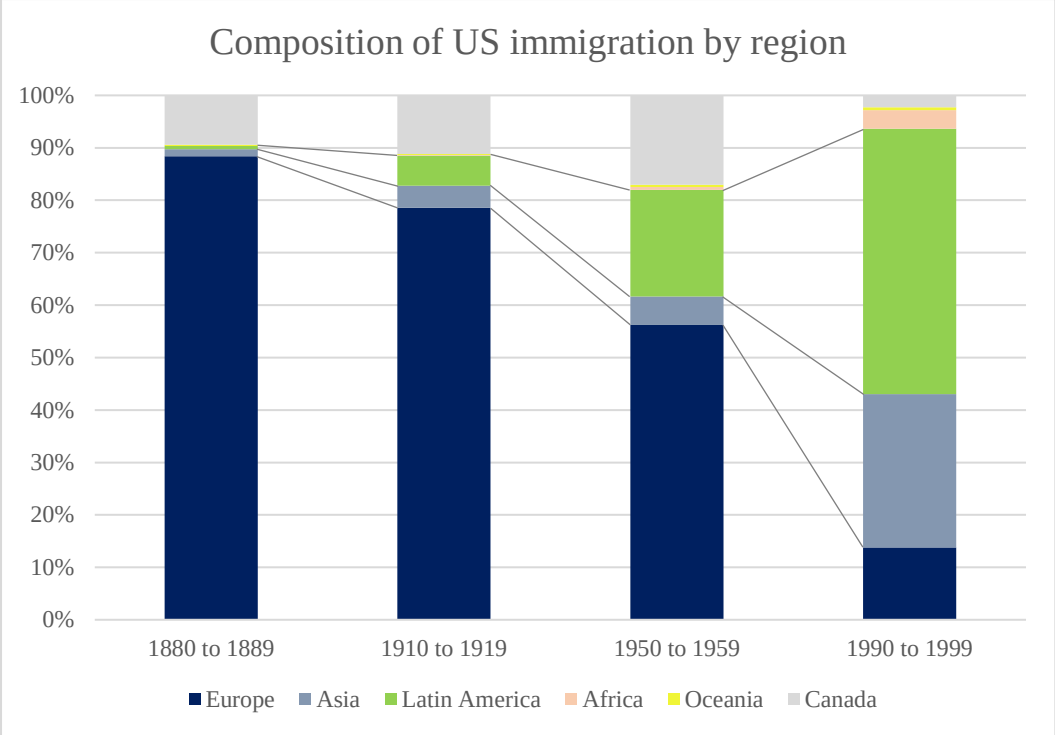


FIGURE 4-6 PERSONS OBTAINING LEGAL PERMANENT RESIDENT STATUS, 1880-1999

Source: United States, Department of Homeland Security (2011), own elaboration

The Latinos arriving in the US were by no means a homogenous mass. Table 4-2 displays their composition at two different points in time.[10] In the 1960s, the large majority immigrated from Mexico, while further sizable groups originated from Cuba, Puerto Rico,[11] Columbia, and the Dominican Republic. While all sought economic opportunity and a better quality of life in the US, the reasons to leave their home countries varied by nationality. Cuban migration surged in the aftermath of the Cuban revolution (Massey, 2008). Puerto Ricans escaped overpopulation and mass unemployment, following the transition from a monocultural plantation economy to a

[10] Table 4-2 only captures the official extent of legal immigration. Undocumented migrants to the US are not included in these numbers, even though they became increasingly relevant from the 1960s onwards.
[11] The Puerto Ricans are not shown in Table 4-2, as they are considered American-born since the Jones-Shafroth Act of 1917 (Bergad and Klein, 2010).

focus on manufacturing and exports. Political violence and rural unrest gave Colombians reason to leave their home country in the 1950s and 1960s. Later, widespread unemployment due to rapid industrialisation and the mechanisation of agriculture grew into main engines of Colombian immigration to the US. Political turmoil in the 1950s and 1960s was the main factor behind emigration from the Dominican Republic. In the 1980s, the reasons for Dominicans to leave their home country shifted due to the disastrous economic conditions following the fall in the price of their main export goods, sugar and rum (Gann, 1986).

Over time, other Latin American countries turned into additional significant sources of immigrants. By 2000, Central and South America represented around 40 percent of the Latin American immigrant community in the US. Peruvians, Salvadorians, Hondurans, Ecuadorians, Guatemalans, and Nicaraguans created sizeable communities in the US, as did immigrants from other Caribbean countries (especially Haiti and Jamaica). “Social problems connected with civil liberty and land” (Gann, 1986: 118), but also internal dissensions and ultimately guerrilla wars led to massive outmigration of these countries. Mexicans, however, kept their overwhelming dominance over incoming migrant flows and have remained the largest Latino community across the US since.

TABLE 4-2 LATIN AMERICANS OBTAINING LEGAL PERMANENT RESIDENT STATUS, 1960-69, 2000-09

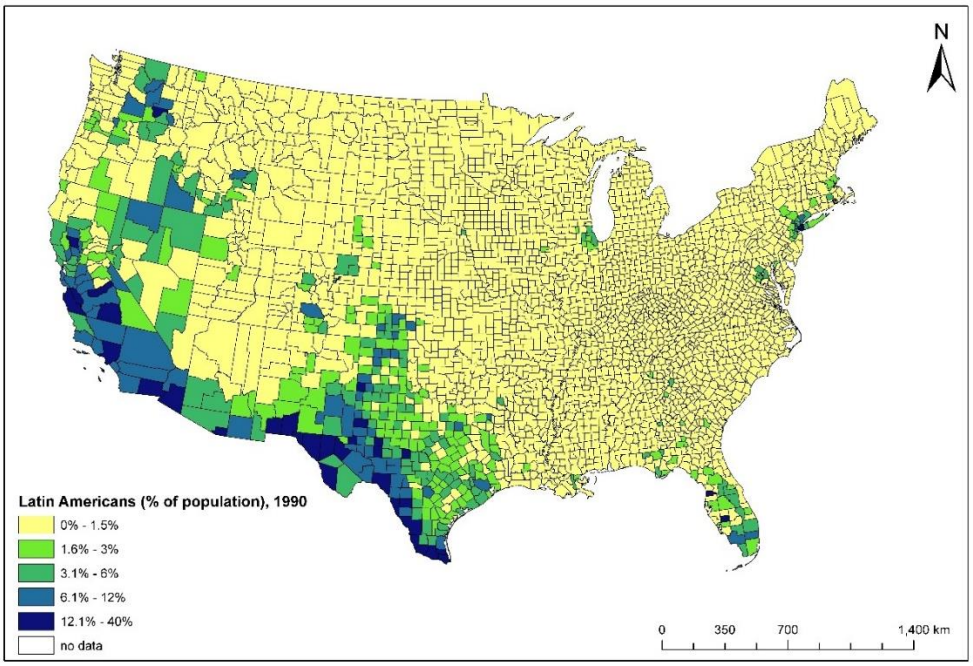
BY COUNTRY OF LAST RESIDENCE

Region and country of last residence	1960 to 1969		2000 to 2009	
	Total number	Share	Total number	Share
Latin America	1.241.044	100%	4.205.877	100%
Mexico	441.824	36%	1.704.166	41%
Caribbean	427.235	34%	1.053.969	25%
Cuba	202.030	16%	271.742	6%
Dominican Republic	83.552	7%	291.492	7%
Haiti	28.992	2%	203.827	5%
Jamaica	62.218	5%	172.523	4%
Other Caribbean	50.443	4%	114.385	3%
Central America	98.560	8%	591.130	14%
Belize	4.185	0%	9.682	0%
Costa Rica	17.975	1%	21.571	1%
El Salvador	14.405	1%	251.237	6%
Guatemala	14.357	1%	156.992	4%
Honduras	15.078	1%	63.513	2%
Nicaragua	10.383	1%	70.015	2%
Panama [1]	22.177	2%	18.120	0%
Other Central America	-	-	-	-
South America	250.754	20%	856.593	20%
Argentina	49.384	4%	47.955	1%
Bolivia	6.205	0%	21.921	1%
Brazil	29.238	2%	115.404	3%
Chile	12.384	1%	19.792	0%
Colombia	68.371	6%	236.570	6%
Ecuador	34.107	3%	107.977	3%
Guyana	4.546	0%	70.373	2%
Paraguay	1.249	0%	4.623	0%
Peru	19.783	2%	137.614	3%
Suriname	612	0%	2.363	0%
Uruguay	4.089	0%	9.827	0%
Venezuela	20.758	2%	82.087	2%
Other South America	28	0%	87	0%
Other America	22.671	2%	19	0%

Data source: United States, Department of Homeland Security (2011)

Despite this heterogeneity in nationalities, Latin American immigrants clustered in just a few distinctive areas of the US. Figure 4-7 displays the settlement pattern of Latin-American-born population across the 48 continental states. Figure 4-7 a), represents the Latino stock in 1980, while Figure 4-7 b), displays Latino settlement in 2010.

a)



b)

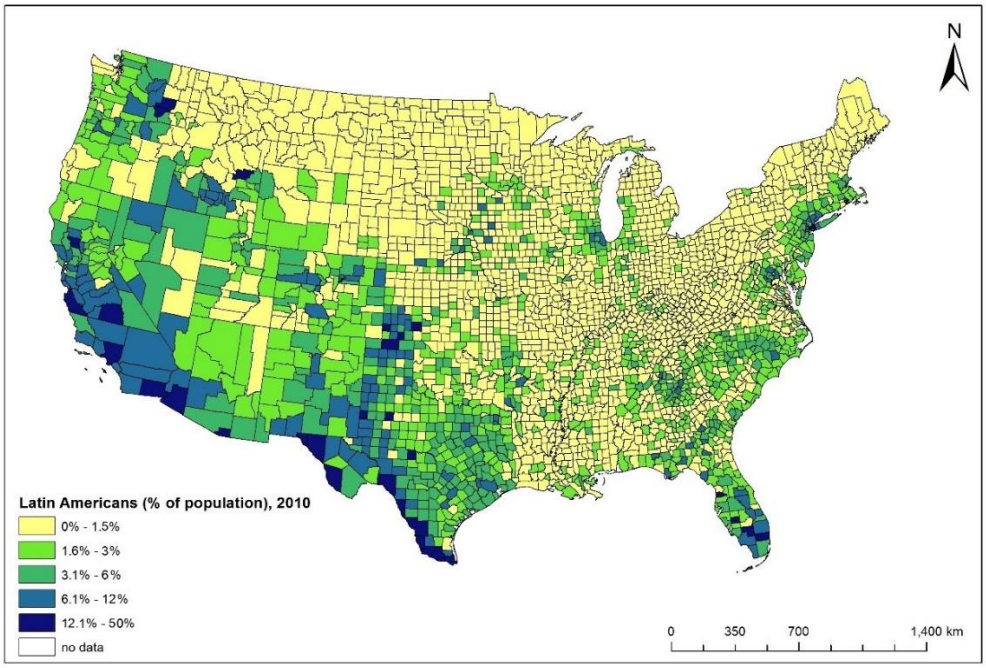


FIGURE 4-7 LATINO FOREIGN-BORN POPULATION AS SHARE OF TOTAL POPULATION, 1980 AND 2010

Data source: Minnesota Population Center (2016); own elaboration

Latinos usually entered the US either via Florida or the border with Mexico. Prior to 2000, the vast majority concentrated in the Western and Southwestern states, such as Texas, New Mexico, Arizona, and California. Further clusters of Hispanics included Florida, New York, but also parts of Illinois, Nevada, Washington, and Idaho. In 1980, over half of the entire Latino population resided in Texas and California alone. Another 17 percent lived in New York (primarily Puerto Ricans) and Florida (especially Cubans) (Bergad and Klein, 2010). Thirty years later, the large majority of Hispanics still lived in this handful of states, however, the overall settlement pattern had evolved.

Apart from traditional Latino destinations such as Texas, Florida, California, New York, and Illinois, Figure 4-7 unveils the presence of large Latino communities in so-called 'new-destination states' such as Colorado, Utah, Wyoming, or the Midwest, including Oklahoma and Kansas, in 2010. Similarly, states such as Georgia, Virginia, or North Carolina – which had traditionally been avoided by international immigrants – start attracting a significant number of Latinos (Bailey, 2005; Bump, 2005; Hansen, 2005). This change in settlement pattern was largely a result of changes in industry structure, amendments to immigration as well as border policies and alterations in the geography of the labour demand. The combination of all three factors incentivised Latinos to establish themselves in regions beyond their traditional settlement pattern (Donato *et al.*, 2008). Thus, while in the 1960s through to the 1990s, growing Hispanic communities were regionally concentrated in only a handful of states, shortly after the turn of the century, Latin Americans turned "from a regional to a national phenomenon" (Massey and Capoferro, 2008: 47), spreading out across the entire US territory. The areas of the Midwest and South, for example, suddenly attracted masses of Southern and Central Americans. This was predominantly due to the strategic

decision of the meat and poultry industries to vertically integrate, deskill and move production facilities from urban into rural areas. Attracted by tax incentives, exemptions from restrictive business policies and lucrative environments, the industry started recruiting especially Latino migrants who were willing to take the low-paying positions with limited benefits and sometimes even dangerous working conditions the local population refused to fill (Kandel and Parrado, 2005; Zúñiga and Hernández-León, 2005).

Their settlement pattern was split between rural areas, on the one hand, where Latinos worked predominantly in the agricultural sector, and a handful of 'gateway' metropolitan areas of the US, on the other. Miami (FL), Los Angeles (CA), New York City (NY), Houston (TX), or Chicago (IL) became key hubs of Latino settlement (Massey, 2008).

Among Latinos different nationalities have tended, to cluster in different areas of the country. New York and New Jersey have been most appealing to Puerto Ricans and Dominicans. Two-thirds of the entire Dominican population in the US lives in these states. Puerto Ricans form the single largest Latino group in New York. Cubans are concentrated predominantly in Florida where they represent the largest Latino community. Central Americans have converged towards the south west (especially California and Texas), Florida, and New York. South Americans have clustered in New York and New Jersey, but live also in Florida and California. More than half of the entire Colombian population in the US, for example, lives in these states (Bergad and Klein, 2010; Sáenz and Morales, 2015). Mexicans constitute the most geographically outspread group. Even though 60 percent of the Mexican population in the US live in California and Texas (Los Angeles being the main centre), a large part of the remaining 40 percent has fanned out beyond the historically inhabited

Latino regions and is living in ‘new-destination areas’, such as North and South Carolina, Virginia, Georgia, or Alabama (Bergad and Klein, 2010; Sáenz and Morales, 2015).

Even though the Latino immigrants to the US originate from a multitude of countries bringing a variety of customs, traditions, and backgrounds along with them, a rough profile of the ‘average Latino’ in the US can be drawn. Generally, at arrival in the United States, the foreign-born population originating from Mexico, the Caribbean, Southern and Central America are young, relatively unskilled, from impoverished rural environments, very religious, and, with the exception of immigrants from some Caribbean countries and Brazil, Spanish-speaking.

Apart from post-Castro Cuban exiles, who predominantly stemmed from the highest ranks of society, the big majority of Latino migrants are of low social rank (Gann, 1986; Sáenz and Morales, 2015). With lowest tertiary education rates and the highest drop-out numbers from high-school compared to all other ethnic groups, Latinos make up the lowest part of the educational spectrum (Bergad and Klein, 2010). Accordingly, only very few Hispanics fill higher-status professional positions. The big majority works in low-paying jobs within “the Latino immigrant occupational niche” (Sáenz and Morales, 2015: 109) such as “agriculture labour, meat, and poultry, and seafood processing; construction; waiters/waitresses; cooks; maids and housekeeping cleaners; and janitors and building cleaners” (Douglas and Sáenz, 2008: 169). Thus, Latino median personal incomes are lowest compared to all other ethnic groups (Bergad and Klein, 2010).

At a median age of around 30, the Latino population is significantly younger than the median white non-Hispanic (42 years of age) (Sáenz and Morales, 2015). In contrast

to their European predecessors one century earlier, women are more strongly represented, resulting into an almost gender-balanced immigrant population, especially from Southern and Central American countries. The majority, roughly 75 percent of Hispanic population above the age of 15 is married and lives mostly in families with more children than the average American one. Latino households consist of an average of 4.3 people. This is significantly higher than for the non-Hispanic white population (3.0 people per household). The Latino foreign-born population constitutes the ethnic group with the largest share of big family households. 33 percent of the foreign-born Latino population lives in families of 5 persons or more (Bergad and Klein, 2010).

The Hispanic community within the US has grown massively over the last four decades. Including second and third generations, it has grown from "14.6 million in 1980 to 22.4 million in 1990 to 35.5 million in 2000 and to 50.5 million in 2010 […] In fact, if the US Latino population were a country in 2010, it would be the 25th largest country in the world" (Sáenz and Morales, 2015: 49). Both natural increase as well as immigration resulted in an average population growth rate of 3.5 percent per year over the last three decades – a growth rate significantly higher than the one of any other ethnic group or the entire US population. In the 2000s alone, Latinos accounted for more than 50 percent of the total US population's growth rate. While in 1980, every 16th US resident had a Hispanic background, by 2010, one in six people did (Sáenz and Morales, 2015). In parallel to their European predecessors more than a century earlier, with the turn of the 20th to the 21st century, Latinos became the single largest minority within the US. Projections estimate that by 2050 the Latino population of the US will increase threefold to represent at that point more than 30 percent of the entire nation's population.

It is crucial to understand these features of US immigration history in order to put the analysis of the long-term economic impact of migration within the remaining part of this dissertation into context. Both the time around the turn of the 19th to the 20th century as well as more recent migration from 1970 onwards, depict crucial eras in US migration history. In both these periods immigration – either in absolute or in relative numbers – rose to unprecedented levels: inward migrant numbers grew almost exponentially and the share of foreign-born in the US population increased to close to 15 percent of the total. As a consequence, the composition of local communities changed markedly. Immigration shaped public opinion, altered the political agenda and eventually affected US immigration policy. Both eras therefore represent key periods for the analysis of the long-term economic impact of migration.

4.3 REFERENCES

Abramitzky, R., Boustan, L., and Eriksson, K. (2014): A nation of immigrants: Assimilation and economic outcomes in the Age of Mass Migration. *Journal of Political Economy, 122*(3), 467–506.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003): Fractionalization. *Journal of Economic Growth, 8*(2), 155–194.

Alexander, J. G. (2007): *Daily life in immigrant America: 1870–1920*. London: Greenwood Press.

Atack, J. and Passell, P. (1994): *A new economic view of American history: From colonial times to 1940*. New York: Norton.

Atack, J., Bateman, F., and Parker, N. (2000): Northern agriculture and the westward movement. In R. G. Engerman (Ed.), *The Cambridge Economic History of the United States: The Long Nineteenth Century*, 285–328. Cambridge: Cambridge University Press.

Bailey, R. (2005): New immigrant communities in the North Carolina Piedmont Triad: Integration issues and challenges. In E. M. Goździak and S. F. Martin (Eds.), *Beyond the gateway: Immigrants in a changing America*, 57–86. Lanham: Lexington Books.

Baines, D. (1994): European emigration, 1815-1930: Looking at the emigration decision again. *The Economic History Review, 47*(3), 525–544.

Bergad, L., and Klein, H. (2010): *Hispanics in the United States: a demographic, social, and economic history, 1980-2005*. Cambridge: Cambridge University Press.

Bertocchi, G., and Strozzi, C. (2006): *The Age of Mass Migration: Economic and institutional determinants*. Discussion paper Series, IZA Institute for the Study of Labour, No. 2499. Bonn, Germany.

Bodnar, J. (1992): *Remaking America: public memory, commemoration, and patriotism in the twentieth century*. Princeton, N.J: Princeton University Press.

Bump, M. (2005): From temporary picking to permanent plucking: Hispanic newcomers, integration, and vhange in the Shenandoah Valley. In E. M. Goździak and S. F. Martin (Eds.), *Beyond the gateway: Immigrants in a changing America*, 137–177. Lanham, MD: Lexington Books.

Calavita, K. (1992): *Inside the state: the bracero program, immigration, and the I.N.S*. New York: Routledge.

Carter, S., and Sutch, R. (2006): *Resident impacts of immigration: Perspectives from America's Age of Mass Migration.* UCR Economics, No. 200808.

Daniels, R. (1990): *Coming to America: A history of immigration and ethnicity in American life*. New York, NY: Harper Perennial.

Donato, K., Tolbert, C., Nucci, A., and Kawano, Y. (2008): Changing faces, changing places: The emergence of new nonmetropolitan immigrant gateways. In D. S. Massey (Ed.), *New Faces in New Places*, 75–98. New York, NY: Russell Sage Foundation.

Douglas, K., and Sáenz, R. (2008): No phone, no vehicle, no English, and no citizenship: The vulnerability of Mexican immigrants in the United States. In A. Hattery, D. G. Embrick, and E. Smith (Eds.), *Race, Human Rights and Inequality*, 161–180. Lanham, MD: Rowman and Littlefield.

Dunlevy, J., and Gemery, H. (1977): The role of migrant stock and lagged migration in the settlement patterns of nineteenth century immigrants. *The Review of Economics and Statistics*, *59*(2), 137–144.

Ferrie, J. (2005): History lessons: The end of American exceptionalism? Mobility in the United States since 1850. *Journal of Economic Perspectives*, *19*(3), 199–215.

Fitzpatrick, D. (1984): *Irish emigration 1801-1921*. Dublin, Ireland: Economic and Social History Society of Ireland.

Gann, L. (1986): *The Hispanics in the United States: a history*. Boulder: Westview Press.

Guinnane, T. (1992): Intergenerational transfers, emigration, and the rural Irish household system. *Explorations in Economic History*, *29*(4), 456–476.

Hansen, A. (2005): Black and white and the other: International immigration and change in Metropolitan Atlanta. In E. M. Goździak and S. F. Martin (Eds.), *Beyond the gateway: Immigrants in a changing America*, 87–110. Lanham, MD: Lexington Books.

Hatton, T., and Williamson, J. (1993): After the famine: Emigration from Ireland, 1850–1913. *Journal of Economic History*, *53*(3), 575–600.

Hatton, T., and Williamson, J. (1994): What drove the mass migrations from Europe in the late nineteenth century? *Population and Development Review*, *20*(3), 533–559.

Joly, D. (2000): Some structural effects of migration on receiving and sending countries. *International Migration*, *38*(5), 25–40.

Kamphoefner, W. (1976): At the corssroads of economic development: Background factors affecting emigration from nineteenth century Germany. In I. A. Glazier and L. DeRosa (Eds.), *Migration Across Time and Nations*. New York: Holmes and Meier.

Kandel, W., and Parrado, E. (2005): Restructuring of the US meat processing industry and new Hispanic migrant destinations. *Population and Development Review*, *31*(3), 447–471.

Levy, M., and Wadycki, W. (1973): The influence of family and friends on geographic labor mobility: An international comparison. *The Review of Economics and Statistics*, *55*(2), 198–203.

Martin, P., and Midgley, E. (1999): Immigration to the United States. *Population Bulletin*, *54*(2), 3.

Massey, D. (2008): *New faces in new places: the changing geography of American immigration*. New York, NY: Russell Sage Foundation.

Massey, D., and Capoferro, C. (2008): The geographic diversification of American immigration. In D. S. Massey (Ed.), *New Faces in New Places*, 25–50. New York, NY: Russell Sage Foundation.

Merk, F. (1978): *History of the westward movement*. New York: Alfred A. Knopf.

Migration Policy Institute. (2015): Top 25 destinations of international migrants. Retrieved September 14, 2017, from http://www.migrationpolicy.org/programs/data-hub/charts/top-25-destinations-international-migrants

Migration Policy Institute. (2017): U.S. Immigrant population and share over time, 1850-present. Retrieved June 28, 2017, from http://www.migrationpolicy.org/programs/data-hub/us-immigration-trends#history

Minnesota Population Center. (2016): National historical geographic information system: Version 11.0 [Database]. https://doi.org/http://doi.org/10.18128/D050.V11.0.

Rodríguez-Pose, A., and von Berlepsch, V. (2014): When migrants rule: The legacy of mass migration on economic development in the United States. *Annals of the Association of American Geographers*, *104*(3), 628–651.

Rodríguez-Pose, A., and von Berlepsch, V. (2015): European migration, National origin and long-term economic development in the United States. *Economic Geography*, *91*(4), 393–424.

Ruggles, S., Genadek, K., Goeken, R., Grover, J., and Sobek, M. (2015): *Integrated Public Use Microdata Series: Version 6.0 [dataset]*. Minneapolis: University of Minnesota.

Sáenz, R., and Morales, M. C. (2015): *Latinos in the United States: Diversity and change*. Cambridge, UK: Polity Press.

Snyderman, M., and Herrnstein, R. (1983): Intelligence tests and the Immigration Act of 1924. *American Psychologist*, *38*(9), 986–995.

Spickard, P. (2007): *Almost all aliens: immigration, race, and colonialism in American history and identity*. New York: Routledge.

United States, Department of Homeland Security (2011): *Yearbook of Immigration Statistics: 2010*. Washington, D.C.

Vedder, R., and Gallaway, L. (1972): The geographical distribution of British and Irish emigrants to the United States after 1800. *Scottish Journal of Political Economy*, *19*(1), 19–35.

Ward, D. (1990): Population growth, migration and urbanization, 1860–1920. In R. Mitchell and P. Groves (Eds.), *North America: The historical geography of a changing continent*, 299–319. Totowa, NJ: Savage.

Wilson, W. (1901): *A History of the American People*. New York: Harper and Brothers.

Zúñiga, V., and Hernández-León, R. (2005): *New destinations: Mexican immigration in the United States*. New York, NY: Russell Sage Foundation.

5 POPULATION DIVERSITY AND ITS LONG-TERM IMPACT FOR ECONOMIC DEVELOPMENT

5.1 INTRODUCTION

In 2015, worldwide migration numbers exceeded expectations and rose to 244 million (UNDESA, 2016). These growing population flows have contributed to a shift both in the political discourse and in the scientific research agenda, bringing the analysis of the economic implications of migration into the fore.

Over the past few decades, a vast amount of new scientific research has led to considerable progress in our understanding of the economic implications of migration. The economic impact of migrants on both their own futures and those of locals (i.e. Borjas, 1994; Card, 2005), on the local labour market and its dynamics (i.e. Altonji and Card, 1991; Friedberg and Hunt, 1995; Bijak et al., 2007), and on public finances (i.e. Kerr and Kerr, 2011) has been extensively analysed. Transmission channels – such as increasing returns to scale (i.e. Borjas, 1995), alterations to the ratio of skilled to unskilled labour (Lundborg and Segerström, 2002), wages (Ottaviano and Peri, 2006), or the stimulation of productivity by means of innovation and specialisation (i.e. Gordon and McCann, 2005; Partridge and Furtan, 2008) – have also been objects of greater scrutiny. The focus of these studies, however, has generally been short-term. Our understanding of the economic implications of migration has commonly been limited to the first five to 10 years after the initial migration wave took place. Analysis of the medium- to long-term impact of migration on economic prosperity has been mostly neglected. Only in recent years have researchers started to address this gap. In

particular, recent work by Rodríguez-Pose and von Berlepsch (2014) has demonstrated how current levels of economic development across the US still depend on migration settlement patterns that took place over 100 years ago. Sequeira et al. (2017) recently confirmed the significance of this relationship. This long-term impact of migration holds in time regardless of the national origin of migrants settling in different territories (Rodríguez-Pose and von Berlepsch, 2015).

Despite this progress in research, one important demographic aspect related to migration has remained firmly anchored in short-term scrutiny: diversity. As formerly homogeneous communities become more diverse by accommodating new individuals bringing their customs, traditions, ideas, abilities and experiences with them, the question of whether more diverse societies facilitate or deter growth has become more prominent. Research on the economic impact of population diversity has flourished, focusing on a multitude of transmission channels ranging from skill variety, social interaction, innovative networks, institutions and the provision of public goods to trust, social participation, social unrest and conflict (i.e. Easterly and Levine, 1997; Alesina and La Ferrara, 2005; Ottaviano and Peri, 2006; Gören, 2014; Alesina et al., 2016; Bove and Elia, 2017; Kemeny and Cooke, 2017). Most of this research unveiled a considerable effect of diversity on growth over the short-term. However, our knowledge about whether population diversity levels generated by past migration waves still affect economic outcomes over the medium- and long-term remains an almost untouched area within the scientific literature. This chapter intends to fill this gap.

We seek to ascertain whether areas that were characterised by a large degree of population diversity more than a century ago are wealthier today than those that remained more homogenous in their population composition. Does having a very

diverse population at one point in time lead to persistently higher levels of economic growth? Or is the economic impact of diversity only evident in the short-term, vanishing once the different population groups become part of the society's 'melting pot'?

In this chapter, we assess the extent to which the high degree of cultural diversity in US counties generated during the Era of Mass Migration of the late 19th and early 20th century has left an enduring impact on the economic development of those US areas that witnessed the greatest heterogeneity in population. Incorporating a two-fold definition of the notion of diversity, encompassing two distinct dimensions of the term (fractionalisation and polarisation), we undertake a decade-by-decade analysis for the US at the subnational level covering the period between 1880 and 2010. We posit that a vibrant, highly diverse population, stemming from a multitude of different backgrounds, nationalities, and cultures, bringing along the risk-seeking and entrepreneurial character of the migrant has the capacity to leave a long-lasting economic impact. We speculate the economic dynamism created by high levels of population diversity to have become embedded not only in local institutions but in the very core of a territory, affecting the subsequent economic development path of the region not only over the short-, but also over the medium- and long-term.

In order to test whether this is the case, the chapter adopts the following structure: Section 5.2 gives an overview of the historical background of the Age of Mass Migration. Section 5.3 summarises previous approaches to the link between diversity and economic development in the literature. In section 5.4, we describe the model, methodological aspects, and the data adopted for our research. We also explain the calculations of the various indices used in the chapter as main variables of interest. Section 5.5 reports the results of our estimation, and section 5.6 concludes.

5.2 MASS MIGRATION TO AND WITHIN AMERICA – A SHORT OVERVIEW

When speaking of the Age of Mass Migration to the US, historians refer to the period between the pre-Civil War years and the mid-1920s. Within this time span more than 40 million Europeans left their homelands as a result of varying degrees of political disturbances, famine, and religious persecution in search of a new and better life. The large majority of these migrants chose the US as their final destination (Hatton and Williamson, 1994; Bertocchi and Strozzi, 2006). With an average annual immigration inflow rate of about 0.75 percent of the total US population at the time (Hatton and Williamson, 1998), the US experienced a population increase of an extent that had been unheard of in modern history.[12] During this period, total US population increased six-fold, from about 17 to 105 million. Meanwhile, the proportion of the foreign-born white population grew from 13 percent in 1850 to approximately 18 percent in 1910 (Table 5-1). Most importantly, "the proportion of people of foreign birth and parentage together reached its maximum level of 45 percent in 1920" (Ward, 1990: 299).

At the time, no legislation existed which restricted migrants from entering the country. Migrants – no matter which nationality – could roam freely and settle wherever they wished.[13] The introduction of the literacy test in the Immigration Act of 1917 led to the first serious restriction to immigration. Quotas for incoming migrants followed in

[12] In peak years, the annual inflow rate of immigrants reached heights of around 1.5% of the total population at the time (Kim, 2007).
[13] With the notable exception of the Chinese after the Chinese Exclusion Act of 1882 (Frazier and Margai, 2010).

1921 (Emergency Quota Act of 1921). By 1924 entry restrictions for all foreigners were passed (Goldin, 1994; Alexander, 2007).

TABLE 5-1 US POPULATION COMPOSITION (IN % OF TOTAL POPULATION), 1840 – 1920

Year	Population (millions)	Black (%)	Foreign parentage (%)	Foreign-born (%)
1840	17.1	16.8	n.d.	n.d.
1850	23.2	15.7	n.d.	12.9
1860	31.4	14.1	n.d.	17.9
1870	39.8	13.5	19.0	19.6
1880	50.2	13.1	22.5	17.8
1890	62.9	11.9	25.0	19.9
1900	76.0	11.6	27.6	18.1
1910	92.0	10.7	27.8	18.0
1920	105.7	9.9	28.0	16.9

n.d. = no data Source: Ward, 1990

Most newcomers settled where relatives and friends had already settled (e.g. Vedder and Gallaway, 1972; Levy and Wadycki, 1973; Dentlevy and Gemery, 1977), creating distinct migrant communities across the country. Hence, regions with large migrant networks attracted further newcomers while others remained nearly untouched by this mass movement of population. The resulting settlement pattern in 1910 is depicted in Figure 5-1.[14]

The north and west of the country attracted the most migrants. Southern states remained, by contrast, largely inhabited by American-born residents. Migrants established themselves in the rural areas of Illinois, Iowa, Michigan, Minnesota, and Wisconsin (Figure 5-1) and further to the west in sparsely populated areas, as well as in southern Texas and the southern tip of Florida in the south. Cities, especially in New England and the Atlantic states as well as Chicago, became big magnets for

[14] The situation in 1880 and 1900 was roughly similar. Maps are available upon request.

migrants, especially for those entering the country during the second migration wave between 1890 and 1920.

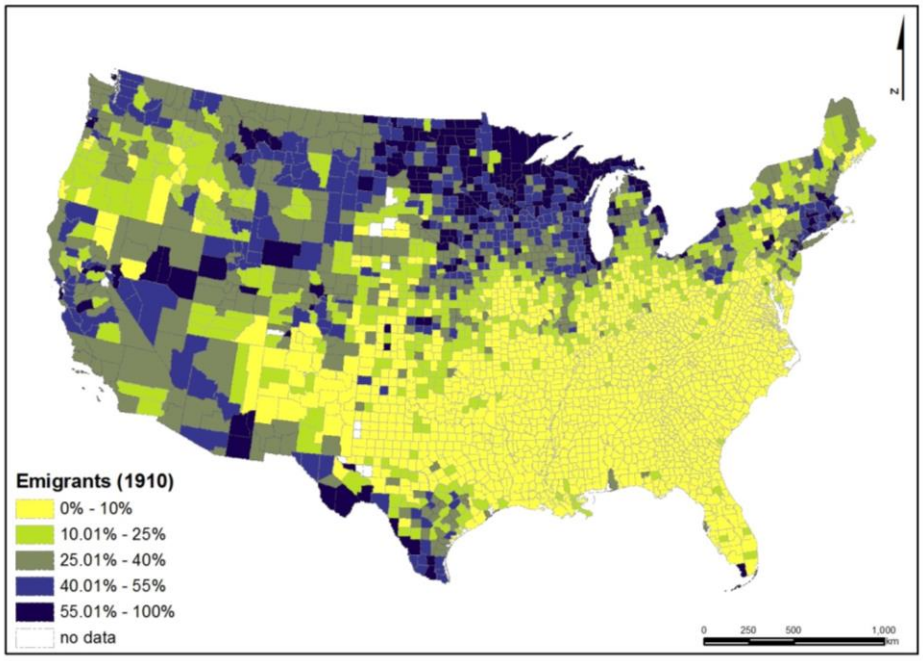


FIGURE 5-1 INTERNATIONAL MIGRANTS AND THEIR CHILDREN, 1910 (AS SHARE OF POPULATION BY COUNTY)

Source: Rodríguez-Pose and von Berlepsch (2014)

While migration from abroad rose rapidly, changing the population composition of large parts of the US, internal migration also picked up speed, reaching exceptionally high geographical population mobility levels. At the end of the 19th century, almost 60 percent of the male US population above the age of 30 had moved across county or state lines and almost a third of those born in the US lived outside their place of birth (Haines, 2000; Ferrie, 2005). Similar to international migrants, American-born individuals moved westwards in search of land to expand wheat, corn, wool, and meat production (Atack et al., 2000).

Most internal migrants of the late 19th century, however, travelled only short distances, with the majority remaining within their state of birth. Twenty percent, however, covered much larger distances, in some cases up to 4,500 km (own calculations). Figure 5-2 depicts their settlement pattern in 1910. The resulting map reveals a different pattern of American-born internal migration than that of international migration. Internal migrants mainly moved from east to west, settling in many states west of the Mississippi (with the exception of Utah, New Mexico, Texas, and parts of California). The majority of the population of mid-western states, such as Oklahoma, Wyoming, Montana, Oregon, Nevada, and Arizona, was thus made up of internal migrants. The entire eastern and southern part of the country (including the growing migrant agglomerations in the eastern cities) remained, with the exception of Florida, outside of internal migration routes.

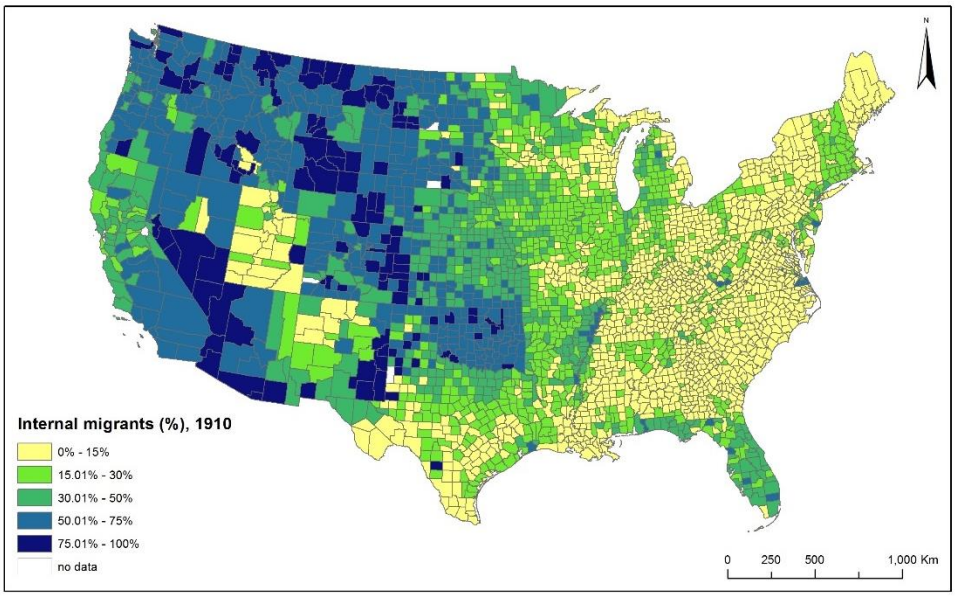


FIGURE 5-2 AMERICAN-BORN INTERNAL MIGRANTS AS SHARE OF POPULATION BY COUNTY, 1910

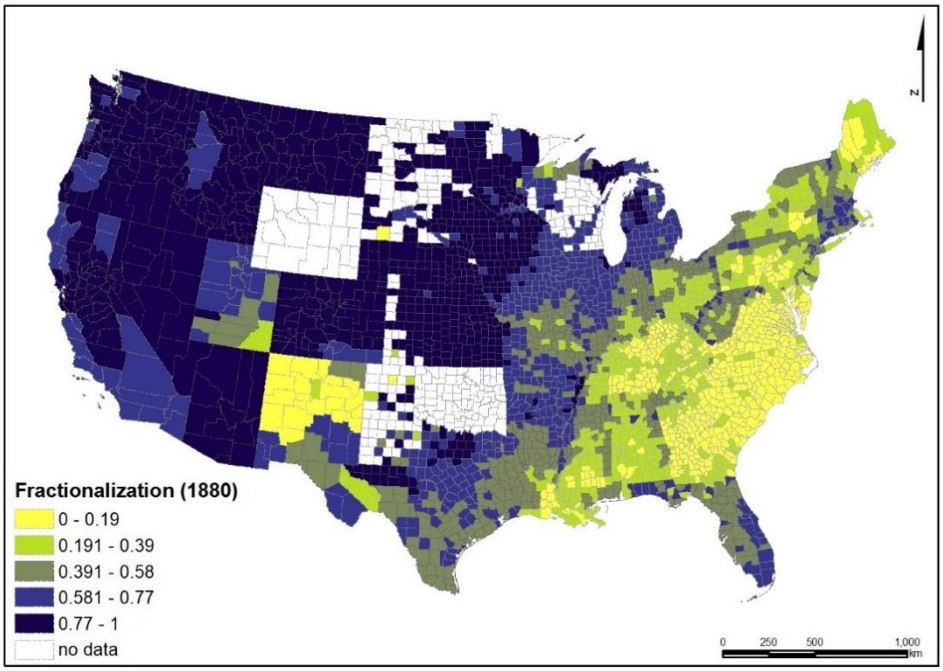
Data source: Ruggles et al., 2015; own elaboration

Both international as well as internal migration movements drastically affected the population composition of the US. In parts of the north-western states, for example, within a few years the population changed from being almost entirely local-born to rates of 90 percent or more having been born in other US states or abroad. Internal migrants originating from locations often thousands of kilometres away were as foreign to the local population as the Germans, Irish, Italians, or Poles settling within the same county. Whilst their language was the same – as was the case for migrants from the British Isles – internal migrants brought habits, customs, traditions, and a lifestyle which was regarded as outlandish and strange by the local population (Merck, 1978).

Some areas of the US were predominantly settled by one or two specific nationalities (Rodríguez-Pose and von Berlepsch, 2015), whereas other regions attracted a multitude of migrants stemming from all over the US as well as from a variety of different countries, leading to high levels of population diversity. Figure 5-3 displays the levels of population diversity – proxied by the widely-used index of fractionalisation, which emphasises the number of different groups within a population – across US counties in both 1880 (a) and 1910 (b).[15] High levels of population diversity became the norm primarily in the west of the country (with the exception of parts of New Mexico, Texas, and Utah), while huge swaths of the old South remained demographically homogeneous. Cities in the North East, such as New York City and Boston, hosted vibrant, mixed migrant communities. By contrast, other areas in the North East, such as Maine, Vermont, or parts of upstate New York, were characterised by low population diversity levels generally ranging between 0 and 0.3.

[15] The calculation of the fractionalisation index will be explained in detail in section 5.4.

a)



b)

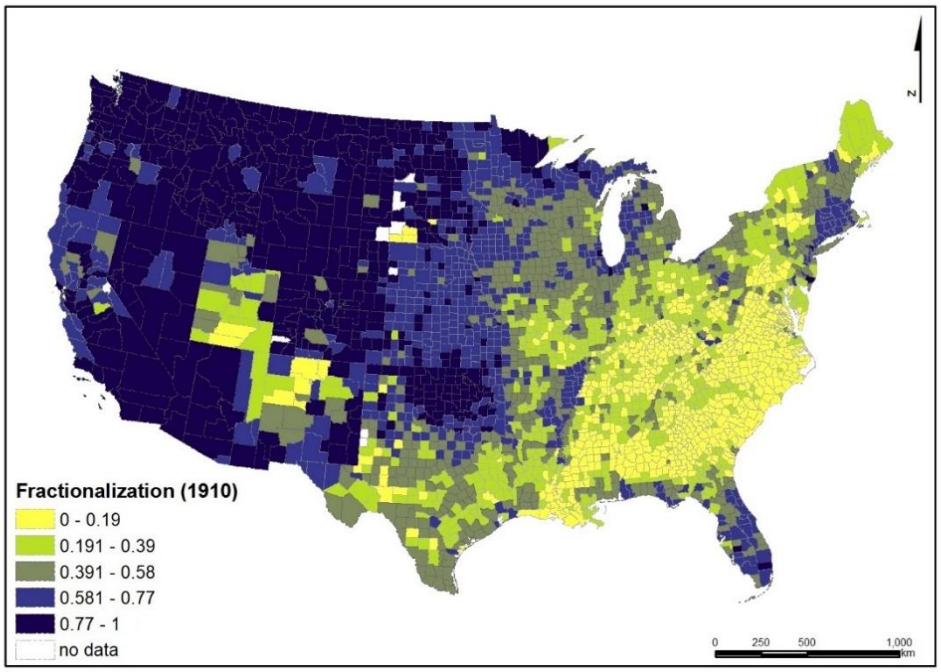


FIGURE 5-3 DIVERSITY IN THE COMPOSITION OF THE POPULATION BY COUNTY, 1880 AND 1910

Data source: Ruggles et al., 2015; own elaboration

5.3 DIVERSITY AND ECONOMIC DEVELOPMENT

With both international as well as internal migrants arriving and moving around the country, population diversity across the US drastically changed (Collier, 2013). How such a rapid shock to diversity levels has affected ensuing economic development is therefore a highly relevant question. Whether population diversity leads to higher or lower growth has turned into a widely discussed and often controversial topic in the theoretical and empirical literature across a wide range of disciplines, ranging from sociology and anthropology to political science, demography, geography, and economics. Overall, conclusions are far from clear cut due to a mixture of different indices, changing geographical units, and varying aggregation levels.

Two opposing strands dominate the debate – one depicting diversity as growth enhancing, the other as growth reducing. As the definition of diversity is far from straightforward, the strongly differing views primarily result from the respective dimension of diversity examined. Both strands choose entirely different angles from which to evaluate the link between diversity and economic development. Consequently, a variety of indices are used as a proxy of diversity, with each indicator measuring a different aspect of the notion. The most popular proxies used are measures of population fractionalisation on the one hand, and polarization and segregation on the other. Hence, whether diversity fosters or deters growth strongly hinges on the indicator employed.

The strand of research which views diversity as growth enhancing generally regards it as the central driver of innovation and creativity, which in turn fosters technological progress and growth. Migrants arriving from diverse locations are depicted as an important input factor in the process of technological progress. They bring in different

skills, ideas, experiences, and abilities to their places of destination. However, the speed of technological progress fuelled by the inflowing population does not depend on the size of the influx but on their composition, transforming diversity into a productivity enhancing and innovation initiating factor (Bove and Elia, 2017).

The connection between diversity and innovation dates back at least to Jacobs (1961, 1969). For her, environments which are characterised by the presence of a large variety of cultural groups provide more fertile soil for new ideas. Within these idea breeding grounds, new innovative concepts can spread more easily to different areas compared to more homogenous places, thereby fostering innovation and growth. Florida's creative class model (2002) supports this line of argument. As skilled, liberal people prefer to live in diverse regions, skilled jobs and innovation will cluster in these same areas. The 'New Argonauts' theory developed by Saxenian (2006) is yet another example of diversity leading to innovation. The concept revolves around foreign-born, technically high-skilled entrepreneurs, travelling back and forth between their home countries and Silicon Valley, boosting economic activity both in the once peripheral regions of their home countries, as well as in the US. Lazear (1999) draws a parallel to a firm context analysing the globalisation of firms. He finds that skill complementarity in a team spanning multiple cultures is key to not only offset the potential costs of diversity, but to significantly raise overall firm productivity. The interaction of a multitude of people with different abilities, ideas, and experiences triggers innovation, technological process, and hence growth.

Empirical research has tended to validate this view. Niebuhr (2010) shows that across Germany patent applications increase in proportion to labour force diversity. Özgen et al. (2011a) find that levels of innovation rise with the degree of diversity in the migrant community across European countries. A diverse labour force and immigrants

originating from a wide range of countries "not only contribute to innovation by means of their high skills and innate abilities, but […] they also bring into firms and host countries new ideas and perspectives from their different cultural backgrounds" (Özgen et al, 2011b: 1). An enlarged diversity of national origins is also linked to improvements in problem-solving, new combinations of ideas, and innovation (i.e. Hong and Page 2001, 2004), while inter-ethnic ties contribute to increased socioeconomic status of migrants (Riedel, 2015). Alesina et al. (2016) report that the productive effects of increasing population diversity are largest for high-skilled migrants and for migrants stemming from wealthier and more culturally similar source countries.

Ottaviano and Peri (2006) demonstrate a significant positive and robust correlation between both wages and rents with regional immigrant diversity in US metropolitan areas, emphasising that a more multicultural environment increases the productivity level of US-born citizens. From a slightly different angle, other studies portray diversity as productivity enhancing not only in regions or cities but also in work establishments. The enlarged skillset of the workforce as well as the interaction of diverse work teams with each other facilitate the production of a larger variety of goods and services and raises labour productivity levels, even when holding average ability constant (i.e. Alesina and La Ferrara, 2005; Hamilton et al., 2003; Trax et al., 2015; Kemeny and Cooke, 2017).

Two common denominators link the above diversity promoting studies. Firstly, the majority place emphasis on the subnational, granular level, evaluating either the impact of diversity at a regional, city, or even individual level. Secondly, and more importantly, this strand of research generally considers diversity on the basis of the number of different population groups – varying by language, religion, or ethnicity –

within a territory. They tend to use an index of fractionalisation – such as Alesina et al.'s (2003) – as the measure of population diversity. This type of index presupposes that the greater the number of groups, the higher the assumed diversity in a society, positively influencing the potential for growth. The groups' size or the distance between them does not enter the calculation of the most frequently used indices.

The strand of research positing that diversity has a negative influence on economic development follows a different line of thought. Rather than considering the positive influence of diversity on idea generation, innovation and productivity, it views the presence of diverse groups as a destabilising factor within a society, enhancing the potential for social unrest and conflict. This body of thought not only takes into account fractionalisation as a proxy for diversity but also increasingly utilises indices of segregation and polarisation.

"When the society is divided by religious, ethnolinguistic, or race differences, tensions emerge along these divisions" (Montalvo and Reynal-Querol, 2005a: 308). Ethno-linguistic fractionalisation has thus been inversely linked to per capita GDP and growth in large cross-country samples (e.g. Easterly and Levine, 1997; Alesina et al. 2003; Churchill and Smyth, 2017). Alesina et al. (2003), for example, found that a difference in economic growth of 1.9 percentage points between a wholly homogenous and a wholly heterogeneous society. The poor economic performance of African countries has been, for example, blamed on ethnic conflict resulting from high levels of national or ethnic polarisation (Easterly and Levine, 1997).

Various channels have been identified as vehicles through which diversity hinders economic development. Gören (2014) emphasises the negative direct effect of diversity on economic growth and considers polarisation as having indirect negative

economic effects via human capital, investment, openness, and civil war. Easterly and Levine (1997) argue there is a reduced probability of adopting 'good policies' in more polarised societies. According to their study, low school attainment, high financial debt and low infrastructure quality are all consequences of high segregation levels. Moreover, diversity is believed to foster rent seeking behaviour by different groups, further undermining the potential for adopting sound public policies. Overall, high polarisation triggers "positive incentives for growth reducing policies, such as financial repression and overvalued exchange rates, that create rents for the groups in power at the expense of society at large" (Easterly and Levine, 1997: 1206).

More fragmented societies are found to curb public sector performance and to generate poor institutions (La Porta et al., 1999; Mauro, 1995; Easterly et al., 2006), leading to regional disparities (Ezcurra and Rodríguez-Pose, 2013), an inefficient provision of public goods and services, a reduction in government transfers and distortionary taxation (Desmet et al. 2009; Azzimonti, 2011), political instability (Alesina et al., 1999; Alesina et al., 2003; Baldwin and Huber 2010), as well as reduced property rights security (Keefer and Knack, 2002), and low quality of government (Alesina and Zhuravskaya, 2011). Enhanced heterogeneity may even lead to the formation of xenophobic political parties (ibid), undermine collective action, and reduce the efficiency of regulation (Baland and Platteau, 2003; Platteau and Seki, 2007).

Diversity is further shown to impact political rights, adversely affecting economic growth (Collier, 2001). Particularly in less democratic societies, polarisation can curtail individual rights and limit overall economic performance (Bluedorn, 2001; Alesina et al., 2003). Further consequences of highly diverse and polarised societies are a reduction of trust and social participation, inefficient communication, less economic integration, lower voting turnout, and a rise in transaction costs for bridging

cultural differences (i.e. Ancona and Caldwell, 1992, Alesina et al., 1999, Alesina and La Ferrara, 2000, Richard et al., 2002; Van Knippenberg and Schippers, 2007; Alesina and Zhuravskaya, 2011; Uslaner, 2011; Mavridis, 2015; Martinez i Coma and Nai, 2017). The resulting rent seeking behaviour leads to slower growth, lower production, reduced investment, and diminished prosperity (Rodrik, 1999; Alesina and La Ferrara, 2005, Montalvo and Reynal-Querol, 2005b). "In extreme cases, diversity can prompt large scale social and economic collapse, sometimes with horrific consequences, as has occurred in recent years in parts of Central Africa, the Balkans, and elsewhere" (Kemeny, 2012: 2136). Highly fragmented societies have been deemed prone to moderate intensity conflict. In highly polarised societies, conflict can be less frequent but of higher intensity (Esteban and Ray, 2008). The likelihood and frequency of civil wars – an extreme example of social collapse – have been associated with high population diversity in terms of polarisation (i.e. Horowitz, 1985; Elbadawi and Sambanis, 2002; Montalvo and Reynal-Querol, 2005a,b).

Focusing on regional data, these results have been held up by a number of studies analysing the case of the US. High diversity in US communities has been connected to a less efficient provision of public goods, lower trust, and less social participation (i.e. Alesina, et al., 1999; Alesina and La Ferrara, 2000; Luttmer, 2001; Alesina and La Ferrara, 2002). Diversity has also been considered a strong and persisting barrier to developing trust across racial, ethnic, or national origins (Glaeser et al., 2000).

Again, a string of common denominators links the above studies. Firstly, within this strand of the literature and with few exceptions, studies tend to use nations as the unit of analysis. Secondly, diversity is increasingly referred to as triggering the negative effects of polarisation and segregation. Different indices have been employed by the literature in order to capture this effect. One of the most commonly used indices,

proposed by Esteban and Ray (1994), finds its roots in the social tension literature. Here, indices measure entirely different aspects of diversity than fractionalisation. Rather than focusing on the number of groups within a population, polarisation indices emphasise their relative size to one another and the distance separating them. The bigger the distance among groups, the more similar their size, and the stronger the lines separating them – the smaller the capacity to communicate and hence the larger the negative impact of diversity on economic development. According to Montalvo and Reynal-Querol (2005a), social unrest is further aggravated if the population is distributed into two separate groups of similar size. Examples of this latter phenomenon would be, amongst others, Northern Ireland or the Basque Autonomous Community.

In short, cultural diversity affects trust among the inhabitants of primarily multinational, multi-ethnic and multi-religious countries. It upsets the coordination of actors and their communication, generating animosity, enlarging differences in preferences and creating situations of conflict. Simultaneously, however, this multitude of ideas, experiences, skills, and abilities can foster technological innovation, create a fertile soil for new ideas, increase productivity levels, and therefore enhance the supply and the quality of goods and services. By influencing both human capital and the process of technological progress, diversity has an undeniable impact on economic growth, although its net effect remains unclear (Bove and Elia, 2017).

One aspect has, however, been largely neglected in all of the aforementioned literature and still needs to be evaluated: the dynamic economic impact of diversity over time. Hence, while there is significant controversy about how fractionalisation and/or polarisation matter for economic growth, to the best of our knowledge, we know only

very little about whether higher or lower initial levels of diversity – regardless of measurement – affect growth differently in the short, medium or very long-term.

Research that examines the time dimension of diversity is limited. Campos and Kuzeyev, (2007) or Campos et al. (2011) analyse changing levels of polarisation over time and evaluate their short-term impact on growth. Both find a negative effect of polarisation. Alesina et al. (2016) and Ager and Brückner (2013a) consider timeframes of 10 and 50 years respectively, evaluating a sample of 120 countries with panel data between 1990 and 2000 (the former) or use a within-county estimation approach for US counties that evaluates the impact of the change in cultural composition over the course of 50 years (1870-1920) on economic growth (the latter). Both find fractionalisation to be positively related to economic prosperity, while polarisation has the opposite effect. However, neither assesses the impact of a fixed initial level of diversity on economic performance across alternating time horizons. Furthermore, while Ager and Brückner (2013a) base their study on the same historical timeframe as used in this chapter, they do not extend their analysis to present levels of economic development.

Studies that come closest to analysing a dynamic effect of diversity over longer time horizons are rare. They include Ager and Brückner (2013b) and Bove and Elia (2017). The former report a significant short- and long-term impact of initial diversity levels on economic development in the US. However, they refer to the use of genetic diversity based on Ashraf and Galor (2013), rather than including the two most frequently discussed proxies of diversity: fractionalisation and polarisation. Bove and Elia (2017) identify a positive association of both fractionalisation and polarisation with real GDP per capita when evaluating a 135 country sample over a 50 year timeframe. The positive link of both indicators – consecutively added to their model

– is significant over the long-term, but fails to retrieve consistent results over the short, 10-20 year, timeframe.

In short, to the best of our knowledge, there is no scientific research treating both dimensions of diversity – fractionalisation and polarisation – which links historical population diversity levels to current economic development and covers a period longer than 50 years. Thus, some key questions remain unanswered: Does diversity, in its two fundamental dimensions of fractionalisation and polarisation, affect growth – if at all – differently in the short- than in the medium- and long-term? Does a high degree of fractionalisation and/or polarisation generated more than a century ago promote growth in the short-run but limit it in the long-term? Or is it vice versa?

5.4 EMPIRICAL APPROACH

The aim of this chapter is precisely to fill this important gap in the literature by examining the extent to which the levels of initial diversity, defined by both fractionalisation and polarisation, generated during the Age of Mass Migration across US counties a) have left a long-lasting economic legacy that can still be identified in the economic development of US counties today and b) whether any positive or negative influence of initial diversity on economic development has waxed or waned with time.

Based on the previous discussion, we adopt two econometric models in order to test our two research questions: one focusing on population heterogeneity, the other on population homogeneity. Following the relevant literature, we employ place-of-origin fractionalisation and polarisation – the two most commonly used indices – to depict

population diversity in Model 1. Its almost opposite, place-of-origin concentration, is used to reflect population homogeneity in Model 2.

We expect both dimensions of diversity – fractionalisation and polarisation – to matter for economic development over very long timeframes. We not only assume that the growth influencing traits of diversity become embedded in the local mentality, traditions, and customs – in short, in local institutions – but also that big diversity shocks in a given period of time can become etched in the core characteristics of a territory and thus persist over extended timeframes.

The implications of this assumption are twofold. First, US counties having received large inflows of both international and internal migrants stemming from a multitude of different origins more than a century ago should be significantly more prosperous today than those which displayed a more homogeneous population composition at the time. Second, we expect US counties marked by a highly polarised population composition during the Age of Mass Migration to have faced considerable barriers to the development of economic activity, deeply limiting their growth potential. Consequently, we assume historical fractionalisation to be positively connected to current income levels across US counties, while historical degrees of polarisation are likely to be negatively and significantly associated to them.

Moreover, in line with Ager and Brückner (2013b), we hypothesise that time will not significantly alter the impact of diversity on economic development. We assume a highly fragmented (highly polarized) society to retain its positive (negative) impact consistently over the short-, the medium- and the long-term. Despite the fact that international migrants become American and internal migrants adopt local traits over time, their cultural baggage brought from their place of origin remains with them and is passed not just to the following generations, but especially to their chosen place of

residence. As diversity becomes embedded in the core character of the county, it permanently influences its subsequent economic development path for decades.

5.4.1 MODEL 1 – POPULATION HETEROGENEITY: THE CASE FOR DIVERSITY

Model 1 is concerned with diversity measured as fractionalisation and polarisation. The model adopts the following form:

$$y_{i,t} = \alpha + \beta Fract_{i,t_0} + \lambda Pol_{i,t_0} + \partial X_{i,t-k} + \theta Z_{i,t_0} + \mu state + \varepsilon_{is}$$

where y represents the income per capita of county i in period t (t=2010, 2000, ..., 1900); *Fract* is the level of fractionalisation in a given county i in t_0, which corresponds to either 1880, 1900, or 1910; *Pol* is the degree of polarisation in a given county i in t_0; X is a vector of variables which are assumed to influence the level of development of any given county at time t-k (k=10); Z represents a similar vector of factors which may have influenced the development of the county at time t_0.[16] Lastly, *state* depicts state controls taking into account unobservable state specific effects and ε represents the error term clustered to the state level s to ensure robustness to arbitrary spatial correlation within one state. Our main coefficients of interest are β and λ describing the relationship of the two dimensions of diversity with economic development.

[16] In order to avoid multicollinearity problems, this vector is only included when analysing the long-term. For the short-term regression analysis, this vector is excluded.

5.4.2 MODEL 2 – POPULATION HOMOGENEITY: THE CASE FOR CONCENTRATION

In order to assess the robustness of our results, Model 2 resorts to an index of concentration to reflect population homogeneity as the main independent variable of interest. All other variables remain the same as in Model 1. In this alternative setup, the model adopts the following form:

$$y_{i,t} = \chi + \delta Conc_{i,t_0} + \varphi X_{i,t-k} + \eta Z_{i,t_0} + \vartheta state + \omega_{is}$$

where *Conc* is defined as the level of concentration within the population of any given county *i* in t_0 corresponding to either 1880, 1900, or 1910 and ω represents the error term clustered to the state level *s*. All other input variables refer to those presented in Model 1.

5.4.3 VARIABLES OF INTEREST – MEASURES OF DIVERSITY AND CONCENTRATION

Diversity: Following the two opposing strands of literature dealing with the link between diversity and economic growth, we resort to the two most commonly employed diversity indices to proxy population heterogeneity: fractionalisation and polarisation.

Fractionalisation (i.e. Easterly and Levine, 1997; Alesina et al., 2003) emphasises the number of different groups within a population. It goes back to the work by the Soviet researchers Bruk and Apenchenko (1964) who crafted an index of ethnic-linguistic

fractionalisation in the *Atlas Narodov Mira* (Atlas of the peoples of the world) based on the shares in total population of ethno-linguistic groups. The modified version of this index by Alesina et al. (2003) is used in this chapter as our first indicator of diversity:

$$Fract_{i,t_0} = 1 - \sum_{g=1}^{n} s_{g,i,t_0}^2$$

where $Fract_{i,t_0}$ is the degree of fractionalisation in county i at time t_0 where s depicts the share of total population of origin group g in county i at time t_0. This index "captures the probability that two randomly selected individuals belong to different groups" (Campos and Kuzeyev, 2007: 622). Hence, $Fract_{i,t_0}$ increases with the number of groups, taking on values between 0 and 1, with $1 - \sigma$ reflecting a highly fractionalised and $0 + \sigma$ a strongly homogeneous society, with $\sigma \to 0$.[17] If each person in a territory belongs to a different group, the index reaches its theoretical maximum.

Polarisation aims to capture the social tension and conflict dimension linked to a heterogeneous population. Esteban and Ray (1994, 1999) from a theoretical, and Montalvo and Reynal-Querol (2005a) from an empirical standpoint, argue that a highly polarised environment maximises the risk of conflict. The measure of polarisation is based on the family of indices developed by Esteban and Ray (1994, 1999), considering not only the number of ethnic groups within a society, but also the distances separating them and their individual size. According to this index type, the degree of polarisation within a population increases as the distance between groups rises, but also when the number of groups increase or when there is convergence in

[17] σ represents any small positive number so that $Fract_{i,t_0}$ becomes arbitrarily close to 1 or 0 respectively.

group size. As the calculation of distance between ethnic groups is highly controversial, we follow Reynal-Querol (2002) for our index, assuming the absolute distance between two groups to be equal and discrete.[18] The polarisation index in this case "measures the normalised distance of a particular distribution of ethnic […] groups from a bimodal distribution" (ibid: 301) and is maximised when two highly distinguishable groups of equal size coexist within the same population.

The polarisation index takes on the following form:

$$Pol_{i,t_0} = 1 - \sum_{g=1}^{n} \left(\frac{0.5 - s_{g,i,t_0}}{0.5}\right)^2 * s_{g,i,t_0}$$

where Pol_{i,t_0} is the degree of polarisation in county i at time t_0 and s depicts the share of total population of origin group g in county i at time t_0. Within this particular specification, it is the size of the groups relative to each other that is of particular importance.

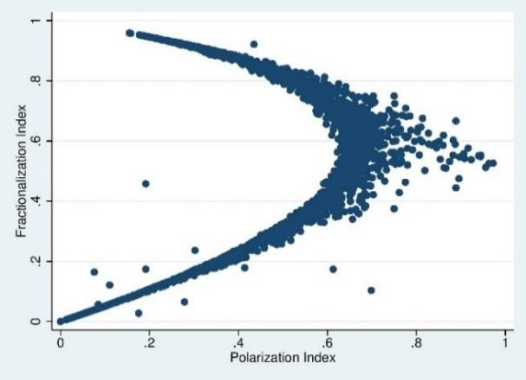
Both indices used in the analysis are aligned to the specifications of our dataset. Instead of ethnic groups as generally used in the literature, we consider the birthplaces of individuals – as defined by the US Census – living in a given county at t_0 as an indicator for different cultural groups. Birthplaces include both European countries as well as American states in order to properly account for international as well as for the high degree of internal migration prevailing at the time. As the historical US Census did not record the county of birth, but solely the state, our indicator does not pick up the bulk of the short distance, intra-state internal migration. Only population groups of internal migrants which travelled large distances leaving their home state are

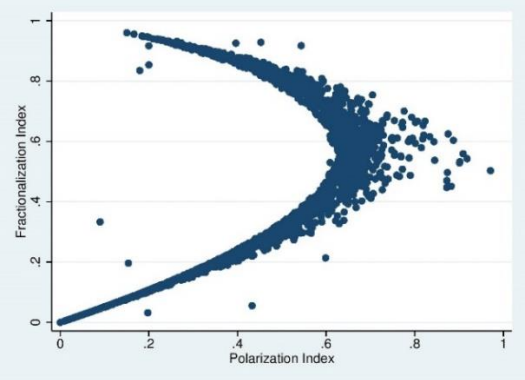
[18] One of the few papers estimating the distance between ethnic groups is Fearon (2003). Language is used as a proxy of cultural distance.

therefore included into the calculation of the index. We therefore take into consideration only the fraction of internal migrants which were identified to be significantly different to the local population (i.e. Merck, 1978).

Figure 5-4 plots the relationship between fractionalisation and polarisation in US counties for all three base years: 1880, 1900, and 1910. It is important to note that varying the size of both indices does not reveal a consistent interdependency. Conditional on the degree of fractionalisation, the extent of the correlation with polarisation varies. Both indices are highly positively correlated at low levels of societal diversity, indicating that adding a further cultural group to an otherwise perfectly homogenous population increases the risk of polarisation and conflict. However, as cultural heterogeneity increases, the positive relationship wanes and becomes irrelevant at medium levels of both fractionalisation and polarisation. The more a population becomes fragmented, the lower the societal standing and influence of a single population group, which reduces the societal polarisation within a given county. At highest levels of fractionalisation, the relationship between both indices turns strongly negative. Once above a certain fractionalisation threshold, the addition of further cultural groups to a population significantly decreases the risk of polarisation. This relationship is consistent across all three base years and in line with previous findings by Montalvo and Reynal-Querol (2005a, b), Ager and Brückner (2013a), or Bove and Elia (2017), underlining the validity of the data.

1880 **1900**





1910

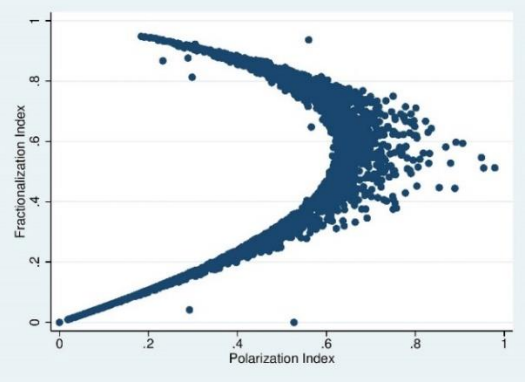


FIGURE 5-4 FRACTIONALISATION VERSUS POLARISATION FOR ALL THREE BASE YEARS

Data source: Ruggles et al., 2015; own elaboration

It is important to stress once again that both indices measure entirely different dimensions of diversity. While one focuses on the number of cultural groups leading to innovation, the other stresses their relative size to another, both provoking social unrest and conflict. Based on these highly distinct theoretical concepts, both indices thus identify independent and distinguishable effects of diversity on economic growth. From a theoretical standpoint, their joint inclusion in our empirical model minimises the risk of omitted variable bias and allows us to capture a more accurate and encompassing effect of the multidimensional notion of diversity on economic growth.

Concerns about the joint inclusion of both variables are addressed from an empirical standpoint in Figure 5-4, which rules out the risk of biased results due to correlation issues. Following Ager and Brückner (2013a), Alesina et al. (2003), Montalvo and Reynal-Querol (2005a), and Gören (2014), we include both indices of fractionalisation and polarisation in our model, as both indices capture a different aspect of diversity.

Concentration: The opposite of diversity is concentration, understood as the marked dominance of one group (based on place of origin) in a given territory. We employ this alternative variable of interest to assess the robustness of the results when analysing the diversity indices. The concentration index is defined as follows:

$$Conc_{i,t_0} = \max(s_{g,i,t_0})$$

where $Conc_{i,t_0}$ is the degree of concentration within the population of county i at time t_0 and s depicts the share of total population of origin group g in county i at time t_0. According to this definition, the index always takes on the population share of the largest represented birthplace group within the population of the particular county i, thereby indicating the degree of concentration within a territory.

5.4.4 CONTROLS – FACTORS INFLUENCING COUNTY DEVELOPMENT

We introduce two sets of control variables into our model. The first group of control variables included is vector Z dating from the time of migration – 1880, 1900, and 1910 – and consists of factors which influenced a county's development at the time of the big migration waves. The controls comprise mean income (as natural log), total population (as natural log), literacy rate, unemployment rate, female participation rate

in the labour force, share of black population, and the percentage of workers employed in agriculture. As these parameters are bound to have influenced the settlement decision of the individual migrants (see i.e. Jennissen, 2003), we can assume they would also have exerted a strong impact on fractionalisation and polarisation at county level in the late 19th and early 20th centuries. Furthermore, if certain population groups predominantly settled in wealthy regions and if this initial prosperity persisted over time, excluding the initial endowment factors of a county would lead to omitted variables and therefore biased estimates.

The second set of control variables in our model, vector X, represents the t-k time dimension, which corresponds to 10 years prior to the period considered in the dependent variable. Again, we control for factors influencing the economic development of the county, such as population size (as natural log), educational attainment, female participation in the labour force, unemployment, the share of black population, infant mortality, and the share of the labour force employed in agriculture. We shift X by 10 years in order to reduce the risk of reverse causality between the control variables and income per capita. As we will show later within this chapter, the results prove to be robust and stable throughout.

5.4.5 THE DATA

For the construction of the dependent variable, we employ income per capita data extracted from the US Bureau of Economic Analysis (BEA) database and the Current Population Survey tables (CPS) of the US Bureau of Labor Statistics (BLS), measured in US dollars. As income per capita was only available for the years 1950 onwards,

we resorted to a proxy for the years 1900 to 1940 and used either the salary income (1940) or calculated an aggregated mean income at county level constructed using the median total income per occupation in 1950 dollars (1880-1930). The construction of these variables uses the total population including all non-participants within the labour force as base in order to remain as comparable as possible with the income per capita variable of later years. The necessary input data for these proxies was extracted from the Integrated Public Use Microdata Series USA database (IPUMS) Version 6.0 (Ruggles et al., 2015). This database provides US microdata covering the censuses and American Community Surveys between 1790 and 2010.[19] We use the natural log of income as the dependent variable.

The main independent variables of interest, fractionalisation and polarisation as well as concentration, were built using the birthplace data at county level of the years 1880, 1900, and 1910, extracted from the IPUMS USA database. The birthplaces of a weighted population sample of 5,791,531 individuals in 1880, 3,852,852 individuals in 1900, and 923,153 individuals in 1910 were aggregated and allocated to the counties of residence of the individual. As the number and size of US counties changed over the period of analysis (2,875 counties in 1880, 3,090 in 1900 and 3,123 in 1910), we matched counties at the time of migration to their 2010 equivalent using US Census Bureau cartographic boundary files of the 48 continental states for each decade included in the analysis.

Data for the control variables were extracted from the IPUMS USA, the US BEA, the US BLS, the Inter-University Consortium for Political and Social Research (ICPSR) database, as well as from the Centers for Disease Control and Prevention (CDC)

[19] The American Community Survey was only initiated in 2005.

databases. In cases involving microdata, the data for individuals was aggregated at the county level. With the exception of mean income and educational attainment, all variables followed the same calculation method based on the same available data points across all years in question. The variable expressing the aggregated mean income at county level in the late 19th and early 20th century is constructed similarly to our dependent variable on the basis of the median total income per occupation (in 1950 dollars). Educational attainment is proxied by the percentage of people completing their college education for the years 1940 to 2000. From 1880 to 1930 we used the literacy rate per county as educational variable. A description of all variables is given in Appendix 5A.

5.4.6 INSTRUMENTAL VARIABLE ESTIMATION

Several endogeneity issues may arise when dealing with long-term migration data. While diversity may affect local GDP per capita, it is also likely that a higher GDP itself attracts more migrants, thereby increasing the level of diversity in the region. Consequently, the direction of causality remains ambiguous: GDP per capita on a regional level may be a function of diversity just as local diversity may be a consequence of local wealth. Moreover, when working with migration data, non-random spatial patterns in the distribution of migrants across space are likely to appear. Regional spillovers in migration may therefore generate clusters of counties with high levels of diversity. This spatial sorting would lead to endogeneity issues in our OLS regressions due to omitted variables. In order to address these endogeneity issues, we resort to instrumental variable (IV) estimation methods with the aim of revealing the true underlying effect of past diversity levels on income levels over time and to ensure

the validity of the least-squares estimations. We employ a shift-share methodology following Card (1999), Ottaviano and Peri (2006), and Saiz (2007). This instrument computes the estimated population composition of a county in 1880, 1900, and 1910 based on the population composition in a previous base year[20] and the US growth rate per population group between base and target year. This implies using the share of inhabitants per birthplace within the population of a county in the base year and multiplying this share by the growth rate of that particular group within the US population for the timeframe between base year and 1880, 1900, or 1910. Hence, we extrapolate predicted population shares under the assumption that migrants settle in areas where their predecessors had already established themselves. With these calculated predicted population shares, we then estimate an imputed fractionalisation, polarisation, and concentration index for each county in the respective target year.

The use of the shift-share instrument is based on the assumption that highly diverse counties in the earlier years of the big migration waves developed a diversity buzz which became a pull factor for new migrants. With the use of the shift-share instrument, we assume these highly diverse counties to have remained attractive to incoming migrants in the following decades, also implying that any changes in the degree of diversity at county level would have been independent of county specific shocks that may have taken place within the timeframe in question.

The results of the Staiger and Stock (1997) test for weak instruments using the first stage F-statistic of joint significance confirm that the shift-share variables for fractionalisation, polarisation, and concentration are all strong instruments. The Kleibergen-Paap Wald F statistics, in combination with the Stock and Yogo (2005)

[20] Diversity levels in 1880 are instrumented by the shiftshare using the population composition in 1870 as base. For 1900 and 1910 values, we used 1880 base values due to the significantly larger data sample available.

critical values for tests of weak instruments, further support the validity of the instruments. The instruments reject the null of weak identification when testing at a nominal 5% significance level. Both the imputed polarisation and fractionalisation indices as well as the shift-share instrument for concentration are identified as strong across the three base years and the various time shifts of the dependent variable.

5.5 ANALYSIS

5.5.1 THE LONG-TERM IMPACT OF DIVERSITY

The analysis starts with an evaluation of the long-term impact of diversity. Our first research question – whether population diversity levels generated in 1880, 1900, and 1910 are connected to county-level income per capita 100 to 130 years later – is first assessed by means of an ordinary least squares regression. As mentioned in the empirical approach, the model controls for wealth influencing factors both at the time of migration and in recent years and includes state controls in order to control for state specific unobserved factors influencing the counties' prosperity. Table 5-2 reports the results of Model 1 for our two main variables of interest, fractionalisation and polarisation, for 1880, 1900, and 1910 with respect to income per capita levels in 2010.

The results in Table 5-2 unveil a positive long-term connection of country-level population diversity with current GDP per capita. The fractionalisation index displays positive coefficients with significance levels below 1% across all three base years. The presence of large numbers of different groups according to place of origin in one county during the age of mass migration is strongly associated with higher levels of income in that county 100 to 130 years later. Polarisation, by contrast, remains insignificant across all three base years. Hence, polarisation at the height of the big migration waves appears unassociated with current levels of county wealth.

The signs and significance of the coefficients of the control variables reinforce the validity of the model, as they are in line with traditional studies on the determinants of growth. A good educational endowment in 2000 is connected to higher levels of

income per capita in 2010. Conversely, levels of unemployment and the percentage of black population – a proxy for poverty – in 2000 are linked to lower county wealth.

Of the factors that may have affected decisions to migrate more than a century ago, few are still connected to county levels of development in recent years. The one exception is the share of black population in 1880, 1900, and 1910. In all cases, counties with a higher percentage of black people at the turn of the century have significantly higher levels of income per head today. We assume this variable to serve as proxy for the economic structure of these largely agrarian counties back at the time of migration rather than indicating the effect of a divided county population into black and white. The coefficient of the share of black population is highly likely to capture the path to convergence of the poor regions of the South to the richer regions in the North (see for example Barro and Sala-i-Martin, 1990, 1995). The total population of a county in 1880, the level of unemployment in 1900, the proportion of agricultural employment in 1910, and the mean income in 1910 are negatively associated with income per head in 2010, while literacy in 1880 displays a positive and significant sign. Population diversity – measured as fractionalisation – hence proves to have a considerably stronger association with future income levels than the large majority of other base year controls. Put differently, the results suggest that a highly diverse population is a better indicator of future regional wealth than, in particular, county wealth at the time of migration.

TABLE 5-2 THE LONG-TERM IMPACT OF DIVERSITY, OLS 1880, 1900, AND 1910

Dep. variable: Income p.c. 2010 (ln)	1880 OLS	1900 OLS	1910 OLS
Fractionalisation ~	0.144***	0.176***	0.155***
	(0.0501)	(0.0474)	(0.0323)
Polarisation~	-0.0365	-0.0470	-0.0301
	(0.0411)	(0.0376)	(0.0308)
Education 2000	0.0125***	0.0125***	0.0123***
	(0.000782)	(0.000826)	(0.000822)
Total population 2000 (ln)	0.00145	0.00296	-0.00257
	(0.00566)	(0.00560)	(0.00562)
Share of black population 2000	-0.00133***	-0.000912*	-0.00141***
	(0.000462)	(0.000489)	(0.000493)
Female participation 2000	-0.000141	0.000245	0.000563
	(0.00117)	(0.00110)	(0.00107)
Unemployment 2000	-0.0247***	-0.0264***	-0.0246***
	(0.00461)	(0.00421)	(0.00452)
Infant mortality 2000	-8.41e-05	-0.000161	-0.000126
	(0.000322)	(0.000288)	(0.000286)
Agriculture 2000	-0.000330	-0.000405	-0.000333
	(0.00208)	(0.00224)	(0.00226)
Mean income (ln)~	-0.000603	-0.000442	-0.00535*
	(0.00333)	(0.00407)	(0.00285)
Literacy~	0.0976**	0.0368	0.0679
	(0.0395)	(0.0850)	(0.0499)
Total population (ln)~	-0.0120**	-0.0125	-0.00806
	(0.00498)	(0.00761)	(0.00834)
Share of black population~	0.219***	0.173***	0.209***
	(0.0459)	(0.0375)	(0.0435)
Female participation~	0.0319	-0.0287	0.000231
	(0.0895)	(0.0910)	(0.0513)
Unemployment ~	-0.00865	-0.0468**	-0.207
	(0.00959)	(0.0189)	(0.161)
Agriculture~	-0.0687	0.000508	-0.0963***
	(0.0531)	(0.0675)	(0.0204)
State controls	*yes*	*yes*	*yes*
Observations	2,825	3,024	3,094
R-squared	0.642	0.637	0.642

~ Variables date from respective year of migration 1880, 1900, or 1910
Note: Robust standard errors in parentheses, clustered at state level
*** p<0.01, ** p<0.05, * p<0.1

The results prove to be robust to the replacement of the diversity variables by a measure of group concentration, as indicated in Model 2 (Table 5-3). The concentration index is significant at the 1% level – as was the case of the fractionalisation index in Table 5-2 – although the association with income per capita

in 2010 is, as expected, negative. Hence, US counties with a more homogeneous population composition (dominated by one large group, regardless of whether the group originates from abroad or from a different American state) more than a century ago seem to have endured a substantially worse economic trajectory over the last 100 to 130 years than those which attracted a large number of people stemming from different birthplaces. In line with the relevant literature, one could speculate the largely homogenous population composition to have hampered the emergence of innovation boosting conditions linked to the buzz of diversity.

As far as both sets of control variables are concerned, there is nearly no change in either the significance levels or in the signs of the coefficients compared to those reported in Table 5-2.

TABLE 5-3 THE LONG-TERM IMPACT OF CONCENTRATION, OLS 1880, 1900, 1910

Dep. variable: Income p.c. 2010 (ln)	1880 OLS	1900 OLS	1910 OLS
Concentration ~	-0.158***	-0.175***	-0.149***
	(0.0486)	(0.0467)	(0.0329)
Education 2000	0.0125***	0.0125***	0.0123***
	(0.000777)	(0.000831)	(0.000822)
Total population 2000 (ln)	0.00196	0.00284	-0.00251
	(0.00572)	(0.00569)	(0.00576)
Share of black population 2000	-0.00133***	-0.000857*	-0.00141***
	(0.000458)	(0.000492)	(0.000496)
Female participation 2000	-0.000129	0.000223	0.000535
	(0.00115)	(0.00110)	(0.00108)
Unemployment 2000	-0.0249***	-0.0268***	-0.0247***
	(0.00476)	(0.00428)	(0.00462)
Infant mortality 2000	-4.60e-05	-0.000160	-0.000127
	(0.000320)	(0.000293)	(0.000290)
Agriculture 2000	-0.000157	-0.000259	-0.000287
	(0.00207)	(0.00227)	(0.00223)
Mean income (ln)~	-0.00108	-0.000371	-0.00504*
	(0.00342)	(0.00408)	(0.00289)
Literacy~	0.104**	0.0522	0.0754
	(0.0393)	(0.0926)	(0.0531)
Total population (ln)~	-0.0116**	-0.0111	-0.00585
	(0.00470)	(0.00758)	(0.00816)
Share of black population~	0.224***	0.172***	0.208***
	(0.0441)	(0.0389)	(0.0427)
Female participation~	0.0339	-0.0308	-0.00295
	(0.0927)	(0.0888)	(0.0514)
Unemployment ~	-0.00965	-0.0463**	-0.207
	(0.00924)	(0.0193)	(0.167)
Agriculture~	-0.0743	-0.00275	-0.0974***
	(0.0525)	(0.0614)	(0.0199)
State controls	*yes*	*yes*	*yes*
Observations	2,826	3,024	3,094
R-squared	0.643	0.636	0.641

~ Variables date from respective year of migration 1880, 1900, or 1910
Note: Robust standard errors in parentheses, clustered at state level.
*** p<0.01, ** p<0.05, * p<0.1

In order to address potential endogeneity issues due to the risk of omitted variable bias as a result of spatial sorting, reverse causality, or unaccounted economic shocks, an instrumental variable estimation is performed using the aforementioned shift-share methodology for all three main variables of interest: fractionalisation, polarisation (Table 5-4, columns 1, 2, and 3) and concentration (Table 5-4, columns 4, 5, and 6).

Table 5-4 reports a positive and strongly significant impact of high levels of fractionalisation in all three base years on income per capita levels in 2010, supporting the validity of previous results. In contrast to the OLS regressions, the polarisation index, while remaining negative, becomes significant at the 5% level for 1880 and the 10% level for 1900 and 1910 respectively. This proves that once we control for endogeneity issues and correct potentially biased estimators, diversity reveals its true underlying two dimensional long-term impact on income per capita levels. On the one hand, the presence of a large number of groups and, thus, considerable population diversity (high fractionalisation) is an important factor behind the long-term economic dynamism of places in the US, provided the diverse groups are not too polarized and, therefore, able to communicate with one another (low polarisation). By contrast, highly homogeneous societies have experienced much lower economic dynamism over the long-term (Table 5-4, regressions 4, 5, and 6). The signs and level of significance of the control variables remain virtually unchanged from those reported in Tables 5-2 and 5-3.

TABLE 5-4 THE LONG-TERM IMPACT OF DIVERSITY AND CONCENTRATION, IV 1880, 1900, 1910

	(1)	(2)	(3)	(4)	(5)	(6)
Dep Var: Inc. p.c. 2010 (ln)	1880 IV	1900 IV	1910 IV	1880 IV	1900 IV	1910 IV
Fractionalisation~	0.371***	0.391***	0.271***			
	(0.0997)	(0.123)	(0.0810)			
Polarisation~	-0.175**	-0.165*	-0.100*			
	(0.0784)	(0.0941)	(0.0570)			
Concentration~				-0.375***	-0.389***	-0.261***
				(0.102)	(0.115)	(0.0683)
Education 2000	0.0125***	0.0126***	0.0123***	0.0124***	0.0126***	0.0124***
	(0.000782)	(0.000857)	(0.000806)	(0.000767)	(0.000880)	(0.000811)
Total population 2000 (ln)	0.00265	0.00107	-0.00327	0.00272	0.000932	-0.00302
	(0.00518)	(0.00594)	(0.00510)	(0.00532)	(0.00623)	(0.00540)
Black population 2000	-0.00160***	-0.00116**	-0.00145***	-0.00156***	-0.00106**	-0.00146***
	(0.000451)	(0.000479)	(0.000456)	(0.000437)	(0.000502)	(0.000467)
Female participation 2000	-4.24e-05	0.000852	0.000747	9.59e-05	0.000814	0.000752
	(0.00115)	(0.000995)	(0.000993)	(0.00111)	(0.00102)	(0.00100)
Unemployment 2000	-0.0227***	-0.0243***	-0.0248***	-0.0242***	-0.0255***	-0.0249***
	(0.00472)	(0.00390)	(0.00440)	(0.00480)	(0.00427)	(0.00461)
Infant mortality 2000	-0.000113	-3.90e-05	-0.000212	-1.49e-05	-5.21e-05	-0.000216
	(0.000336)	(0.000303)	(0.000287)	(0.000317)	(0.000308)	(0.000293)
Agriculture 2000	9.45e-05	-0.00109	-0.000798	0.000716	-0.000661	-0.000415
	(0.00220)	(0.00212)	(0.00228)	(0.00228)	(0.00235)	(0.00232)
Income (ln) ~	-0.00337	-0.00617	-0.00764**	-0.00367	-0.00634	-0.00749**
	(0.00364)	(0.00431)	(0.00333)	(0.00382)	(0.00448)	(0.00337)
Literacy~	0.0451	0.0221	0.0302	0.0694	0.0476	0.0523
	(0.0537)	(0.0783)	(0.0499)	(0.0492)	(0.0939)	(0.0536)
Total population (ln)~	-0.0171***	-0.0106	-0.0112	-0.0150***	-0.00818	-0.00769
	(0.00522)	(0.00929)	(0.00880)	(0.00496)	(0.00908)	(0.00822)
Black population ~	0.226***	0.196***	0.212***	0.254***	0.204***	0.218***
	(0.0411)	(0.0362)	(0.0403)	(0.0405)	(0.0380)	(0.0394)
Female participation ~	0.0533	-0.0315	0.0115	0.0278	-0.0338	0.00928
	(0.0875)	(0.0929)	(0.0519)	(0.0924)	(0.0903)	(0.0513)
Unemployment~	-0.00115	-0.0486**	-0.225	-0.00603	-0.0456**	-0.218
	(0.0128)	(0.0214)	(0.150)	(0.0105)	(0.0221)	(0.158)
Agriculture~	-0.0650	0.0238	-0.0958***	-0.0716	0.0344	-0.0913***
	(0.0543)	(0.0763)	(0.0196)	(0.0545)	(0.0667)	(0.0189)
State Controls	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Observations	2,817	2,826	3,067	2,820	2,827	3,069
First-stage F-statistic	22.63	36.95	99.28	87.19	74.10	213.74

Variables date from respective year of migration: 1880, 1900, or 1910
Robust standard errors in parentheses, clustered at state level.
*** p<0.01, ** p<0.05, * p<0.1

5.5.2 THE DYNAMIC IMPACT OF DIVERSITY

The second part of the analysis is dedicated to examining the dynamic impact of diversity on income levels. Starting with income levels in 1900, the dependent variable in Model 1 is changed each time by 10 years in order to account for potential changes

in the influence of original population diversity on income per head. For this part of the analysis, however, vector Z, including the base year controls, is dropped from the estimation in order to avoid issues of multicollinearity in the earlier years considered. By means of an ordinary least squares regression, followed again by two rounds of robustness checks, we seek to analyse if the impact of diversity on county-level income per head varies over time. The results of the analysis are reported in Table 5-5.

The results point towards an enduring and positive association between population diversity and local income levels in the US. With the exception of the 1920s to 40s, heavily affected by the great depression and both world wars[21], the link between fractionalisation and income per capita at the county-level remains positive and strong, with no evidence of a waning or shifting connection over time. As in Table 5-2, with the exception of one year in the 1900 base year regression, there is no significant connection between population polarisation and income levels.

To test the validity of these results, we conduct the same exercise substituting fractionalisation and polarisation by concentration levels within the county population (Table 5-5). In line with the previous long-term results (Table 5-3), a strong negative association between high levels of concentration and regional income levels emerges not only in the long-, but also in the short- and medium-term. Similarly, the effect of concentration is negative and significant over the whole 100-year timeframe considered, with the exception of 1920 to 1940 (see footnote 21).

[21] The 1920s to 1940s were heavily influenced by the Great Depression and both world wars. These macroeconomic shocks seem to have dominated the diversity effect to such an extent that the significance levels are lost during these years. Normal economic rules were suspended during these years with an economy more and more nationalised focusing on the production of war supplies and the arms and defence industry. Later, starting with the 1950s, with an economy slowly back to 'normal' times, the diversity variables pick up their significance levels again.

Both OLS regressions emphasise the importance of a county's population composition at the time of the great migration waves to the US for its subsequent economic development. The results suggest that counties which failed to attract a large variety of groups from different origins – both international as well as national – seem to have suffered negative economic consequences for more than a century, as indicated by the significantly lower income levels, than those counties that succeeded in establishing vibrant and diverse communities.

Is this dynamic connection purely an association or is there a causal relationship? To answer this question, we resort to the use of an instrumental variable estimation, using, once again, a shift-share instrument. The results for both Model 1 and 2 are displayed in Table 5-6.

Again, and with the exception of the period between 1920 and 1940 (see footnote 21 for a potential explanation), the results depict a strong and robust association across time between population diversity and regional income levels. As in the OLS estimations, the coefficient for the fractionalisation index remains, across all three base years, positive and highly significant at the 1 percent level. In contrast to the OLS regressions, the use of an IV estimation makes the coefficient of polarisation significant for the early (1900-1910) and later years (1960-2000) of the analysis. Polarisation has, as expected, a negative influence on county-level economic development in line with Alesina et al. (2003), Montalvo and Reynal-Querol (2005a, b), or Ager and Brückner (2013a). Wherever strong barriers across place-of-birth origins were evident among population groups at the time of the great migration, local development has lagged behind.

Moreover, we now find a dynamic effect related to the size of the coefficients. This is particularly evident in the case of the 1880 and 1900 base year regressions. Columns 1 and 2 (income in 1900 and 1910 as dependent variables) in Table 5-6 display coefficients up to almost 5 times larger than those presented in columns 10 or 11 (1990 and 2000). The results indicate that high levels of fractionalisation and polarisation in the composition of a population had a more powerful effect on income levels within the first 10 to 30 years, while, in the longer term, despite remaining significant, the extent of this effect becomes significantly smaller (Figure 5-5). Hence, a high degree of population diversity, generated by mass internal and international migration at the turn of the 20th century, seems to be at the origin of some sort of diversity *buzz*. We presume such local *buzz* to have produced fertile grounds for long-term increases in productivity and innovation (Jacobs 1961, 1969). But the impact of population diversity has not been constant over time. There are many possible ways to explain this result. We find the influence of diversity on county-level wealth to be particularly strong during the years when migrants were still economically active and kept the local population culturally diverse. We assume that as long as the different population groups remained clearly distinct from one another and immersed in the culture of their home countries and home regions, the economic impact of diversity remained high. We argue the assimilation of migrants and, especially, of their children into the American melting pot to have reduced population diversity and, consequently, attenuated its positive economic effects. As the cultural distance between previously highly different population groups might have decreased with adaptation to the American way of life, the economic premium linked to past local diversity seems to have waned. However, the positive effect of past diversity *buzz* did not disappear completely: formerly diverse counties remained more dynamic over time than counties

that had, by and large, stayed mostly homogeneous in their population composition. Hence, diversity shocks at local level seem to have triggered economic mechanisms that – one might argue - became engraved in the territory and have proved enduring – leaving traces that can still be detected more than a century after the initial shock took place. In spite of the assimilation of former migrants into American culture, the rapid 'Americanisation' of their children and the loss of local diversity over successive generations, high historical population diversity levels in the late 19th and early 20th century still affect current local economic development across the US. Diversity linked to migration has left a very long-lasting trace on local wealth which is still measurable in terms of higher average income levels today.

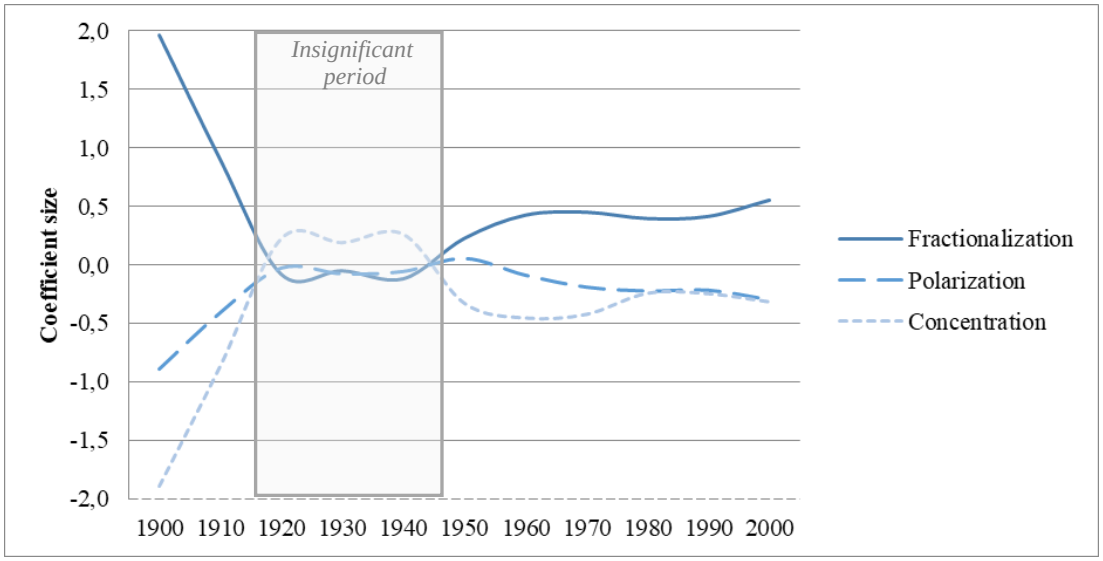


FIGURE 5-5 EVOLUTION OF COEFFICIENTS FOR FRACTIONALISATION, POLARISATION AND CONCENTRATION OVER TIME (IV, BASE YEAR 1880)

Source: Own elaboration

TABLE 5-5 THE DYNAMIC EFFECT OF DIVERSITY AND CONCENTRATION, OLS[22]

Var. of interest	Base year	Dep Var.: Income	(1) 1900	(2) 1910	(3) 1920	(4) 1930	(5) 1940	(6) 1950	(7) 1960	(8) 1970	(9) 1980	(10) 1990	(11) 2000
Diversity	1880	*Fractionalisation*	0.247***	0.199*	0.087	0.012	0.0042	0.161***	0.236***	0.233***	0.300***	0.282***	0.322***
			(0.057)	(0.114)	(0.073)	(0.064)	(0.069)	(0.0555)	(0.0477)	(0.0510)	(0.0668)	(0.0605)	(0.0691)
		Polarisation	-0.078	-0.028	-0.054	-0.046	-0.059	-0.0269	-0.0864**	-0.106*	-0.164**	-0.123*	-0.132*
			(0.054)	(0.073)	(0.042)	(0.037)	(0.048)	(0.0430)	(0.0393)	(0.0541)	(0.0614)	(0.0673)	(0.0712)
		Observations	2,835	2,848	2,858	2,844	2,561	2,817	2,870	2,857	2,865	2,865	2,871
		R-squared	0.571	0.551	0.587	0.624	0.626	0.844	0.802	0.634	0.490	0.432	0.303
	1900	*Fractionalisation*		0.474***	-0.036	-0.0155	-0.068	0.185**	0.273***	0.289***	0.244**	0.254**	0.274***
				(0.081)	(0.069)	(0.071)	(0.076)	(0.0802)	(0.0864)	(0.0969)	(0.102)	(0.116)	(0.101)
		Polarisation		-0.126*	-0.015	-0.058	-0.048	0.0161	-0.0243	-0.0855	-0.102	-0.0704	-0.0704
				(0.069)	(0.052)	(0.047)	(0.057)	(0.0537)	(0.0533)	(0.0667)	(0.0847)	(0.0933)	(0.0879)
		Observations		3,046	3,094	3,070	2,750	3,046	3,103	3,085	3,098	3,098	3,103
		R-squared		0.542	0.589	0.628	0.627	0.835	0.794	0.606	0.476	0.405	0.300
	1910	*Fractionalisation*			-0.049	0.0053	-0.016	0.160**	0.159***	0.157**	0.198***	0.215***	0.283***
					(0.058)	(0.055)	(0.070)	(0.0630)	(0.0580)	(0.0724)	(0.0641)	(0.0721)	(0.0602)
		Polarisation			-0.019	-0.054	-0.095*	0.0162	0.0133	-0.000155	-0.0365	-0.0354	-0.0569
					(0.048)	(0.044)	(0.052)	(0.0633)	(0.0574)	(0.0739)	(0.0617)	(0.0671)	(0.0555)
		Observations			3,117	3,089	2,757	3,071	3,128	3,111	3,123	3,123	3,128
		R-squared			0.591	0.628	0.628	0.833	0.788	0.600	0.474	0.400	0. 305
Concentration	1880	*Concentration*	-0.250***	-0.209**	-0.054	0.022	0.025	-0.158***	-0.204***	-0.199***	-0.219***	-0.205***	-0.218***
			(0.069)	(0.102)	(0.068)	(0.059)	(0.062)	(0.0578)	(0.0476)	(0.0471)	(0.0605)	(0.0601)	(0.0647)
		Observations	2,836	2,849	2,859	2,845	2,561	2,818	2,872	2,859	2,867	2,867	2,873
		R-squared	0.565	0.551	0.587	0.624	0.625	0.844	0.800	0.631	0.482	0.423	0.292
	1900	*Concentration*		-0.471***	0.075	0.065	0.135	-0.236***	-0.319***	-0.306***	-0.183**	-0.200**	-0.195**
				(0.067)	(0.075)	(0.079)	(0.081)	(0.0829)	(0.0730)	(0.0872)	(0.0786)	(0.0831)	(0.0779)
		Observations		3,046	3,094	3,070	2,750	3,046	3,104	3,086	3,099	3,099	3,104
		R-squared		0.541	0.590	0.628	0.627	0.835	0.793	0.606	0.471	0.398	0.291
	1910	*Concentration*			0.086	0.035	0.098	-0.209***	-0.203***	-0.184***	-0.180***	-0.196***	-0.240***
					(0.055)	(0.054)	(0.069)	(0.0500)	(0.0420)	(0.0486)	(0.0564)	(0.0513)	(0.0545)
		Observations			3,117	3,090	2,757	3,072	3,129	3,112	3,124	3,124	3,129
		R-squared			0.591	0.628	0.626	0.833	0.788	0.599	0.471	0.396	0.296
		Lag. contr.	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
		Base year contr.	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*no*
		State controls	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*

Robust standard errors in parentheses, clustered at state level *** p<0.01, ** p<0.05, * p<0.1

[22] Detailed estimation results including control variable coefficients can be made available upon request

TABLE 5-6 THE DYNAMIC EFFECT OF DIVERSITY AND CONCENTRATION, IV[23]

Var. of interest	Base year	Dep Var.: Income	(1) 1900	(2) 1910	(3) 1920	(4) 1930	(5) 1940	(6) 1950	(7) 1960	(8) 1970	(9) 1980	(10) 1990	(11) 2000
Diversity	1880	*Fractionalisation*	1.961***	0.899***	-0.081	-0.050	-0.118	0.227***	0.425***	0.450***	0.396***	0.414***	0.553***
			(0.181)	(0114)	(0.1005)	(0.095)	(0.110)	(0.0708)	(0.0663)	(0.0727)	(0.0798)	(0.0848)	(0.0914)
		Polarisation	-0.893***	-0.407***	-0.023	-0.073	-0.053	0.0593	-0.0849	-0.188***	-0.221***	-0.214**	-0.296***
			(0.172)	(0.110)	(0.092)	(0.086)	(0.104)	(0.0679)	(0.0665)	(0.0729)	(0.0805)	(0.0835)	(0.0858)
		Observations	2,831	2,844	2,849	2,832	2,553	2,806	2,857	2,845	2,852	2,852	2,858
		First-stage F-stat	73.55	78.41	70.38	87.16	61.99	72.93	82.92	82.88	76.85	77.17	74.78
	1900	*Fractionalisation*		0.661***	0.129	0.0332	0.054	0.327***	0.417***	0.451***	0.553***	0.463***	0.539***
				(0.119)	(0.107)	(0. 099)	(0.101)	(0.0792)	(0.0683)	(0.0745)	(0.0829)	(0.0825)	(0.0916)
		Polarisation		-0.332***	-0.062	-0.065	-0.104	-0.105	-0.106*	-0.191***	-0.297***	-0.198***	-0.236***
				(0.095)	(0.087)	(0.0802)	(0.085)	(0.0674)	(0.0579)	(0.0636)	(0.0708)	(0.0726)	(0.0791)
		Observations		2,848	2,853	2,836	2,557	2,810	2,861	2,849	2,856	2,856	2,862
		First-stage F-stat		135.81	136.13	145.29	123.39	153.70	151.99	156.91	142.44	144.29	137.57
	1910	*Fractionalisation*			-0.060	-0.018	-0.075	0.220***	0.378***	0.409***	0.355***	0.349***	0.358***
					(0.072)	(0.067)	(0. 072)	(0.0499)	(0.0511)	(0.0570)	(0.0594)	(0.0611)	(0.0607)
		Polarisation			-0.019	-0.064	-0.058	0.0194	-0.0675	-0.151***	-0.185***	-0.150***	-0.140**
					(0. 061)	(0.057)	(0.060)	(0.0454)	(0.0446)	(0.0502)	(0.0553)	(0.0559)	(0.0555)
		Observations			3,090	3,064	2,744	3,041	3,096	3,079	3,091	3,091	3,096
		First-stage F-stat			246.59	261.98	255.31	282.16	279.32	279.97	286.17	293.68	309.42
Concentration	1880	*Concentration*	-1.895***	-0.868***	0.223	0.188	0.259*	-0.330**	-0.456***	-0.425***	-0.245*	-0.249*	-0.318**
			(0.272)	(0.154)	(0.148)	(0.150)	(0.147)	(0.145)	(0.120)	(0.129)	(0.141)	(0.136)	(0.154)
		Observations	2,836	2,849	2,854	2,837	2,557	2,811	2,862	2,850	2,857	2,857	2,863
		First-stage F-stat	85.22	80.40	81.84	94.15	95.99	93.91	101.21	111.46	106.47	104.99	98.26
	1900	*Concentration*		-0.618***	-0.066	0.041	0.042	-0.356***	-0.429***	-0.408***	-0.407***	-0.355***	-0.398***
				(0.194)	(0.144)	(0.124)	(0.125)	(0.131)	(0.0974)	(0.106)	(0.125)	(0.109)	(0.133)
		Observations		2,849	2,854	2,837	2,557	2,811	2,862	2,850	2,857	2,857	2,863
		First-stage F-stat		63.38	65.03	72.04	70.39	76.30	73.25	79.03	70.62	69.16	65.78
	1910	*Concentration*			0.157	0.123	0.190*	-0.297***	-0.410***	-0.380***	-0.227**	-0.235**	-0.218**
					(0.098)	(0.100)	(0.1003)	(0.114)	(0.0992)	(0.110)	(0.105)	(0.109)	(0.102)
		Observations			3,094	3,069	2,748	3,046	3,101	3,084	3,096	3,096	3,101
		First-stage F-stat			238.21	307.38	235.88	298.02	254.42	281.25	300.42	293.77	293.81
		Lag. contr.	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
		Base year contr.	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*no*	*no*
		State controls	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*

Robust standard errors in parentheses, clustered at state level *** p<0.01, ** p<0.05, * p<0.1

[23] Detailed estimation results including control variable coefficients can be made available upon request

Replacing fractionalisation and polarisation with concentration yields results which are almost the reverse carbon copy of the fractionalisation coefficients. Concentration proves yet again to have a negative, enduring, and strongly robust impact on local economic development. Just like the indexes of fractionalisation and polarisation, high levels of concentration within a county's population reveal a dynamic impact on income levels over time. The large coefficients in the short-term decrease over time, despite keeping strong significance levels throughout.

The above findings reinforce and extend the findings of Hong and Page (2001), Florida (2002), and Niebuhr (2010) that more diverse places, measured by the number of population groups, are more economically dynamic and productive than more homogeneous places. Diversity fosters economic growth – not only in the short-, but also in the medium- and even in the very long-term. However, there seems to be a strong need for channels of dialogue between the different groups, as the relative size and distance between groups of different origins interacting in a territory proves to be detrimental for sustainable economic development. If the lines separating groups are too deep and insuperable, communication lines fail, bridging between groups becomes difficult, resulting in social unrest and conflict, highly polarised societies, and thus low economic growth for decades to come.

5.6 CONCLUSION

The question of whether and how population diversity impacts the economic trajectory of territories has recently attracted increasing attention (i.e. Easterly and Levine, 1997; Alesina and La Ferrara, 2005; Ottaviano and Peri, 2006; Gören, 2014; Alesina et al., 2016; Bove and Elia, 2017). The literature dealing with the topic has focused on a

multitude of factors, ranging from the labour force setup and skill endowments to the provision of public goods. Two opposing views have emerged – one referring to diversity as growth enhancing, the other as growth reducing. Each view is to some extent dependent on the respective diversity indicator employed. The most frequently used are population fractionalisation and measures of polarisation and segregation.

Most analyses have, however, typically considered the short-term economic impact of diversity rather than evaluating its effects over longer timeframes. Despite an undeniable effect on growth over the short-term, whether past population diversity levels still affect economic outcomes over the medium- or long-term and whether there is a time varying impact on regional economic prosperity remains an almost untouched area within the scientific literature. This chapter has aimed to fill in the gap. The objective has been to assess the extent to which diversity, measured as two dimensional notion of fractionalisation and polarisation, in the population composition of US counties during the Age of Mass Migration between 1880 and 1910 has left an imprint on the region's economic development and whether that potential imprint can still be felt today, more than 130 years later. It also evaluates if the dimension and the direction of the impact of diversity on economic development trends over time by considering the impact of diversity over a time span of 130 years, shifting the focus of the analysis 10 years at a time between 1900 and 2000.

The results of the analysis identify the presence of a strong and very long-lasting impact of diversity on county-level economic development. Counties that attracted migrants from very diverse national and international origins over a century ago are significantly richer today than those that were marked by a more homogeneous population at the time. Highly diverse counties after the big migration waves of the late 19th and early 20th centuries strongly benefited from the enlarged skillset, the

different perspectives and experiences, the dynamic and risk-seeking character of the arriving migrants and from the interaction among those different groups. The result seems to have been a surge of new ideas and of a newfound dynamism that was quickly translated into lofty short-term economic gains. These gains proved durable and, albeit in a reduced way, can still be felt today.

Yet the benefits of diversity come with a strong caveat: our retrieved results indicate the gains of having a large number of groups from different origins within a territory (fractionalisation) only to materialise if the diverse groups are able to communicate with one another (low polarisation). Deep cut lines separating the groups (high polarisation) emerge as an important barrier for economic development. Hence, diversity becomes a double-edged sword: One might argue that it works only if the different groups can interact, that is, if the 'melting pot' really happens. Where it is not possible to build a dialogue between the different groups, where bridging does not occur, groups and communities remain in their own physical or mental ghettoes, undermining any economic benefits from a diverse environment.

In the US context, the benefits from diversity have remained over time. Where high levels of diversity are coupled with 'bridging' across groups – high population fractionalisation with low polarisation – economic gains were generated that were felt in the short-, medium-, and long-term. With the exception of the highly turbulent 1920s to 1940s, a strongly positive and robust association between fractionalisation and regional income levels, as well as a negative association of polarisation, is evident in the analysis. The only change in this enduring relationship is that both connections, while remaining strongly statistically significant, become weaker after the 1920s. While the initial spark of diversity at the turn of the 20th century is revealed a strong

booster of economic dynamism for a period of between 10 to 30 years, its impact, albeit decreasing, has not yet faded entirely.

We can only speculate cautiously about the reasons as to why this is the case. Among the many different possible views on the mechanisms at work, we assume that as successive generations of migrants have blended into the American 'melting pot' and often moved away from where their ancestors settled, the seeds of diversity coupled with the risk-seeking and entrepreneurial character of the migrant may have grown roots not only in local institutions, but also in places. Diversity in those places where it facilitated the bridging among groups more than a century ago has, in all likelihood, generated more welcoming, vibrant, entrepreneurial, and economically dynamic territories. This vibrancy has, in a way, together with the mentality of the migrant become embedded in the very core of the territory, a factor which guarantees that transformations which took place a very long time ago are still felt today. However, further case study based anthropological research will be needed in order to firmly prove this point.

The results of the analysis also have implications for policy. Even though the conditions and circumstances today do not correspond to those in the US in the late 19th and early 20th centuries, our results appeal for pause and thought in a period when migration policies are fast changing and have often become driven by extreme parties and the tabloid press. At a time when many developed countries are rapidly closing down their borders to immigration (Beine et al., 2016), trying to shield what – particularly in the case of Europe and Japan – are still rather homogeneous populations from external influences and the perceived security, economic, and welfare threats often unjustly associated with migrants, restricting migration will limit diversity and is bound to have important and long-lasting economic consequences. By foregoing

new migration, wealthy societies may be jeopardising, as our research shows, not only the short-term positive impact associated with greater diversity, but also the enduring positive influence of diversity on economic development. The large, positive, and persistent impact of societal diversity on economic development seen in the US would therefore be difficult to replicate – something that ageing and lethargic societies across the world cannot relinquish. However, if migration is to be encouraged, it is of utmost importance that mechanisms facilitating the dialogue across groups and hence the integration of migrants are in place to guarantee that diversity is transformed into higher and durable economic activity over the short-, medium- and long-term.

5.7 REFERENCES

Ager, P. and M. Brückner (2013a): Cultural diversity and economic growth: Evidence from the US during the age of mass migration. *European Economic Review, 64*, 76-97.

Ager, P. and M. Brückner (2013b): Immigrants' genes: Genetic diversity and economic development in the US, *Munich Personal RePEc Archive*, MPRA Paper No. 51906.

Alesina A., R. Baqir and W. Easterly (1999): Public goods and ethnic divisions. *Quarterly Journal of Economics*, *114*(4), 1243-1284.

Alesina A. and E. La Ferrara (2000): Participation in heterogeneous communities. *Quarterly Journal of Economics*, *115*(3), 847-904.

Alesina A. and E. La Ferrara (2002): Who trusts others? *Journal of Public Economics*, *85*, 207-234.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003): Fractionalization. *Journal of Economic Growth, 8*(2), 155-194.

Alesina, A. and E. La Ferrara (2005): Ethnic diversity and economic performance. *Journal of Economic Literature*, *43*, 762-800.

Alesina, A. and E. Zhuravskaya (2011): Segregation and the quality of government in a cross section of countries. *American Economic Review*, *101*, 1872-1911.

Alesina, A., J. Harnoss and H. Rapoport (2016): Birthplace diversity and economic prosperity. *Journal of Economic Growth*, *21*, 101-138.

Alexander, J. (2007): *Daily life in immigrant America: 1870–1920*. London: Greenwood.

Altonji, J. and D. Card (1991): The effects of immigration on the labor market outcomes of less-skilled natives. In: J. M. Abowd and R. B. Freeman (Eds.), *Immigration, trade and the labor market*, 201-234, Chicago: University of Chicago Press.

Ancona D. and D. Caldwell (1992): Demography and design: predictors of new product team performance. *Organization Science, 3*, 321-341.

Ashraf, Q. and O. Galor (2013): The "Out of Africa" hypothesis, human genetic diversity, and comparative economic development. *American Economic Review, 103*(1), 1-46.

Atack, J., Bateman, F., and Parker, N. (2000): Northern agriculture and the westward movement. In R. Engerman and R. Gallman (Eds.), *The Cambridge Economic History of the United States: The Long Nineteenth Century*, 285–328. Cambridge: Cambridge University Press.

Azzimonti, M. (2011): Barriers to investment in polarized societies. *The American Economic Review, 101*(5), 2182-2204.

Baland, J. and J. Platteau (2003): Institutions and the efficient management of environmental resources. In: D. Malher and J. Vincent (Eds.), *Handbook of Environmental Economics*, 127-190, North Holland, Amsterdam.

Baldwin, K. and J. Huber (2010): Economic versus cultural differences: Forms of ethnic diversity and public goods provision. *The American Political Science Review, 104*(4), 644-662.

Barro, R and X. Sala-i-Martin (1990): Economic growth and convergence across the United States. *NBER Working Paper Series*, No. 3419.

Barro, R and X. Sala-i-Martin (1995): Capital mobility in neoclassical models of growth. *The American Economic Review, 85*(1), 103-115.

Beine, M., Boucher, A., Burgoon, B., Crock, M., Gest, J., Hiscox, M., McGovern, P., Rapaport, H., Schaper, J. and E. Thielemann (2016): Comparing immigration policies: An overview from the IMPALA Database. *International Migration Review*, *50*(4), 827–863.

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

Bijak, J., D. Kupiszewska, M. Kupiszewski, K. Saczuk and A. Kicinger (2007): Population and labour force projections for 27 European countries, 2002–2052: Impact of international migration on population ageing. *European Journal of Population, 23*(1), 1-31.

Bluedorn, J. (2001): Can democracy help? Growth and ethnic divisions. *Economics Letters, 70*(1), 121–126.

Borjas, G. (1994): The economics of immigration. *Journal of Economic Literature, 32*, 1667-1717.

Borjas, G. (1995): The economic benefits from immigration. *Journal of Economic Perspectives, 9*(2), 3-22.

Bove, V. and L. Elia (2017): Migration, diversity and economic growth. *World Development, 89*, 227-239.

Bruk, S. and V. Apenchenko (1964): Atlas Narodov Mira, Glavnoe upravlenie geodezii i kartografii gosudarstvennogo geologicheskogo komiteta SSSR and Institut etnografii im. H. H. Miklukho-Maklaia, Akademiianauk, SSSR, Moscow.

Campos, N. and V. Kuzeyev (2007): On the dynamics of ethnic fractionalization. *American Journal of Political Science, 51*(3), 620-639.

Campos, N., A. Saleh and V. Kuzeyev (2011): Dynamic ethnic fractionalization and economic growth. *The Journal of International Trade and Economic Development, 20*(2), 129-152.

Card, D. (1999): The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of labor economics*, 1801–63. Amsterdam: North-Holland.

Card, D. (2005): Is the new immigration really so bad? Economic Journal, *115*(507), 300-323.

Churchill, A. and R. Smyth (2017): Ethnic diversity and poverty. *World Development*, *95*, 285-302.

Collier, P. (2001): Implications of ethnic diversity. *Economic Policy*, *16*(32), 127-166.

Collier, P. (2013): *Exodus: How migration is changing the world*. Oxford University Press.

Dentlevy, J., and H. Gemery (1977): The role of migrant stock and lagged migration in the settlement patterns of nineteenth century immigrants. *Review of Economics and Statistics, 59*(2), 137–144.

Desmet, K., S. Weber and Ortuño-Ortín, I. (2009): Linguistic diversity and redistribution. *Journal of the European Economic Association*, *7*(6), 1291-1318.

Easterly, W. and R. Levine (1997): Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics*, *112*(4), 1203–1250.

Easterly, W., J. Ritzen and M. Woolcock (2006): Social cohesion, institutions, and growth, *Economics & Politics*, *18*, 103-120.

Elbadawi, I. and N. Sambanis (2002): How much war will we see? Explaining the prevalence of Civil War. *Journal of Conflict Resolution*, *46*(3), 307–334.

Esteban, J. and D. Ray (1994): On the measurement of polarization. *Econometrica, 62*(4), 819-851.

Esteban, J. and D. Ray (1999): Conflict and distribution. *Journal of Economic Theory*, *87*, 379-415.

Esteban, J. and D. Ray (2008): Polarization, fractionalization and conflict. *Journal of Peace Research, 45*(2), 163-182.

Ezcurra, R. and Rodríguez-Pose, A. (2013): Does economic globalization affect regional inequality? A cross-country analysis. *World Development, 52*, 92-103.

Ferrie, J. (2005): History lessons: The end of American exceptionalism? Mobility in the United States since 1850. *Journal of Economic Perspectives, 19*(3), 199-215.

Florida, R. (2002): *The rise of the Creative Class: And how it's transforming work, leisure, community and everyday life*. New York: Perseus Book Group.

Frazier, J. and F. Margai (2010): *Multicultural Geographies: The changing racial/ethnic patterns of the United States.* Albany: State University of New York Press.

Friedberg, R., and J. Hunt (1995): The impact of immigrants on host country wages, employment and growth. *Journal of Economic Perspectives, 9*(2), 23-44.

Glaeser, E., D. Laibson, J. Sheinkman and C. Soutter (2000): Measuring trust. *Quarterly Journal of Economics, 65*, 811-846.

Goldin, C. (1994): The political economy of immigration restriction in the United States, 1890 to 1921. In C. Goldin, G. Libecap (Eds.), *The regulated Economy: An Historical Approach to Political Economy*, 223-258. Chicago: University of Chicago Press.

Gordon, I. R and P. McCann (2005): Innovation, agglomeration, and regional development. *Journal of Economic Geography, 5*, 523-543.

Gören, E. (2014): How ethnic diversity affects economic growth. *World development, 59*, 275-297.

Haines, M. (2000): The Population of the United States, 1790–1920. In S. Engerman and R. Gallman (Eds.), *The Cambridge Economic History of the United States: The long Nineteenth Century*, 143–206. Cambridge: Cambridge University Press.

Hamilton, B., J. Nickerson and H. Owan (2003): Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy, 111*(3), 465-497.

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

Hatton, T. and J. Williamson (1998): *The Age of Mass Migration: Causes and economic impact*. New York, NY: Oxford University Press.

Hong, L. and S. Page (2001): Problem solving by heterogeneous agents. *Journal of Economic Theory*, *97*(1), 123-163.

Hong, L. and S. Page (2004): Groups of diverse problem solvers can outperform groups of high-ability problem solvers. Proceedings of the National Academy of Sciences of the United States of America, *101*(46), 16385-16389.

Horowitz, D. L. (1985): *Ethnic groups in conflict*. Berkley: University of California Press.

Jacobs, J. (1961): *The death and life of the great American cities*. New York: Random House.

Jacobs, J. (1969): *The economy of cities*. New York: Random House.

Jennissen, R. (2003): Economic determinants of net international migration in western Europe. *European Journal of Population, 19*(2), 171-198.

Keefer, P. and S. Knack (2002): Polarization, politics and property rights: Links between inequality and growth. *Public Choice*, *111*, 127–154.

Kemeny, T. (2012): Cultural diversity, institutions, and urban economic performance. *Environment and Planning A*, *44*, 2134-2152.

Kemeny, T. and A. Cooke (2017): Urban immigrant diversity and inclusive institutions. *Economic Geography*, *93*(3), 267-291..

Kerr, S. and W. Kerr (2011): Economic impacts of immigration: A survey. *Finnish Economic Papers*, Finnish Economic Association, *24*(1), 1-32.

Kim, S. (2007): Immigration, industrial revolution and urban growth in the United States, 1820-1920: Factor Endowments, Technology and Geography. NBER Working Paper No. 12900.

La Porta R., F. Lopez de Silanes, A. Shleifer and R. Vishny (1999): The quality of government. *Journal of Law, Economics and Organization*, *15*(1), 222-279.

Lazear, E. (1999): Globalisation and the market for team-mates. *The Economic Journal, 109*(454), 15-40.

Levy, M., and W. Wadycki (1973): The influence of family and friends on geographic labor mobility: An international comparison. *The Review of Economics and Statistics, 55*(2), 198–203.

Lundborg, P. and P. Segerström (2002): The growth and welfare effects of international mass migration. *Journal of International Economics, 56*, 177-204.

Luttmer, E. (2001): Group loyalty and the taste for redistribution. *Journal of Political Economy*, *109*(3), 500-528.

Martinez i Coma, F. and A. Nai (2017): Ethnic diversity decreases turnout. Comparative evidence from over 650 elections around the world. *Electoral Studies, 49*, 75-95.

Mauro, P. (1995): Corruption and growth. *Quarterly Journal of Economics, 110*(3), 681–712.

Mavridis, D. (2015): Ethnic diversity and social capital in Indonesia. *World Development*, *67*, 376-395.

Merk, F. (1978): *History of the Westward Movement*. New York: Alfred A. Knopf.

Montalvo, J. and M. Reynal-Querol (2005a): Ethnic diversity and economic development. *Journal of development economics, 76*, 293-323.

Montalvo, J. and M. Reynal-Querol (2005b): Ethnic polarization, potential conflict, and civil wars. *American Economic Review*, *95*(3), 796-816.

Niebuhr, A. (2010): Migration and innovation: does cultural diversity matter for regional R&D activity? *Papers in Regional Science, 89*, 563–585.

Ottaviano, G. and G. Peri (2006): The economic value of cultural diversity: Evidence from US cities. *Journal of Economic Geography*, *6*(1), 9-44.

Özgen, C., P. Nijkamp and J. Poot (2011a): Immigration and innovation in European regions. *IZA Working Papers*, 5676, Forschungsinstitut zur Zukunft der Arbeit, Bonn.

Özgen, C., P. Nijkamp and J. Poot (2011b): The impact of cultural diversity on innovation: Evidence from Dutch firm-level data. *IZA Working Paper*, 6000, Forschungsinstitut zur Zukunft der Arbeit, Bonn.

Partridge, J. and H. Furtan (2008): Increasing Canada's international competitiveness: Is there a link between skilled immigrants and innovation? Paper presented at the American Agricultural Economics Association annual meeting, Orlando, FL.

Reynal-Querol, M. (2002): Ethnicity, political systems, and civil wars. *Journal of Conflict Resolution*, *46*, 29-54.

Richard O., T. Kochan and A. McMillan-Capehart (2002): The impact of visible diversity on organizational effectiveness: disclosing the contents in Pandora's black box. *Journal of Business and Management, 8,* 265–292.

Riedel, S. (2015): The interrelation of immigrants' interethnic ties and socioeconomic status in Germany. An autoregressive panel analysis. *European Journal of Population*, *31*, 287-307.

Rodríguez-Pose, A. and V. von Berlepsch (2014): When migrants rule: The legacy of mass migration on economic development in the United States. *Annals of the Association of American Geographers, 104*(3), 628-651.

Rodríguez-Pose, A. and V. von Berlepsch (2015): European migration, national origin and long-term economic development in the United States. *Economic Geography, 91*(4), 393-424.

Rodrik D. (1999): Where did all the growth go? External shocks, social conflict, and growth collapses. *Journal of Economic Growth, 4*, 385–412.

Ruggles, S., K. Genadek, R. Goeken, J. Grover and M. Sobek (2015): *Integrated Public Use Microdata Series: Version 6.0*. Minneapolis: University of Minnesota.

Saiz, A. (2007): Immigration and housing rents in American cities. *Journal of Urban Economics, 61*(2), 345–371.

Saxenian, A. (2006): *The new argonauts: regional advantage in a global economy*. Cambridge, MA: Harvard University Press.

Sequeira, S., Nunn, N., and Qian, N. (2017): Migrants and the making of America: The short- and long-run effects of immigration during the Age of Mass Migration. *National Bureau of Economic Research Working Paper Series*, No. 23289.

Staiger, D. and J. Stock (1997): Instrumental variables regression with weak instruments. *Econometrica, 65*(3), 557–586.

Stock, J., and M. Yogo (2005): Testing for weak instruments in linear IV regression. In Donald, A. and J. Stock (Eds.), *Identification and Inference for Econometric Models*. New York, NY: Cambridge University Press.

Trax, M., S. Brunow and J. Suedekum (2015): Cultural diversity and plant-level productivity. R*egional Science and Urban Economics*, *53*, 85-96.

UNDESA (2016): *International migration report 2015: Highlights*. ST/ESA/SER.A/375, United Nations, New York.

Uslaner, E. (2011): Trust, diversity, and segregation in the United States and the United Kingdom. *Comparative Sociology*, *10*(2), 221-247.

Van Knippenberg, D. and M. Schippers (2007): Work group diversity. *Annual Review of Psychology*, *58*, 515-541.

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, 19*, 19–35.

Ward, D. (1990): Population growth, migration and urbanization, 1860-1920, in R. Mitchell and P. Groves (Eds.), *North America: The historical Geography of a changing continent*. 299-319, Savage.

5.8 APPENDIX 5A

TABLE 5-7 VARIABLE DESCRIPTIONS AND SOURCES

Variable	Description	Source
Main variables of interest		
Dependent variable:		
Income (ln)	Natural log of average income aggregated at county level in county *i* in year *t* and t_0 respectively*	
	• 1950-2010: Income per capita in current dollars - not adjusted for inflation	US BEA, US BLS
	• 1940: Family wage and salary income in current dollars – not adjusted for inflation. Calculation based on size of total population	IPUMS USA
	• 1880-1930: aggregated mean income constructed on the basis of median total income score per occupation in hundreds of 1950 dollars. Calculation based on size of total population	IPUMS USA
Independent variables (Model 1):		
Fractionalisation	Level of fractionalisation in population of county *i* in year t_0	IPUMS USA own construction
Polarisation	Level of polarisation in population of county *i* in year t_0	IPUMS USA own construction
Independent variables (Model 2):		
Concentration	Level of concentration in county *i* in year t_0	IPUMS USA own construction
Instruments		
*Shiftshare_diversity**	Fractionalisation index based shiftshare in year t_0	IPUMS USA own construction
*Shiftshare_polarisation**	Polarisation index based shiftshare in year t_0	IPUMS USA own construction

Variable	Description	Source
*Shiftshare_concentration**	Concentration index based shiftshare in year t_0	IPUMS USA own construction
Controls included in *X* and *Z*		
Education	1940-2000: Percentage of population of county *i* with college degree relative to total population above 25 years in year *t-k*	ICPSR
Literacy	1880-1930: Literacy rate in county *i* relative to total population in year t_0	IPUMS USA
Total population (ln)	Natural log of total population in county *i* in year *t-k* and t_0 respectively	ICPSR and US BLS
Share of black population	Percentage of black population of county *i* relative to total population in year *t-k* and t_0 respectively	ICPSR and IPUMS USA
Female participation	Female participation rate in the labour force in county *i* in year *t-k* and t_0 respectively	ICPSR and IPUMS USA
Unemployment	Unemployment rate in county *i* in year *t-k* and t_0 respectively	ICPSR, US BEA and US BLS
Infant mortality	Infant mortality rate in county *i* in *t-k* measured as number of deaths among infants aged <1 year per 1,000 live births	CDC and ICPSR
Agriculture	Percentage of the labour force employed in agriculture in county *i* in year *t-k* and t_0 respectively	ICPSR, US BEA and US BLS
Mean income (ln)	Initial mean income in county *i* in 1880, 1900 and 1910 constructed on the basis of median total income score per occupation in hundreds of 1950 dollars based on individual occupational data, as natural log	ICPSR and IPUMS USA
State controls	State dummies	Own construction

***** Base years: 1870 for 1880 and 1880 for subsequent years | t_0 refers to the years of migration either 1880, 1900 or 1910 | *t-k* refers to the time period 10 years prior to the dependent variable

6 A WOMAN'S TOUCH? FEMALE MIGRATION AND LONG-TERM ECONOMIC DEVELOPMENT

6.1 INTRODUCTION

In 2015, 244 million people globally lived outside their countries of origin. Women represented almost half of this, that is 48 percent or 117 million of total migrants worldwide (United Nations, 2016). The salience of women migrants has led to the development of a large literature focusing on female migrants, how their experience differs from their male coutnerparts, and how they are differentially embedded in family and social structures (e.g. Chant and Radcliffe, 2003; Kofman et al., 2011). Similarly, the economic impact of migration has been thoroughly researched and, more recently, studies have begun to consider its long-term economic impact, finding that it has a significant and positive economic effect on the places of destination and that the positive implications endure more than a century later (Rodríguez-Pose and von Berlepsch, 2014, 2015). Yet, despite the growing literature on gender and incipient research into the long-term economic impact of migration, the interaction between these two factors has largely been ignored. Migration economics has too often considered the migrant as 'genderless' and assumed that the dominant – e.g. 'male' – perspective of the migrant experience was identical for women. Hence, the question of whether there is a gender specific dimension to the long-term macroeconomic impact of migration remains unanswered.

This chapter intends to tackle this gap in our knowledge. Focusing on the United States during the Age of Mass Migration around the turn of the 19th to the 20th century – a

time when millions of people left their homelands and established themselves on American soil – we assess the link between female migration and economic development. We assume that women migrants have a *two-fold* economic impact: A direct effect via territorial concentrations of female migration and a second, indirect one via their children. These two influences may affect economic development differently in the short- and long-term. In line with the relevant literature, we assume migrant women to have played an essential role in migrant communities. While we speculate the migrant woman's role in the short-term to be predominantly connected to an enlargement of the labour force, we assume them to have acted as 'cultural carriers' of the mentality, customs, traditions, and social capital associated to the risk-seeking and economically dynamic character of the migrant over the long-term. In this role, we suspect them to have positively shaped the economic development of the places where they settled. Following our line of thought, one might argue that the institutional constructs brought to the US by migrants in connection to their migrant mentality transferred from mother to child left a territorial imprint on their settlement regions and that this imprint still affects the economic wellbeing of communities and territories in the US today.

In order to test whether this is the case, this chapter adopts the following structure: Section 6.2 provides a historical overview of women migrant settlement patterns in the US at the turn of the 20th century. Section 6.3 summarises the literature linking female migration and economic development. In Section 6.4, we explain the methodology and the data employed in our analysis. Section 6.5 presents the results, while Section 6.6 concludes.

6.2 MIGRANT WOMEN DURING THE AGE OF MASS MIGRATION

The period between the mid-19th century and World War I is commonly referred to as the 'Age of Mass Migration' – a time where population movements reached an unprecedented level in modern history. More than 40 million people left Europe in search of better lives elsewhere. The large majority of these migrants headed towards the Americas and, in particular, to the US (Hatton and Williamson, 1994, 1998; Bertocchi and Strozzi, 2006).

Between 1860 and 1920, the stock of foreign-born population in the US multiplied by more than a factor of three, from roughly 4 million in 1860 to nearly 14 million in 1920 (Gibson and Jung, 2006). During this time, migration to the US was strongly gendered: only one in three migrants was a woman. However, despite a higher male entry rate, greater male return rates and higher male mortality resulted in a relatively equal gender balance of the US foreign-born population with a gender ratio close to 1 over the entire period (Figure 6-1) (Gabaccia, 1994).

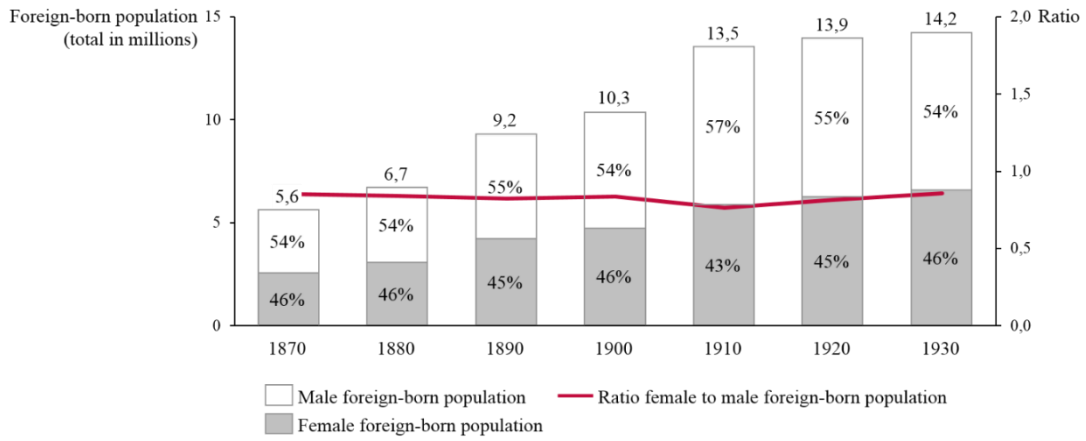


FIGURE 6-1 US FOREIGN-BORN POPULATION BY GENDER, 1870-1930

Data source: Gibson and Jung, 2006; own elaboration

Women often shared the same motives as men for leaving Europe: poverty, population increase, food shortages, drought, political upheaval, and economic and religious oppression. They, however, also escaped other "forms of oppression unique to them as women" (Schwartz Seller, 1981: 6): unwanted arranged marriages, unequal wages and working conditions, discrimination, sexual harassment, and the restrictiveness of female domestic roles. At that time, "daughters without dowries faced declining social status at best, hunger at worst" (Schwartz Seller, 1981: 16).

Entering the US as migrants proved more difficult for women than men. Although US immigration practices were gender neutral in theory, in practice they effectively hindered the unrestricted entry of women into the country. US law made entry difficult for those 'likely to become a public charge'. Particularly when travelling unaccompanied, women were scrutinised far more than men, especially regarding their marital status, intended residence, and financial situation. The absence of a male 'provider' often meant that women were viewed as economic dependents and sent back home (Friedman-Kasaba, 1996).

In spite of this, many foreign-born women made it to the US. The origin of these women was highly diverse in terms of social class, age, national origin, religion, and education. However, a series of common traits can be identified. Most women migrants were young, generally between 25 and 44 (Gibson and Jung, 2006). In 1900, over 70 percent of them stemmed from north western Europe – mainly Germany, Ireland, or the United Kingdom. After 1900, the share of southern and eastern European women – mainly from Russia, Austria-Hungary, and Italy – grew (Table 6-1).

TABLE 6-1 US FOREIGN-BORN POPULATION BY GENDER AND NATIONAL ORIGIN (IN %)

	1900*		1910	
	Men	Women	Men	Women
North-western Europe	**65,6**	**70,5**	**46,5**	**54,3**
Great Britain[1]	11,3	11,3	8,7	9,5
Ireland	13,2	18,5	8,0	12,7
Germany	26,8	27,7	17,4	19,9
Scandinavian countries[2]	10,7	10,0	9,2	9,3
Others[3]	3,6	3,1	3,1	3,0
Southern & Eastern Europe	**19,6**	**15,5**	**40,9**	**32,7**
Italy	5,6	3,6	11,5	7,9
Russia & Finland	6,6	5,3	13,1	12,4
Austria-Hungary	6,6	5,6	13,3	11,1
Others[4]	0,8	1,0	3,0	1,2
Europe "not specified"	**0,2**	**0,2**	**0,0**	**0,0**
Totals				
Europe	85,5	86,2	87,4	87,0
Asia	2,0	0,2	2,1	0,5
Americas	12,2	13,4	10,1	12,2
All other[5]	0,3	0,3	0,3	0,3

* Gender split unavailable for earlier years

[1] Including England, Wales and Scotland

[2] Including Norway, Sweden and Denmark

[3] Including the Netherlands, Belgium, Luxemburg, France and Switzerland

[4] Including the Balkan Peninsula, Portugal and Spain

[5] Including Africa, Australia, Atlantic Islands, born at sea and unspecified country

Source: US Bureau of the Census (1910); own calculations

The large majority of migrant women originated from agrarian and extractive regions, such as the potato and wheat fields of Ireland, the orchards of southern Italy, or the mining towns of the UK. Most were poor, unskilled, and could neither read nor write, having worked either as farmhands or domestic servants in addition to their duties within the home. For those from outside the British Isles, familiarity with the English language was rare (Schwartz Seller, 1981).

In 1900, close to 70 percent of women migrants living in the US were married while only 20 percent were single (US Bureau of the Census, 1900a). The latter was considered a great disadvantage at the time. Being part of a family meant not only support in the new unknown country but, most importantly, financial security. Only very few women had enough savings to pay for the journey and to survive on their own after arrival. Unmarried women travelling alone generally did so to re-join parents, siblings, or other relatives (Weatherford, 1986). Thus, as a way to survive in the unknown, most women migrants became wives or mothers shortly after arrival. The marital age for foreign-born women (around 20-21) was significantly lower than that of American-born women (usually 24-25) (US Bureau of the Census, 1900a). While 70 percent of foreign-born women chose a partner from the same home country, those marrying outside their national origin tended to remain within their cultural group (e.g. an Irish woman marrying a Scot) or chose American-born men (Carpenter, 1927).

Migrant families were usually larger than American ones. Foreign-born women had on average, one child every 3.2 years, while the figure for American women was one child every 5.3 years (US Immigration Commission, 1911). The younger the mother, the higher the likelihood of one pregnancy per year. "[…] The child of a German migrant was three times as likely to be the tenth child in its family as the American baby" (Weatherford, 1986: 2).

Most migrant women at first received support from relatives and kinfolk already settled in the US. "Kinship became the single most important link in the construction of migration 'chains' from specific locations [back in Europe] to specific locations within the United States" (Gabaccia, 1994: 62). After arriving in the US, migrant women followed in their relatives' footsteps and joined them in their area of

settlement. Figure 6-2 shows their settlement pattern based on 1880 Census data. High shares of female migrants were the norm in the north east of the country. In contrast to men, women tended to settle in the highly urbanised arc, extending from Maine to Minnesota and, in particular, in New England, New York, Pennsylvania, Ohio, Michigan, Indiana, Illinois, and Wisconsin. Outside this belt, Utah and parts of southern Texas also had high concentrations of migrant women.

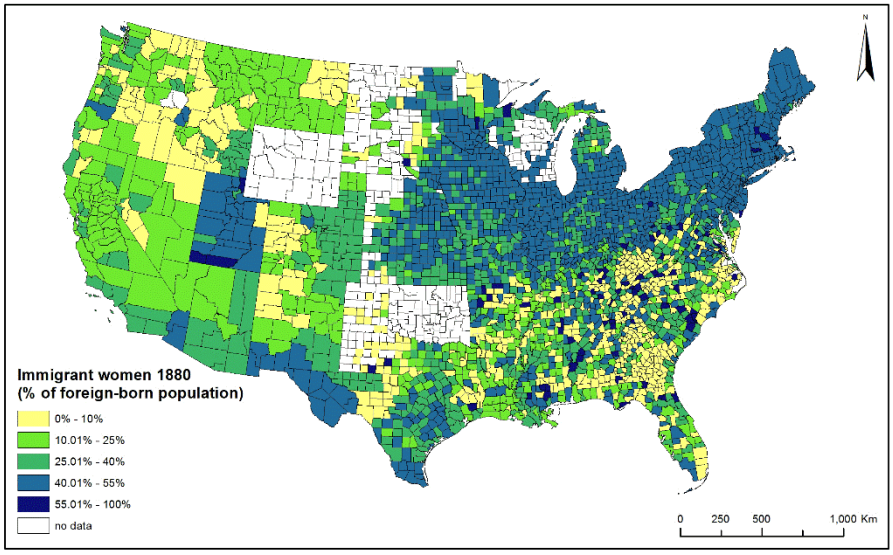


FIGURE 6-2 SETTLEMENT PATTERN OF MIGRANT WOMEN (% OF TOTAL FOREIGN STOCK), 1880

Source: Ruggles et al., 2015; own elaboration

In contrast, migrant men largely outnumbered women in the west and in agricultural regions (Figure 6-3). Urban and industrial areas had a more balanced migrant gender ratio than rural ones – predominantly those in the east (see Appendix 6A). This bias was mostly due to the availability of 'female jobs' in these areas. The combination of domestic services and a concentration of female worker dominated industries, such as

the textile and garment industry in New England and the Mid-Atlantic States, drew migrant women in large numbers. Sixty-two percent of white foreign-born women gainfully employed around the turn of the century were located in these regions of the country (Hill, 1929). These settlement patterns for both female and male migrants remained similar over time, as displayed in Appendix 6B.

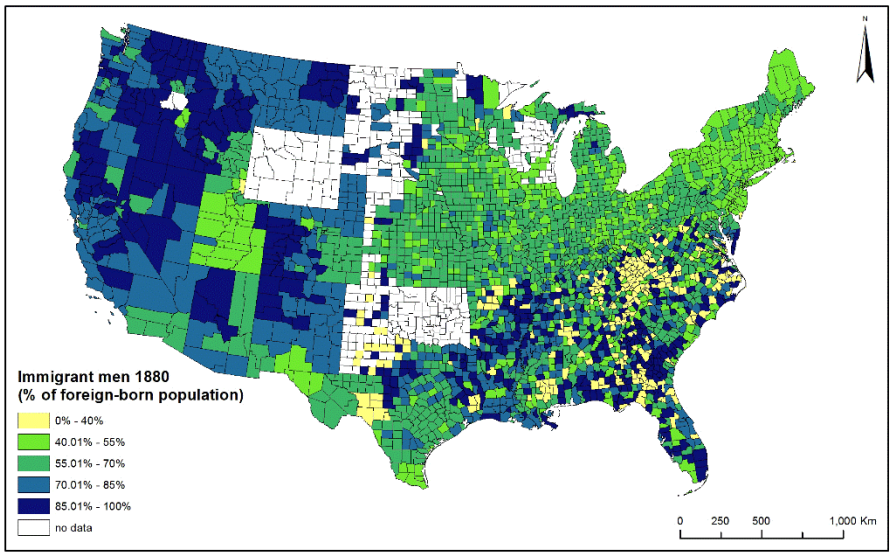


FIGURE 6-3 SETTLEMENT PATTERN OF MIGRANT MEN (% OF TOTAL FOREIGN STOCK), 1880

Data source: Ruggles et al., 2015; own elaboration

Job opportunities affected female foreign-born settlement, yet, only around one in five migrant women were gainfully employed and therefore active in the official labour market. In 1900, a mere 19 percent of the total adult – 15 years and above – female, white, foreign-born population was an 'active breadwinner' (US Bureau of the Census, 1900a, Table 6-2). Migrant female labour force participation also had a very distinctive structure. Predominantly young women – aged under 24 – were in employment. Marriage and the birth of children marked a watershed, as migrant wives

and mothers generally stayed out of the official labour force. Return to official gainful employment was mainly associated with the loss of a husband as a consequence of separation, divorce, or death (Figure 6-4).

TABLE 6-2 FOREIGN-BORN WHITE FEMALES IN GAINFUL EMPLOYMENT BY AGE (IN %), 1890, 1900

	1890	1900
Share of total 15+	**19,8**	**19,4**
15 to 24 years	50,4	48,9
25 to 34 years	19,8	19,8
35 to 44 years	12,0	13,0
45 to 54 years	10,5	11,7
55 to 64 years	9,4	9,8
65+	6,1	6,2
Age unknown	37,5	26,3

Source: US Bureau of the Census, 1900a; own calculations

This implies that the family and the rearing of children, in particular, were at the heart of the life of most migrant, adult women, above 25 years. The tasks all were clearly divided: wives worked inside the home, providing childcare, food, shelter, and clothing while the husbands worked outside, earning the family income. Oftentimes, it was the mother who managed the family finances, making the decisions on how to allocate the family budget. It was commonplace for "the workman to turn over his wages to his wife on pay day and to ask no questions as to what it goes for" (Weatherford, 1986: 104).

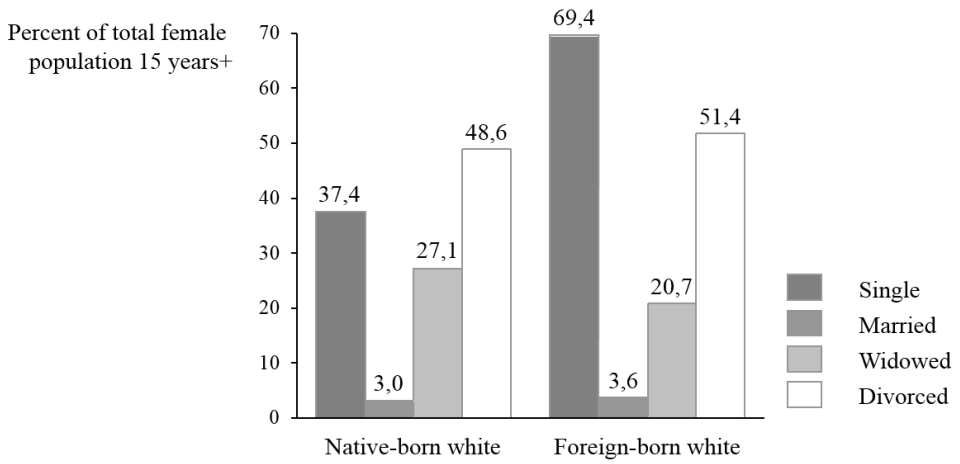


FIGURE 6-4 OFFICIAL EMPLOYMENT FEMALE POPULATION 15 YEARS+ BY MARITAL STATUS, 1900

Source: US Bureau of the Census, 1900a; own elaboration

Officially employed women were concentrated in so called ‘female industries’, such as domestic services, manufacturing, textiles, fabric mills, tobacco, or artificial flower making. While foreign-born men were employed in more skilled, higher paying jobs, the wide range of foreign-born female occupations was at the low end of the occupational ladder and hence low skilled and low paying. Many of the ‘female industries’, such as domestic services and manufacturing, were largely dominated by migrant women (Table 6-3). In 1905, 84 percent of migrant women arriving at the port of New York named domestic occupations, such as servant girls, maids or cooks, as their profession (Kellor, 1907). Employment often helped migrant women to become assimilated into the American society much more than it did for migrant men. Young girls in particular, often hired at very young age of 11 or 12 and who worked as domestic servants, cooks, or maids within American households, integrated quickly. “Domestic service not only augmented family incomes […], but it brought an important segment of the migrant population – the future mothers of the second and

third generations – into intimate contact with middle-class American home life" (Conzen, 1976: 93).

TABLE 6-3 WHITE FEMALE FOREIGN-BORN WORKERS IN MAJOR OCCUPATIONAL CATEGORIES, 1900 AND 1910

Major occupational category	1900 (%)*	1910 (%)*
Agriculture	10,7	7,8
Professional Services	6,3	8,3
Domestic & Personal Service	33,1	33,5
Trade & Transportation (inc. Clerks)	12,6	10,3
Manufacturing & mechanical	22,1	24,3

*Numbers in % of total white females (100%) all nativities gainfully employed in sector

Source: US Bureau of the Census, 1900b; US Bureau of the Census, 1940; own elaboration

Occupational structure, however, varied with age and marital status. While young migrant women largely worked in the domestic service and manufacturing industry, older women, if and when re-entering the labour force, tended to work at night in order to combine employment and child rearing responsibilities. They often undertook unofficial wage earning activities inside their home, commonly including their children as helpers. These homeworkers typically "[packed] food into jars, stripped feathers, basted pants, made buttonholes, crocheted slippers, assembled toys" (Gabaccia, 1994: 50), rolled cigars, or made artificial flowers. Sometimes migrant women within a neighbourhood formed informal cooperatives in order to maximise their work output and support each other (Weatherford, 1986). This contribution to the

family income, however, was often acquired unofficially and thus remained excluded from official records.

6.3 MIGRANT WOMEN AND ECONOMIC DEVELOPMENT

The wide literature on the economics of international migration largely views an influx of migrants as positive for economic development (i.e. Borjas, 1994; Card, 2005). Transmission mechanisms, such as increasing returns to scale (i.e. Borjas, 1995), alterations to the ratio of skilled to unskilled labour (Lundborg and Segerström, 2002), increasing wages (Ottaviano and Peri, 2006), and the stimulation of productivity by means of innovation and specialisation (i.e. Gordon and McCann, 2005; Partridge and Furtan, 2008), are considered important drivers of economic growth in the receiving country. In the case of the US, 19th century migration was behind an increase of 13 to 42 percent in capital stock (Neal and Uselding 1972), making migration essential for the take-off of the US economy (Hirschman and Mogford, 2009).

These findings have, in our opinion, two key drawbacks. First, economic research has tended to downplay the gender dimension of the economic impact of migration, and, second, studies have focused on the short-term, largely ignoring the long-term economic impact.

6.3.1 GENDER AND THE ECONOMIC IMPACT OF MIGRATION

There is growing literature across the social sciences that considers the gender dimension of migration. Prevailing topics have been, for example, income differentials

between husband and wife (Cooke, 2003); differences in remittance patterns and specific household compositions (i.e. Cerrutti and Massey, 2001; de la Brière et al., 2002); working habits and hours, education and brain drain (i.e. Edwards and Ureta, 2003; Özden and Schiff, 2005; Amuedo-Dorantes and Pozo, 2006); or comparisons of women and men migrant social capital and how these affected individual employment opportunities and settlement patterns (i.e. Davis and Winters, 2001; Curran and Rivero-Fuentes, 2003; Linvingston, 2006). This research has shown that the specific characteristics of women, the reasons behind their migration decision, their migratory patterns, the impact on their places of origin, and their assimilation and participation at their places of destination differ from those of men (i.e. Ravenstein, 1889; Hondagneu-Sotelo, 1994; Cerrutti and Massey, 2001; Mahler and Pessar 2003; Oishi, 2005; Andall, 2013).

While the majority of this gender based research has focused on the individual, macroeconomic approaches differentiating along gender lines have been neglected almost entirely. Although female migration has been shown to strongly influence factors behind economic development, such as an increase in the country's labour force (i.e. Cuberes and Teignier, 2012; Lechman and Kaur, 2015), improvements to gender equality (i.e. Berik et al, 2009; Klasen and Lamanna, 2009), and the empowerment of women (i.e. Duflo, 2012), we know little about the specific effect of migrant women on regional growth. Studies analysing the macroeconomic impact of migration have tended to simply pool male and female migrants together to form 'the migrant' – one homogenous group. Women migrants are considered 'tied movers', naturally perceived within the family frame, economic dependents following the male head of the household. In many studies, gender is simply relegated to a control

variable, if considered at all (Pfeiffer et al., 2008). There is an implicit assumption that men and women play equal roles in shaping the economic outcomes of migration.

Of the literature, which does investigate the macroeconomic influence of female migrants, some studies have hinted at a positive economic impact of a strong presence of women migrants in the labour force (Blau et al., 2003, evaluating the US) or of the incidence of foreign women entrepreneurs (Collins and Low, 2010, analysing Australia). Others, however, have pointed in the opposite direction. Smith and Bailey (2006) show a negative economic impact associated to a gender gap between native-born and foreign-born families. Poor integration of migrant women into the labour market and a tendency to end up in occupations below their skill levels also indicate an unused economic potential linked to migrant women (Riaño and Baghdadi, 2007). In short, the limited research on the macroeconomic impact of female migrants goes in different directions, providing no clear answers as to how female migrants shape the economy wherever they settle.

6.3.2 MIGRANT WOMEN AND THEIR LONG-TERM IMPACT FOR ECONOMIC DEVELOPMENT

Most studies on the macroeconomic impact of migration have focused on the relatively short-term. In general, these analyses have found that migration is economically beneficial, improving economic dynamism in the receiving area. The long-term economic effect of migration has, by contrast, been almost entirely neglected. The exceptions are Rodríguez-Pose and von Berlepsch (2014, 2015), who analysed the effect of 19th century migration on US economic development today. Their findings,

robust across different migration waves, underline that the effect of migration is very long-lasting, with positive effects detected at local level long after the first generation of migrants became assimilated into US society: areas of the US where migrants settled in large numbers at the turn of the 20th century are considerably more prosperous today than those largely bypassed by migration. Sequeira et al. (2017) support these results.

One possible line of argumentation among a range of potential explanations for the persistence of a positive economic effect of migration is linked to the institutional 'constructs' built by migrants in their places of destination (Rodríguez-Pose and von Berlepsch, 2014). Higher levels of economic development linked to migration are contingent on the cultural institutions brought and developed by migrants from their places of origin, such as inherited traditions, habits, trust, or customs within the community (Tabellini, 2010; Algan and Cahuc, 2010) coupled with the migrant's character being more risk-seeking, entrepreneurial and economically dynamic than its native-born counterpart. Studies by Acemoglu et al. (2001) have underlined the role cultural institutions play for economic development, as the institutions connected to different European colonial powers have significantly shaped the economic trajectory of countries on the American continent.

Similar mechanisms may have been at play during the Age of Mass Migration. Migrants not only brought their skills, labour, and work ethics, but also their "baggage […] in the way of culture, religion, social networks and links with the society of origin" (Joly, 2000: 30), which helped them model entire institutional structures according to the "national blueprint" (Rodríguez-Pose and von Berlepsch, 2015: 399) they left behind. Not only was language preserved, so too were customs, habits, mentality, and traditions imported (Rice and Feldman; 1997). "[Immigrants] came not

to establish something new but to re-establish something old" (Daniels, 1990: 146). Hence, it is logical to speculate cautiously that the migrants and the institutional frameworks they established have transformed the territories where they settled. Eventually, "as migrants settle, they rework the destination itself" (Silvey and Lawson, 1999: 125).

Institutional constructs have been shown to persist over very long time frames. Putnam (1993), studying social capital, and Duranton et al. (2009), focusing on family structures, have demonstrated that institutional structures built in the Middle Ages still shape regional development today. Algan and Cahuc (2010) also find that values, norms, and beliefs of second and third generation migrants are highly correlated to those of their country of origin. Mass migration has led to the formation of 'ethnic landscapes' (i.e. Conzen 2001; Nostrand and Estaville, 2001), where economic success is a consequence of "the cohesive bond provided by shared values and common backgrounds [of migrants, which] took root and remain strong to this day" (Harwick, 2009: 237). Given the spatial clustering of 19th century migrants arriving in the US, their sheer volume, and the strong migrant networks across the country, it could be assumed that the institutional mechanisms developed during the Age of Mass Migration resulted in the formation of institutional constructs that might still shape current economic development.

But do women play a distinctive role in the transfer of habits and institutional traits from one generation to another? There are a multitude of views on this issue. One argument, for example, builds on the positioning of the woman within the family construct (Zlotnik, 1995). According to Anthias and Yuval-Davis (1989), women are the "ideological reproducers, [...] the 'cultural carriers' of the ethnic group" (p. 9). In their role as mothers, women transmit the cultural heritage, way of life, and history of

their ancestors to the next generation. Within this line of argumentation, tit is mainly the migrant woman who conveys traditions, customs and habits, establishes strong ties within the ethnic community, and plays an essential role in preserving the native culture (i.e. Yuval-Davis, 1993; Pettman, 1996). Migrant women thus bear "the responsibility for the maintenance and generational transmission of culture" (Gray, 2003: 34). Studies across a wide array of disciplines, ranging from economics (i.e. Schüller, 2015) and sociology (i.e. Killian and Hegtvedt, 2003) to psychology (i.e. Sabatier, 2008), have confirmed the role of the mother as cultural carrier and show that the role of the mother in the transmission of ethnic and cultural identity as well as cultural habits to the next generation is significantly stronger than that of the father.

If one evaluates history following this theory, many of the traits of the life of migrant women during the Age of Mass Migration seem to support this theory. While fathers were mostly absent from their children's lives due to long working hours and leisure activities outside the home, mothers were seen as the "guardian of the family" (Friedman-Kasaba, 1996: 130). Mothers taught their children the language of their country of origin and transmitted their way of life, customs, and traditions. Children learned from their mothers about the hardships of an migrant's life, thus contributing to the diffusion of the 'spirit of a migrant' – being more risk-seeking, entrepreneurial and dynamic than their American counterparts – to the second generation. Furthermore, migrant women, far more than their male counterparts or native-born Americans, cultivated strong bonds within the migrant community. Reaching out to neighbours, distant kin, their ethnic group, and their church or synagogue community, foreign-born women created large and dense cultural networks (Gabaccia, 1994). As Ryan (2011) argues, immigrant women often formed localised networks which are different from those of men but no less important in passing on cultural norms. This

was particularly the case with immigrant mothers who often formed networks based on mutual support rather than economic production (Ryan et al., 2008). Within these networks, traditions and cultural festivities were kept alive. Solidarity, support, and cooperation across generations, social class or gender, sharing of information, household equipment, or food within community networks helped in withstanding the hardships of migrant life (Garbaccia, 1994).

In this chapter, we hypothesise that the institutional constructs formed within the period of mass migration have left a territorial imprint that can still be felt in the economic development of US counties today. We speculate that this effect was stronger in areas with higher concentrations of migrant women, possibly acting as 'cultural carriers' of ethnic identity and behaviours to the next generation. According to our theory, large clusters of female migrants are assumed to have fostered the 'spirit of the migrant' to become embedded into the character of the territory, transforming the institutional constructs at the destination areas to become more dynamic, entrepreneurial, and economically active. In line with our hypothesis, these conditions would have been likely to have resulted in an enduring economic dynamism, translated in a greater level of development of those territories that attracted more migrant women and where second generation migrant children born to migrant mothers became prevalent.

6.4 EMPIRICAL APPROACH

In order to assess whether migrant women settling in the US during the Age of Mass Migration left a trace on the subsequent county level economic development, we

estimate two different econometric models. Model 1 focuses on the direct impact of migrant women. We seek to determine whether a large share of female migrants in a given US county affects its economic growth in the short-term and whether this potential impact shifts over time. We also consider the ratio of female to male migrants, inquiring whether a greater migrant gender balance in a county has left a significantly different imprint on the county's economic trajectory than in counties where migrant men predominated.

Model 2 evaluates the indirect effect of migrant women over the long-term, focusing on their children. We analyse the first generation born to migrants on American soil and their imprint on economic growth. We test the notion of the migrant mother as potential 'cultural carrier' and compare the impact on economic development of a large share of children born to migrant mothers, relative to those born to migrant fathers and two American parents.[24]

Our hypothesis is that larger shares of migrant women in a given US county during the Age of Mass Migration will have had a significant and positive impact on the growth trajectory of the county, both directly and indirectly. Following the literature on gender equality, women empowerment, and female participation in the labour force (i.e. Berik et al, 2009; Duflo, 2012), a large share of migrant women could have acted as a driver of regional growth in the short- and long-term. In the short-term, women migrants expanded the labour force, especially in traditional 'female' industries, such as the textile or garment industry, possibly contributing a non-negligible share to the US GDP in the late 19th and early 20th centuries. Their manpower might have boosted

[24] We use the notion of American or 'native-born' to refer to people born on US soil. 'Foreign-born' refers to a birthplace outside of the US.

economic activity shortly after their arrival in the US and served as an important fillip for the receiving region's economy.

In the long-term, the impact of migrant women on economic growth would have adopted a more indirect form via their children. We follow one possible explanation and hypothesise that the role of migrant mothers as carriers of culture and harbingers of the mentality of the migrant (Anthias and Yuval-Davis, 1989) would have made their children and the territories where they settled more dynamic than those where women migrants were absent or in a minority. We speculate the cultural institutions passed on by migrant mothers to their children over 100 years ago to have therefore left an indelible territorial impression on US counties which is still evident today. According to this hypothesis, the migrant women's institutional baggage would have affected the counties' territory in a way that part of the migrants' mentality – being more dynamic, more risk-seeking and entrepreneurial – might have become engraved into the territory's very own character leaving a long-lasting trace for economic development.

6.4.1 MODEL 1: THE DIRECT IMPACT

Model 1 evaluates the *direct* impact of women migrants on their areas of settlement. It adopts the following form:

$$y_{i,t} = \alpha + \beta Femig_{i,t_0} + \gamma Mig_{i,t_0} + \delta Ratio_{i,t_0} + \partial X_{i,t-k} + \theta Z_{i,t_0} + \mu state + \varepsilon_{i,s,t} \qquad (1)$$

where y represents the natural log of income per capita in county i at period t (t = 2010, 1910, 1880); *Femig* depicts the share of female migrants in the total population of

county *i* at time t_0 (t_0 = 1880, 1910); *Mig* is the percentage of foreign-born relative to the total population of county *i* at time t_0; *Ratio* depicts the ratio of female to male migrants in county *i* at time t_0[25]; *X* is a vector of variables associated with the level of economic development of county *i* at time *t-k* (*k* =10);[26] *Z* represents a similar vector of factors considered to have had an effect on the county's economic development at time t_0 and that may also have influenced its attractiveness to migrants; *State* represents state fixed effects controlling for any unobserved factors at state level, while ε depicts the county specific error term, clustered at state level *s*, ensuring robustness to arbitrary spatial correlation within one state. The correlation coefficient between *Femig* and *Mig* and *Mig* and *Ratio* is, respectively, below 0.1 and below 0.4, implying a very limited risk of multicollinearity. All three variables are thus included in the regression analysis.

6.4.2 MODEL 2: THE INDIRECT IMPACT

Model 2 estimates the *indirect* long-term impact of female migration at the turn of the 20th century on regional economic growth in 2010, focusing on children below the age of 16 born to migrant women. While the dependent and control variables remain the same as in Model 1, we exchange the variables of interest for different combinations of parentage. The model takes the following form:

[25] Using a single measure of migration provides analytical clarity, but may fail to reflect the significant diversity between and within different migrant groups. While we cannot fully address this challenge in this chapter; this represents an important strand for further research. We are grateful to a reviewer for raising this point.
[26] In order to avoid issues with multicollinearity, *X* is included only in the long-term analysis.

$$y_{i,t} = \alpha + \beta Mforeign_{i,t_0} + \gamma Fforeign_{i,t_0} + \delta Aparents_{i,t_0} + \partial X_{i,t-k} + \theta Z_{i,t_0} + \mu state + \varepsilon_{i,s,t} \quad (2)$$

where *Mforeign* is defined as the share of children born in a given county *i* to a foreign-born mother and an American-born father relative to the total number of children in the same county at time t_0; *Fforeign* represents the share of American-born children with a foreign-born father and an American-born mother in a given county *i*; and *Aparents* corresponds to the share of children with both American-born parents. The base category in this second model is the share of children with both foreign-born parents. Children are defined as all individuals in a county below the age of 16.

6.4.3 THE DATA

The dependent variable – the natural log of income per capita at county level – was extracted for 2010 from the US Bureau of Economic Analysis (BEA) database, measured in US dollars. As income per capita data more than a century ago are unavailable, we referred to the Integrated Public Use Microdata Series USA database (IPUMS) Version 6.0 (Ruggles et al. 2015) for the construction of the 1880 and 1910 income variables. This database provides US microdata covering the censuses and American Community Surveys between 1790 and 2010.[27] A proxy, aggregated at county level and based on individual data of median total income scores per occupation in 1950 dollars (as natural log), is used. The construction uses the size of the total population including all non-participants within the labour force as base in

[27] The American Community Survey was only initiated in 2005.

order to remain as comparable as possible with the income per capita variable of later years.

Migration and parentage variables in both models were generated using the IPUMS database. A weighted sample of the US population of 5,791,531 individuals in 1880, covering 11 percent of the total US population at the time, and 923,153 individuals in 1910, representing one percent of the population, was employed to construct the main variables of interest. All data were allocated to the individuals' county of residence and aggregated at county level. All residents with a non-US birthplace were classified as 'migrants'.

As US county size, quantity, and geography changed over the period of analysis (2,875 counties or equivalent territorial units in the 48 contiguous states in 1880; 3,123 in 1910; 3,109 in 2010), counties in 1880 and 1910 were matched to their regional equivalent in 2010. Using cartographic boundary files of the 48 continental states (excluding Alaska and Hawaii) for every decade of analysis provided by the US Census Bureau, we normalized county boundaries to the 2010 situation and calculated historical county averages, weighted by the population density during the boundary change. 1880 and 1910 were chosen as reference years for our regression analysis, as both represent a peak in foreign-born population stock – 1880 covering the first migration wave; 1910 covering the second.

Two differentiated sets of control variables are included in the model (vectors *X* and *Z*). Vector *X* comprises factors dating from period *t-k* (year 2000). All variables in vector *Z* date from the time of the two historic censuses, 1880 and 1910 respectively. Vector *X* accounts for variables directly determining the current income per capita levels across US counties. Vector *Z* is incorporated to consider factors that may have

influenced the level of economic activity and therefore growth in a given US county at the time of migration, but also to control for variables that served as pull factors, affecting the initial settlement decision of the migrant women.

Both vectors contain, wherever possible, the same variables measured at county level: educational attainment, unemployment rate, share of black population, urban share, percentage of the labour force employed in manufacturing, and female participation rate in the labour force. While educational attainment is measured as the literacy rate for the historical years, for the *t-k* dimension we resort to the share of people with tertiary education. The share of women in the total county population is included in the model as a way to prevent the main variable of interest, female migration, from picking up effects related to the size of the overall female population in a county.[28] Furthermore, in the long-term analysis, we control for the initial average income at county level at the time of migration by incorporating an income proxy based on individual data reflecting the median total income per occupation in 1950 dollars (IPUMS USA database). All 2000 controls were extracted from the US BEA, the Current Population Survey tables of the US Bureau of Labor Statistics (BLS), and the 2000 Census Summary files. The IPUMS USA database and the Inter-University Consortium for Political and Social Research database (ICPSR) were used as sources for the historical variables. In cases where data were only available at the individual level, the same method as for the migration and parentage variables was used for the construction of new variables. A detailed description of all variables and sources can be found in Appendix 6C.

[28] As the share of female population in t_0 is highly correlated with the stock of migrant women in 1880 and 1910, this variable is discarded from the analysis in the early years. Only the size of the female population in time dimension *t-k* is included.

6.4.4 INSTRUMENTAL VARIABLE (IV) ESTIMATION

Any analysis involving income and migration data over long timeframes is prone to potential endogeneity issues. Prosperous counties may have attracted large numbers of migrant women, but a large share of migrant women, in turn, can be behind the economic dynamism and GDP of these counties. The direction of causality is therefore difficult to ascertain. Furthermore, immigrant women might have purely settled in counties which either had higher income levels or showed good growth prospects, resulting in spatial sorting. Lastly, any model analysing data spanning more than 100 years is highly likely to suffer from omitted variable bias. In order to address these concerns, factor out the true underlying impact of migrant women on economic development, and ensure the validity of our least squares estimates, we refer to instrumental variable estimation methods (IV). Two different types of instruments are proposed for estimating the direct and indirect effect of female migration: Socioeconomic factors and the path dependency of migrant women.

In the case of the *direct* effect of women migrants, we differentiate in Model 1 instruments by time horizon. For the short-term analysis, we employ the share of married individuals and the mean number of distinct generations living in the same household to instrument for the share of migrant women in a given county. Both socioeconomic instruments are taken from the respective year in question (1880 or 1910) and were extracted from the IPUMS USA and ICPSR databases. For the long-term analysis, we add the share of population in urban areas in 1910.

These instruments are uncorrelated with the disturbance term and explain the variation in settlement choice by women migrants. Spinsterhood was considered a great disadvantage at the time, meaning that migrant women to the US either were already

married upon arrival or married shortly thereafter. The outlook "to a land where they could marry quickly and relatively well, and where they could exercise more choice in acquiring a spouse" (Gabaccia, 1994: 34) encouraged many women who had few hopes of a favourable marriage in their places of origin to cross the Atlantic. Marriage rates among the migrant population were distinctly higher than among the American-born population, just as foreign-born women married at a significantly younger age – women migrants married, on average, five years earlier than American-born women (Dickinson, 1980). This implies that marriage rates are likely to be correlated with the presence of women migrants, without simultaneously being correlated with the error term.

On average, migrant women were also younger when they had their first child than their American counterparts. "An annual pregnancy was a fact of life for a great many immigrant women" (Weatherford, 1986: 2). Children were considered an economic asset. They worked for the family from an early age and helped make ends meet. To save money, the large majority of children remained part of their parents' household long after they had grown up. Migrant women also generally moved with their in-laws upon marriage. As migrant women were frequently responsible for establishing and maintaining migrant social networks, their 'kinship-work' often involved taking care of the older generation in their own homes (Weatherford, 1986; Gabaccia, 1994). Consequently, migrant women tended to end up living with their children, their children's children, and their parents, making the average number of generations under one roof a suitable instrument to assess endogeneity. As family structure and living arrangements in the US drastically evolved over the past 130 years, the number of distinct generations in one household more than a century ago does not have any independent influence on income per capita levels in 2010.

By adding the urbanisation level 1910 as instrument for the share of migrant women in the long-term analysis, we exploit the distinct settlement pattern of migrant women depicted in Figure 6-1. Women migrants clustered predominantly in the highly urbanised areas of the northeast, where 'female industries' concentrated. As many current major urban clusters were merely small towns 100 years ago, the share of urban population at the turn of the 20th century is uncorrelated to county GDP levels in 2010, meaning that the instrument does not violate the exogeneity condition.

For the *indirect* effect of migrant women on economic development, a third type of instrument is used in Model 2. Exploiting the fact that migrant women followed the footsteps of their next of kin, creating a path-dependency in migration patterns (Gabaccia, 1994), we use the supply-push component of children with foreign-born mothers and native-born fathers (see also Card and DiNardo, 2000; Ager and Brückner, 2013). This instrument assumes that migrant women in 1880 and 1910 respectively a) follow the same settlement patterns as their average female predecessor and b) integrate in society in a similar manner and frequency (i.e. marrying American men) as migrant women had done in the past.

The supply-push instrument is calculated using the US growth rate of the share of children born to foreign-born mothers and American-born fathers between a base year and the target year 1880 (1910) multiplied by the initial share of this population group in the base year in a given county *i*.[29] As initial base year population shares are used in the computation, the instrument is exogenous to any county specific shocks that

[29] The supply-push component of children with foreign-born mothers for target year 1880 is computed using initial shares in 1870. For 1910, we use 1880 initial base year values due to the significantly larger data sample.

may have affected the share of children born to foreign-born mothers in any given county between 1870 and 1910.

The Staiger and Stock (1997) test for weak instruments and Kleibergen-Paap Wald F statistics, in combination with Stock and Yogo (2005) critical values, are used in order to make sure that all chosen instruments are non-weak instruments. In almost all cases the weak instrument hypothesis is rejected. The only exception is the number of distinct generations per household when used in the long-term analysis of Model 1. We therefore limit the use of this instrument to the short-term analysis.

6.5 ANALYSIS

6.5.1 THE DIRECT IMPACT OF MIGRANT WOMEN

The first part of the analysis focuses on whether large shares of migrant women settling in the US around the turn of the 20th century had a significant direct impact on economic development both in the short- and long-term. We first assess Model 1 for the short-term using ordinary least squares (OLS), followed by an instrumental variable (IV) estimation controlling for wealth influencing factors at the time of migration and including state fixed effects in order to minimise potential issues of spatial correlation and control for unobserved state specific factors. Standard errors are clustered at state level to control for arbitrary spatial correlation within a given state. As instruments, we employ the share of married individuals in a county and the mean number of distinct generations living in a household. Table 6-4 reports the short-term results for our main variable of interest – *female migrants* – with respect to

income per capita in 1880 and 1910 respectively. Columns (1) through (3) display the results with base year 1880, while columns (4) through (6) cover the year 1910.

Contrary to expectations, Table 6-4 reveals a negative association between higher shares of women migrants and economic development across both base years. For the 1880 regressions, negative and strongly significant (at the 1% level) coefficients for OLS regressions are estimated. While high concentrations of migrants, regardless of gender, emerge as an important growth enhancing factor, a higher share of women migrants in a county's population is connected to significantly lower short-term levels of economic development. This result is corroborated and strengthened when correcting for potential endogeneity issues using IV estimation methods. Both instruments retrieve highly significant and negative coefficients for both base years. The ratio of women to men migrants, however, displays a positive sign across OLS and IV 1880 regressions at levels of significance below 5% in the majority of cases. Hence, while migration was a strong driver of economic development, both a large presence of migrant women or a high imbalance between men and women at the turn of the 20th century led to lower levels of economic growth in the short-term. Results for the 1910 regressions, albeit displaying slightly weaker coefficients, point in the same direction.

The validity of these results is reinforced by the significance levels and coefficients of the control variables across both base years and OLS as well as IV regression results. With the exception of the unemployment rate, all controls show the expected signs. A better educated, more urbanised county with a higher share of manufacturing employment, a larger black population, and a strong female labour force participation was significantly richer 100 years ago than a more rural one with lower literacy, less manufacturing, less black population, and fewer women in employment. We assume

the black population variable to serve as proxy for the economic structure of the largely agrarian counties of the South, rather than indicating the effect of a polarised county population into black and white. The coefficient of the share of black population is highly likely to capture the path to convergence of the poor regions of the South to the richer regions in the North (see for example Barro and Sala-i-Martin, 1990, 1995).

One variable which may potentially explain the – at first rather perplexing – result of a seemingly negative impact of women migrants on economic development is female participation in the labour force, which is both positive and statistically significant. During the Age of Mass Migration, only around 19 percent of the total female, white, foreign-born adult population was active in the official labour market (US Bureau of the Census, 1900a). Most foreign-born women were 'homemakers' or worked in the shadow economy. Hence, the yield of the average migrant woman's work was often not recorded in traditional measures of economic activity, such as mean income. Moreover, 'success' for women migrants at the time was often linked to marriage and bearing children. This type of success generally implied leaving the labour force and making a less measurable contribution – at least in official records – to the economy. In any case, the positive and significant coefficient of the female participation in the labour force variable might display that, once gainfully employed, even if in the low skill and low pay 'female jobs' dominant at the time, women migrants made a positive contribution to regional economic growth by enlarging the labour force in the region and boosting economic activity.

TABLE 6-4 THE DIRECT IMPACT OF MIGRANT WOMEN IN THE SHORT-TERM, OLS 1880 AND 1910

Dep. Var: Mean income per capita 1880/1910 (ln)	1880			1910		
	(1) OLS	(2) IV Married	(3) IV Generations	(4) OLS	(5) IV Married	(6) IV Generations
Female Migrants ~	-0.395***	-0.448***	-4.226***	-0.0458	-2.386***	-1.453*
	(0.137)	(0.131)	(0.318)	(0.0434)	(0.135)	(0.782)
Total migrants ~	0.582***	0.592***	1.339***	0.498***	1.608***	1.200***
	(0.141)	(0.126)	(0.426)	(0.113)	(0.263)	(0.326)
Ratio ~	0.0170**	0.0209*	0.297***	-0.00128	0.143***	0.0861*
	(0.00782)	(0.0117)	(0.0365)	(0.00225)	(0.0120)	(0.0502)
Manufacturing ~	0.739***	0.738***	0.675***	0.606***	0.316**	0.422***
	(0.128)	(0.124)	(0.244)	(0.0629)	(0.126)	(0.114)
Urban Share ~	0.157***	0.161***	0.450***	0.350***	0.658***	0.533***
	(0.0469)	(0.0456)	(0.105)	(0.0239)	(0.0761)	(0.0977)
Literacy ~	0.479***	0.488***	1.142***	0.478***	0.765**	0.662**
	(0.0881)	(0.0937)	(0.382)	(0.0810)	(0.300)	(0.274)
Unemployment ~	0.103**	0.101**	-0.0368	0.635***	1.623***	1.212**
	(0.0464)	(0.0448)	(0.0491)	(0.145)	(0.341)	(0.495)
Black Population ~	0.367***	0.372***	0.755***	0.213***	0.738***	0.532***
	(0.0989)	(0.104)	(0.246)	(0.0640)	(0.0892)	(0.206)
Fem. Participation ~	1.147***	1.169***	2.726***	0.988***	1.342***	1.192***
	(0.186)	(0.199)	(0.383)	(0.115)	(0.246)	(0.174)
State Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	2,851	2,851	2,843	3,120	2,950	2,942
R-squared	0.6625	-	-	0.5661	-	-
First-stage F-stat	-	54.19	42.78	-	173.09	14.18

~ Variables date reflect the respective year of migration: 1880 or 1910.
Robust standard errors in parentheses, clustered at state level.
*** p<0.01, ** p<0.05, * p<0.1

When considering the long-term impact of women migrants, wealth influencing controls for both the time of migration as well as 10 years prior to the dependent variable – the natural log of income per capita in 2010 – are included in Model 1. Table 6-5 reports the results for 1880 (results for 1910 are presented in Appendix 6D).

When taking the whole US county sample into account, OLS estimations for both base years display insignificant coefficients for the main variable of interest (Column 1). While overall levels of migration remain positively and significantly associated with higher income per capita levels 100 to 130 years later, large shares of migrant women settling in a given US county around the turn of the 20th century seem to have no

bearing on current levels of county development. This hints at a potential waning of the very strong initial negative impact of female migration numbers over time. Similar findings are obtained for the ratio between female and male migrants. Despite very high significance levels in the short-term displayed in Table 6-4, the wealth declining effect of high imbalances between men and women migrants all but disappears in the long-term.

TABLE 6-5 THE DIRECT IMPACT OF MIGRANT WOMEN IN THE LONG-TERM, 1880

Dep. Var.: Income per capita 2010 (ln)	OLS		IV			
			Whole sample		Female migration ≥ 10%	
	(1) Whole sample	(2) Female migration ≥ 10%	(3) Married	(4) Urban share	(5) Married	(6) Urban share
Female Migrants ~	-0.0336	-0.0652**	-1.063***	-0.562**	-0.838***	-1.008**
	(0.0234)	(0.0297)	(0.284)	(0.236)	(0.189)	(0.491)
Total migrants ~	0.234***	0.284***	0.467***	0.353***	0.395***	0.420***
	(0.0684)	(0.0650)	(0.0739)	(0.0757)	(0.0473)	(0.0933)
Ratio ~	0.00288	0.00386	0.0745***	0.0397**	0.0486***	0.0584*
	(0.00279)	(0.00287)	(0.0228)	(0.0176)	(0.0136)	(0.0311)
Manufacturing 2000	-0.144***	-0.142***	-0.0919*	-0.117***	-0.0989***	-0.0896**
	(0.0253)	(0.0318)	(0.0471)	(0.0300)	(0.0343)	(0.0386)
Education 2000	0.0113***	0.0108***	0.0120***	0.0117***	0.0114***	0.0115***
	(0.000912)	(0.00100)	(0.00111)	(0.000913)	(0.000995)	(0.000962)
Female 2000	0.0153***	0.0153***	0.0155***	0.0154***	0.0152***	0.0152***
	(0.00157)	(0.00181)	(0.00241)	(0.00178)	(0.00238)	(0.00260)
Unemployment 2000	-0.0220***	-0.0249***	-0.0206***	-0.0213***	-0.0230***	-0.0225***
	(0.00436)	(0.00421)	(0.00445)	(0.00401)	(0.00378)	(0.00368)
Black Population 2000	-0.00104***	-0.00100***	-0.00197***	-0.00152***	-0.00101***	-0.00102**
	(0.000344)	(0.000342)	(0.000440)	(0.000332)	(0.000392)	(0.000437)
Fem. Participation 2000	0.00112	0.00121	0.00485**	0.00304**	0.00239*	0.00265*
	(0.00107)	(0.00127)	(0.00208)	(0.00139)	(0.00143)	(0.00145)
Income ~	0.0157	-0.0204	-0.113**	-0.0504	-0.120***	-0.141*
	(0.0240)	(0.0219)	(0.0550)	(0.0418)	(0.0395)	(0.0736)
Manufacturing ~	-0.0118	0.00540	0.117	0.0544	0.0750	0.0902
	(0.0588)	(0.0710)	(0.103)	(0.0829)	(0.0891)	(0.105)
Literacy ~	0.167***	0.247***	0.421***	0.297***	0.291***	0.301***
	(0.0373)	(0.0617)	(0.0862)	(0.0739)	(0.0572)	(0.0696)
Unemployment ~	-0.00510	0.0139	-0.0104	-0.00783	0.00442	0.00235
	(0.0123)	(0.0177)	(0.0190)	(0.0142)	(0.0266)	(0.0280)
Black Population ~	0.233***	0.284***	0.360***	0.298***	0.186***	0.164**
	(0.0358)	(0.0450)	(0.0613)	(0.0482)	(0.0619)	(0.0785)
Fem. Participation ~	-0.0872	-0.129	0.645***	0.289	0.256	0.340
	(0.153)	(0.154)	(0.250)	(0.237)	(0.210)	(0.359)
State Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	2,444	1,998	2,444	2,444	1,998	1,998
R-squared	0.684	0.684	–	–	–	–
First stage F-statistic	–	–	10.84	26.32	39.10	12.40

~ Variables date from respective year of migration, 1880 | Robust standard errors in parentheses, clustered at state level | *** p<0.01, ** p<0.05, * p<0.1

If one assumes, however, that a minimum critical mass of migrant women could be necessary in order to leave a long-lasting territorial imprint on the regions where they settled, we reach different results. Column (2) in Table 6-5 displays the results for those counties where the share of migrant women in the total foreign-born population was at least 10 percent in 1880. In contrast to the whole sample regressions, the coefficient of female migration remains negative but becomes significant at the 5 percent level for 1880 and at the 10 percent level for 1910 respectively (see Appendix 6D). Both the coefficients for total migration as well as for the ratio of migrant women to migrant men remains nearly unchanged. Hence, female migration seems to be significantly and negatively associated with economic development in the long-term only if the number of migrant women settling in any given county surpasses a 10 percent threshold. Below this threshold, it can be assumed that women migrants were simply too few in number to leave a direct long-lasting legacy on economic growth.

As would be expected, most of the year 2000 controls are significant (with the exception of the share of black population in the 1910 base year regressions and female participation in the labour force across both base years) and show the expected signs. While the share of population with a college degree and the share of women in a county's population are positively associated with economic development, a high unemployment rate, a large black population, and high employment in manufacturing are negatively linked to economic growth.

The base year controls, which display significant associations with income per capita 100 to 130 years later are the literacy rate, the percentage of black population, and employment in manufacturing; the latter however only for 1910. The only 1880 or 1910 variable with a strong and significant positive association across all samples and base years over the very long time is the share of total migrants, supporting earlier

work in this area (Rodríguez-Pose and von Berlepsch, 2014). The previously highly significant coefficient of initial female participation in the labour force fails to leave a long-lasting effect and becomes irrelevant for the determination of income per capita levels in 2010. A simple enlargement of the labour force by migrant women adding their manpower to the aggregate production function seems to leave no long-lasting positive impact on regional economic growth.

As a means to address potential endogeneity issues involving models including both migration and income per capita variables, we perform an IV estimation using again the share of married individuals residing in any given county and the share of population living in urban areas. The results for 1880 are displayed in Table 6-5 columns (3) through (6); those for 1910 in Appendix 6D. We use the same two samples as in the OLS analysis, focusing first on the whole county sample and then limiting it to those with a minimum critical mass of 10 percent women migrants in the two historical US Censuses.

The coefficients for women migrants in both base years prove to be robust to correcting for endogeneity and potentially biased estimators and replicate the short-term results. A large presence of foreign-born women at the turn of the 20th century seems to have a significant and negative impact on county income per capita levels in 2010. Consequently, the results suggest that counties where large numbers of foreign-born women established themselves during the Age of Mass Migration have endured a substantially worse economic trajectory over the last 100 to 130 years than those which had been largely bypassed by female migration. In contrast to the OLS regressions, the coefficient of the ratio of women to men migrants is positive and significant for all IV regressions in Table 6-5. As in the short-term, these coefficients allude to the fact that a higher gender balance in the migrant community served as a

powerful driver of economic dynamism over the very long-term. Counties largely dominated by male migrants (equivalent to a very low ratio of female to male migrants) have consequently had a worse economic trajectory than those with a greater gender balance amongst migrants. The coefficients for both sets of control variables in the IV regressions largely coincide with those obtained using OLS.

6.5.2 THE INDIRECT IMPACT OF MIGRANT WOMEN

A seemingly negative direct impact of female immigration on officially recorded levels of economic development does not necessarily mean that women did not have other channels to influence economic outcomes. The second part of the analysis focuses on whether migrant women became drivers of development indirectly via their children in a society that, at the time, shunned their direct contribution to the economy. To do that, we apply Model 2, focusing on the main variable of interest 'children with a foreign mother', including again the two sets of control variables. Table 6-6 reports the results for both base years 1880 and 1910.

The results confirm the hypothesis of the role of the migrant mother as 'cultural carrier' (Anthias and Yuval-Davis, 1989) of the ethnic capital and mentality of the migrant. Across both base years, the OLS coefficients of our main variable of interest is significant at the 1 percent level for 1880 and at the 10 percent level for 1910 respectively (Columns 1 and 2). A large share of children with a foreign mother and an American father is positively associated with higher levels of income per capita in 2010, relative to the base category: children with two foreign parents. Children with a foreign-born father, by contrast, and an American mother seem to contribute no more

to the long-term economic development of a county than the presence of children with two foreign parents. Counties with a higher share of children born to two American parents have, by contrast, performed decisively worse over the long-term. One possible explanation for this result might be that migrant women marrying American men seem to have been the most successful in securing both a better future for themselves – through marriage to Americans who, on average, had higher wages than migrants (Abramitzky et al., 2014) – and for their communities of adoption through their capacity to transmit the spirit of the migrant and their 'cultural baggage' to their children coupled with their ability to become assimilated in the society of their receiving region. One might speculate that the dynamism of the migrant transferred onto the first American-born generation by their migrant mother was thus reinforced by the inter-cultural character of the relation between migrant mothers and American fathers.

TABLE 6-6 THE INDIRECT IMPACT OF MIGRANT WOMEN IN THE LONG TERM, 1880 AND 1910

Dep. Var.: Income per capita 2010 (ln)	OLS		IV	
	(1) 1880	(2) 1910	(3) 1880	(4) 1910
Children with foreign mother ~	0.195***	0.110*	0.219*	0.138**
	(0.0603)	(0.0598)	(0.121)	(0.0542)
Children with American parents ~	-0.142***	-0.0790*	-0.144***	-0.0757*
	(0.0366)	(0.0427)	(0.0225)	(0.0409)
Children with foreign father ~	-0.0568	0.00892	-0.0580	0.0120
	(0.0694)	(0.0448)	(0.0550)	(0.0445)
Manufacturing 2000	-0.152***	-0.147***	-0.152***	-0.147***
	(0.0256)	(0.0286)	(0.0240)	(0.0279)
Education 2000	0.0113***	0.0113***	0.0113***	0.0113***
	(0.000910)	(0.000907)	(0.000665)	(0.000885)
Female 2000	0.0154***	0.0159***	0.0154***	0.0159***
	(0.00161)	(0.00163)	(0.00151)	(0.00158)
Unemployment 2000	-0.0207***	-0.0208***	-0.0207***	-0.0207***
	(0.00432)	(0.00435)	(0.00319)	(0.00423)
Black Population 2000	-0.00118***	0.000137	-0.00118***	0.000136
	(0.000348)	(0.000415)	(0.000330)	(0.000406)
Fem. Participation 2000	0.000882	0.00192*	0.000880	0.00192*
	(0.00107)	(0.00110)	(0.000749)	(0.00108)
Income ~	0.0376	0.0302	0.0377**	0.0303*
	(0.0288)	(0.0187)	(0.0164)	(0.0182)
Manufacturing ~	-0.0306	-0.0962**	-0.0305	-0.0960**
	(0.0605)	(0.0398)	(0.0350)	(0.0388)
Literacy ~	0.166***	0.144	0.166***	0.143
	(0.0511)	(0.0900)	(0.0310)	(0.0881)
Unemployment ~	0.000466	0.0281	0.000484	0.0288
	(0.0146)	(0.134)	(0.0123)	(0.131)
Black Population ~	0.249***	0.144***	0.250***	0.144***
	(0.0365)	(0.0437)	(0.0370)	(0.0426)
Fem. Participation ~	-0.169	0.0392	-0.170**	0.0382
	(0.157)	(0.0646)	(0.0863)	(0.0634)
State Controls	*Yes*	*Yes*	*Yes*	*Yes*
Observations	2,437	2,617	2,437	2,617
R-squared	0.685	0.672	–	–
First stage F-statistic	–	–	188.22	276.60

~ Variables date from respective year of migration 1880 or 1910

Robust standard errors in parentheses, clustered at state level | *** p<0.01, ** p<0.05, * p<0.1

To further test these results while correcting for potential issues of endogeneity and to check whether the effect of the first American-born generation on long-term economic development is causal, rather than a mere association, we conduct an IV estimation

using the calculated supply-push as instrument. The results of the analysis are displayed in Table 6-6, columns (3) and (4).

The outcomes confirm the results of the OLS analysis. The results suggest once again that a large share of children of migrant women marrying American men have a distinctly higher impact on long-term county economic growth than those born to women marrying other migrants or foreign men marrying American women. The presence of higher shares of children with two American-born parents leaves the lowest long-term imprint on economic development of the four different groups considered.

In short, rather than directly, our retrieved findings suggest that migrant women played an important role in the long-term economic development of US counties indirectly, via their children, when successfully integrating in the receiving community. We assume that once actively integrated into the community of their chosen county, 'melting' into US society and marrying native-born Americans, migrant women transmitted the migrant mentality and cultural baggage to their children which set up the foundation for long-term economic dynamism wherever they settled. In line with this way of argumentation, counties endowed with a large share of children born to migrant women and American men more than 100 years ago are significantly better off today than those where the share of children born to women migrants was substantially lower. One way to explain this result is to refer to the ability of migrant women to integrate by establishing inter-cultural ties, paired with their skill to transmit the spirit of the migrant being more dynamic, more entrepreneurial, and risk-seeking onto their children. This ability might have acted as a long-term growth accelerator for those regions which not only succeeded in attracting female migration but also managed to successfully integrate foreign-born women into their communities. While

our research clearly shows the long term impact onto economic development, more qualitative research is needed to accurately prove the assumed transfer mechanism.

6.6 CONCLUSION

Despite a large body of research on migration and economic development, at least two important factors have been neglected. First, the macroeconomic impact of the gender dimension has been overlooked, under the implicit assumption that the economic effects of migration are identical for both men and women. Second, the long-term impact of migration attracted limited attention – perhaps because short-term impacts are politically more urgent. This chapter has aimed to fill both gaps.

We have assumed that women migrants have a *two-fold* impact on economic development: a) a direct one, triggered by large concentrations of migrant women, and b) an indirect one, via their children. With this fundamental assumption in mind, we analysed two research questions: (1) What has been the short- and long-term territorial economic impact of large shares of migrant women settling in the US around the turn of the 20th century? and (2) Do migrant women have a different direct and indirect effect than the migrant in general? The analysis confirms that migrant women have indeed exerted an important and differential direct and indirect effect on the short- and long-term economic development of US counties.

This impact, however, is not always in line with expectations. The concentration of migrant women in specific counties at the turn of the 20th century proved to be connected with significantly *lower* levels of regional economic development both immediately after settling in the US as well as many decades later. Regions which were largely bypassed by female migration have performed significantly better over the following 130 years than those heavily targeted by migrant women. This seems to be mainly due to the low female participation rate in the official labour force. Consequential to large numbers of 'homemakers' and migrant women employed in the shadow economy, the yield of the average migrant woman's work remained unrecorded in traditional measures of economic activity.

The positive influence of women on long-term economic development has come in an indirect manner: via their children. Counties where migrant mothers bore more children – especially those married to American fathers – and which have successfully managed to integrate them into their communities, have been more dynamic over the next century than specifically those where the majority of children were born to two American parents. Communities where large clusters of children were born to foreign-born mothers and American-born fathers more than 100 years ago have become a motor of local economic development. This economic legacy of migrant women is also significantly stronger than that of foreign-born fathers.

Consequently, migrant women have been a distinct force for development in the US since at least the late 19th century. While habits and customs – early, often arranged, marriages which generally led to child bearing and exclusion from the labour market – prevented them from making a measurable,[30] positive contribution, their economic

[30] The average migrant woman's work was not recorded in traditional measures of economic activity.

contribution came in more intricate and indirect ways: via their capacity to shape the gender ratio between foreign-born women and men, via their ability to integrate into their chosen communities and, especially, via their children. We assume women migrants, more than migrant men, to have transferred the risk seeking, hardworking, entrepreneurial spirit of the migrant and the culture of their places of origin to their offspring. Following this line of thought, this transmission might have contributed to trigger an economic dynamism within the children of women migrants that could have become etched in the institutions of the places where they settled. This impact was strongest amongst the more integrated of women migrants: those who married locals. Whereas many foreign women at the time came to America following kin or as a consequence of pre-arranged marriages to men from their villages that had already made the transatlantic crossing, those marrying locals were more independent and determined to make a future for themselves.

Our results, even though embedded into the particular historical and geographical context of the later 19th and early 20th century America, have important policy implications for the present. In times of increasing aversion to mass migration, special attention needs to be paid to policies that lead to inequalities generating male dominated migrant inflows.[31] Huge gender imbalances in the migrant population not only have serious social consequences but, as this chapter shows, important short- and long-term negative economic effects. The results show the need to establish mechanisms to quickly integrate migrant women into society, both by encouraging their labour force participation and by creating the mechanisms to allow them to make free choices in issues such as marriage. Successfully absorbing and integrating women migrants will guarantee current and future economic prosperity. Failing to achieve this

[31] For an analysis of inequalities, migration and government policies, see McGovern (2012)

integration would, in all likelihood, undermine the positive economic influence of migration for decades to come.

6.7 REFERENCES

Abramitzky, R., Boustan, L., and Eriksson, K. (2014): A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy*, 122(3), 467-506.

Acemoglu, D., S. Johnson, and J. Robinson (2001): The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*(5), 1369-1401.

Ager, P, and M. Brückner (2013): Cultural diversity and economic growth: Evidence from the US during the age of mass migration. *European Economic Review*, *64*, 76-97.

Algan, Y. and P. Cahuc (2010): Inherited trust and growth. *American Economic Review*, *100*, 2060-2092.

Amuedo-Dorantes, C. and S. Pozo (2006): Migration, remittances, and male and female employment patterns. *American Economic Review*, *96*(2), 222-226.

Andall, J. (2013): Gendered mobilities and work in Europe: An introduction. *Journal of Ethnic and Migration Studies*, *39*(4), 525-534.

Anthias, F. and Yuval-Davis, N. (1989): Introduction. In Yuval-Davis, N. and Anthias, F. (eds.), *Woman-Nation-State*, 1-15. London: Macmillan.

Barro, R and X. Sala-i-Martin (1990): Economic growth and convergence across the United States. *NBER Working Paper Series*, No. 3419.

Barro, R and X. Sala-i-Martin (1995): Capital mobility in neoclassical models of growth. *The American Economic Review*, *85*(1), 103-115.

Beine, M, Docquier, F. and Rapoport, H (2010): On the robustness of brain gain estimates. *Annals of Economics and Statistics, 97/98,* 143-165.

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.

Berik, G., Y. van der Meulen Rodgers and S. Seguino (2009): Feminist economics of inequality, development and growth. *Feminist Economics, 15*(3), 1-33.

Blau, F., L. Kahn, J. Moriarty and A. Portela Souza (2003): The role of the family in immigrants' labour-market activity: An evaluation of alternative explanations: Comment. *American Economic Review*, *93*(1), 429-447.

Borjas, G. (1994): The economics of immigration. *Journal of Economic Literature, 32*, 1667-1717.

Borjas, G. (1995): The economic benefits from immigration. *Journal of Economic Perspectives*, *9*(2), 3-22.

Card, D. and J. DiNardo (2000): Do immigrant inflows lead to native outflows? *American Economic Review: Paper and Proceedings, 90,* 360-367.

Card, D. (2005): Is the new immigration really so bad? *Economic Journal*, *115*(507), 300-323.

Carpenter, N. (1927): Immigrants and their children 1920. *Census Monographs,* no. 7, Department of Commerce, Bureau of the Census, Government Printing Office, Washington, D.C.

Cerrutti, M. and D. Massey (2001): On the auspices of female migration from Mexico to the United States. *Demography, 38*(2), 187-200.

Chant, S. and Radcliffe, S. (2003): Migration and development: the importance of gender, in Chant, S. (ed.), *Gender & migration in Developing Countries*. 1-29, London: John Wiley & Sons Ltd.

Collins, J. and A. Low (2010): Asian female immigrant entrepreneurs in small and medium-sized businesses in Australia. *Entrepreneurship & Regional Development, 22*(1), 97-111.

Conzen, K. (1976): *Immigrant Milwaukee, 1836-1860: Accommodation and community in a frontier city*. Cambridge, MA: Harvard University Press.

Conzen, M. (2001): American homelands: A dissenting view. In Nostrand, R. and L. Estaville (eds.), *Homelands: A geography of culture and place across America*. 238-271, Baltimore: Johns Hopkins University Press.

Cooke, T. (2003): Family migration and the relative earnings of husbands and wives. *Annals of the Association of American Geography, 93*(2), 338-349.

Cuberes, D. and M. Teignier (2012): Gender gaps in the labour market and aggregate productivity. *Sheffield Economic Research Paper*, SERP 2012017.

Curran, S. and E. Rivero-Fuentes (2003): Engendering migrant networks: The case of Mexican migration. *Demography, 40*(2), 289-307.

Daniels, R. (1990): *Coming to America. A history of immigration and ethnicity in American life*. New York: Harper Perennial.

Davis, B. and P. Winters (2001): Gender, networks, and Mexico-U.S. migration. *Journal of Development Studies, 38*(2), 1-26.

De la Brière, B., A. de Janvry, E. Sadoulet and S. Lambert (2002): The Roles of destination, gender and household composition in explaining remittances: An analysis for the Dominican Sierra. *Journal of Development Economics*, *68*(2), 309-328.

Dickinson, J. (1980). *The role of the immigrant women in the U.S. labor force, 1890-1910*. New York: Arno Press.

Duflo, E. (2012): Women empowerment and economic development. *Journal of Economic Literature*, *50*(4), 1051-1079.

Duranton, G., A. Rodríguez-Pose and R. Sandall (2009): Family Types and the Persistence of Regional Disparities in Europe. *Economic Geography*, *85*(1), 23-47.

Edwards, A. and M. Ureta (2003): International Migration, Remittances, and Schooling: Evidence from El Salvador. *Journal of Development Economics*, *72*(2), 429-461.

Friedman-Kasaba, K. (1996): *Memories of Migration – Gender, Ethnicity, and Work in the Lives of Jewish and Italian Women in New York, 1870-1924*. State University of New York Press, New York.

Gabaccia, D. (1994): *From the Other Side: Women, Gender, and Immigrant Life in the U.S.*, 1820-1990. Indiana University Press, Bloomington and Indianapolis.

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, U.S. Bureau of the Census, Washington, DC.

Gray, B. (2003): Too close for comfort – Remembering the forgotten diaspora of Irish women in England. In Alfonso, C., W. Kokot and K. Tölölyan (eds.), *Diaspora, Identity and Religion: New Directions in Theory and Research*. 33-52, London, New York, NY: Routledge.

Gordon, I. and P. McCann (2005): Innovation, agglomeration, and regional development. *Journal of Economic Geography*, *5*, 523-543.

Harwick, S. (2009): Inscribing ethnicity on the land. In M. Conzen (ed.), *The making of the American landscape.* 229-252, Boston: Unwin Hyman.

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

Hatton, T. and J. Williamson (1998): *The Age of Mass Migration: Causes and Economic Impact.* Oxford University Press, New York, NY.

Hill, J. (1929): Women in Gainful Occupations 1870 to 1920. *Census Monographs*, no. 9, Department of Commerce, Bureau of the Census, Government Printing Office, Washington, D.C.

Hirschman, C. and E. Mogford (2009): Immigration and the American industrial revolution from 1880 to 1920. *Social Science Research*, *38*, 897-920.

Hondagneu-Sotelo, P. (1994): *Gendered Transitions: Mexican Experiences of Immigration*. Berkeley: University of California Press.

Joly, D. (2000): Some structural effects of migration on receiving and sending countries. *International Migration*, *38*(5), 25-40.

Killian, C. and K. A. Hegtvedt (2003): The role of parents in the maintenance of second generation Vietnamese cultural behaviors. *Sociological Spectrum, 23*, 213-245.

Kellor, A. (1907): The immigrant women. *The Atlantic monthly*, 401-407.

Klasen, S. and F. Lamanna (2009): The impact of gender inequality in education and employment on economic growth: New evidence for a panel of countries. *Feminist Economics*, *15*(3), 91-132.

Kofman, E., Kraler, A., Kohli, M. and Schmoll., C (2011): Issues and debates on family-related migration and the migrant family: A European perspective, in Kraler, A., Kofman., E., Kohli, M. and Schmoll, C. (eds.) *Gender, Generations and the Family in International Migration*. 13-54, Amsterdam University Press, Amsterdam.

Lechman, E. and H. Kaur (2015): Economic growth and female labour force participation – verifying the u-feminization hypothesis. New evidence for 162 countries over the period 1990-2012. *Economics and Sociology*, *8*(1), 246-257.

Livingston, G. (2006): Gender, Job Searching, and Employment Outcomes among Mexican Immigrants. *Population Research and Policy Review*, *25*(1), 43-66.

Lundborg, P. and P. Segerström (2002): The growth and welfare effects of international mass migration. *Journal of International Economics*, *56*, 177-204.

Mahler, P. and S. Pessar (2003): Transnational Migration: Bringing Gender in. *International Migration Review, 37*(3), 812-46.

McGovern, P. (2012): Inequalities in the (De-)Commodification of Labour: Immigration, the Nation State, and Labour Market Stratification. *Sociology Compass*, *6*(6), 485-498.

Neal, L. and P. Uselding (1972): Immigration: A neglected source of American economic growth: 1790 to 1912. *Oxford Economic Papers*, *24*, 66-88.

Nostrand, R. and L. Estaville (2001): *Homelands: A geography of culture and place across America*. Baltimore: Johns Hopkins University Press.

Oishi, N. (2005): *Women in Motion*. Stanford: Stanford University Press.

Ottaviano, G. and G. Peri (2006): The economic value of cultural diversity: evidence from US cities. *Journal of Economic Geography*, *6*, 9-44.

Özden, C. and M. Schiff (2005): *International Migration, Remittances, and the Brain Drain*. New York: Palgrave Macmillan.

Partridge, J. and H. Furtan (2008): Increasing Canada's international competitiveness: Is there a link between skilled immigrants and innovation? Paper presented at the American Agricultural Economics Association annual meeting, Orlando, FL.

Pettman, J. (1996): *Worlding Women: A Feminist International Politics*. London, New York, NY: Routledge.

Pfeiffer, L., S. Richter, P. Fletcher and J. Taylor (2008): Gender in Economic Research on International Migration and Its Impacts: A Critical Review. In: A. R. Morrison, M. Schiff and M. Sjöblom (eds.), *The international migration of women*. The International Bank for Reconstruction and Development / The World Bank, Palgrave Macmillan, NY.

Putnam, R. (1993): *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton, NJ.

Ravenstein, E. (1889): The Laws of Migration. *Journal of the Royal Statistical Society, 52*(2), 241-305.

Riaño, Y. and N. Baghdadi (2007): Understanding the labour market participation of skilled immigrant women in Switzerland: The interplay of class, ethnicity, and gender. *International Migration & Integration*, *8*, 163-183.

Rice, T. and J. Feldman (1997): Civic culture and democracy from Europe to America. *Journal of Politics*, *59*, 1143-1172.

Rodríguez-Pose, A. and V. von Berlepsch (2014): The legacy of mass migration on economic development in the United States. *Annals of the Association of American Geographers*, *104*(3), 628-651.

Rodríguez-Pose, A. and V. von Berlepsch (2015): European migration, national origin and long-term economic development in the United States. *Economic Geography*, *91*(4), 393-424.

Ruggles,S., K. Genadek, R. Goeken, J. Grover and M. Sobek (2015): *Integrated Public Use Microdata Series: Version 6.0* [Machine-readable database]. Minneapolis: University of Minnesota.

Ryan, L. (2011): Migrants' social networks and weak ties: accessing resources and constructing relationships post-migration. *The Sociological Review*, *59*(4), 707-724.

Ryan, L., Sales, R., Tilki, M. and Siara, B. (2008): Social networks, social support and social capital: The experiences of recent Polish migrants in London. *Sociology, 42*(4), 672-690.

Sabatier, C. (2008): Ethnic and national identity among second-generation immigrant adolescents in France: The role of social context and family. *Journal of Adolescence, 31*(2), 185-205.

Schüller, S. (2015): Parental ethnic identity and educational attainment of second-generation immigrants. *Journal of Population Economics*, *28*(4), 965-1004.

Schwartz Seller, M. (1981): *Immigrant Women*. Temple University Press, Philadelphia, PA.

Sequeira, S., Nunn, N., and Qian, N. (2017): Migrants and the Making of America: The Short- and Long-Run Effects of Immigration during the Age of Mass Migration. *National Bureau of Economic Research Working Paper Series,* No. 23289.

Silvey, R. and V. Lawson (1999): Placing the migrant. *Annals of the Association of American Geographers*, *89*(1), 121-132.

Smith, D. and A. Bailey (2006): International family migration and differential labour-market participation in Great Britain: Is there a 'gender gap'? *Environment and Planning A, 38*, 1327-1343.

Staiger, D., and J. Stock (1997): Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586.

Stock, J. and M. Yogo (2005): Testing for weak instruments in linear IV regression. In: Donald, A. and Stock, J. (Eds.), *Identification and Inference for Econometric Models*. Cambridge University Press, New York, NY.

Tabellini, G. (2010): Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association*, *8*(4), 677-716.

United Nations (2016): *International Migration Report 2015*. Department of Economic and Social Affairs, United Nations, New York.

US Bureau of the Census (1900a): *US Census Special Report – Statistics on Women at Work*. Department of Commerce, Bureau of the Census, Government Printing Office, Washington, D.C.

US Bureau of the Census (1900b): *US Census Special Report – Occupations*. Department of Commerce, Bureau of the Census, Government Printing Office, Washington, D.C.

US Bureau of the Census (1910): *US Census Report – Vol. 1 Population*. Department of Commerce, Bureau of the Census, Government Printing Office, Washington, D.C.

US Bureau of the Census (1940): *US Census Report –Population. Comparative Occupation Statistics for the United States, 1870 to 1940*. Department of Commerce, Bureau of the Census, Government Printing Office, Washington, D.C.

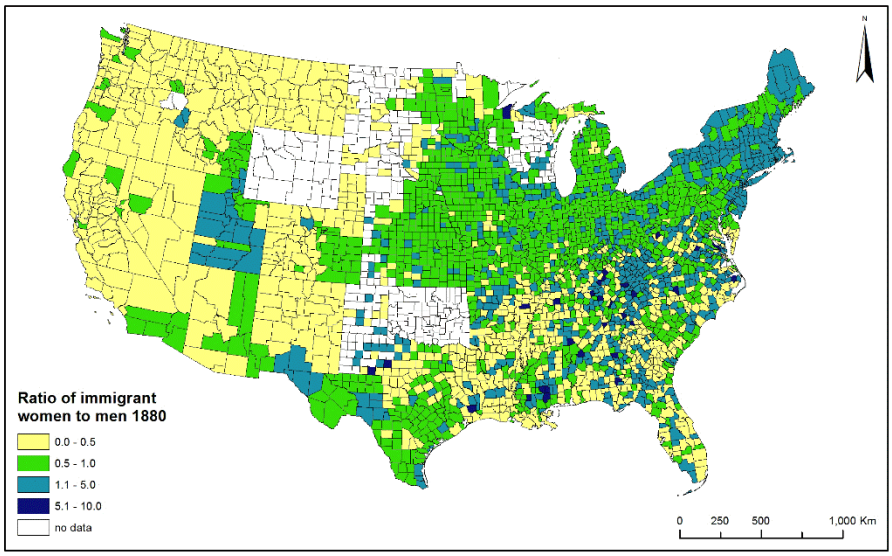
US Immigration Commission (1911): *Abstracts of Reports of the Immigration Commission*, Vol. I and II, 61st Congress, 3rd session, Senate Doc. No. 747. Government Printing Office, Washington, D.C.

Weatherford, D. (1986): *Foreign and Female – Immigrant Women in America*, 1840-1930. Schocken Books, New York.

Yuval-Davis, N. (1993): Gender and Nation. *Ethnic and Racial Studies*, *16*(4), 621-632.

Zlotnik, H. (1995): Migration and the family: The female perspective. *Asia and Pacific Migration Journal*, *4*(2-3), 253-271.

6.8 APPENDIX 6A



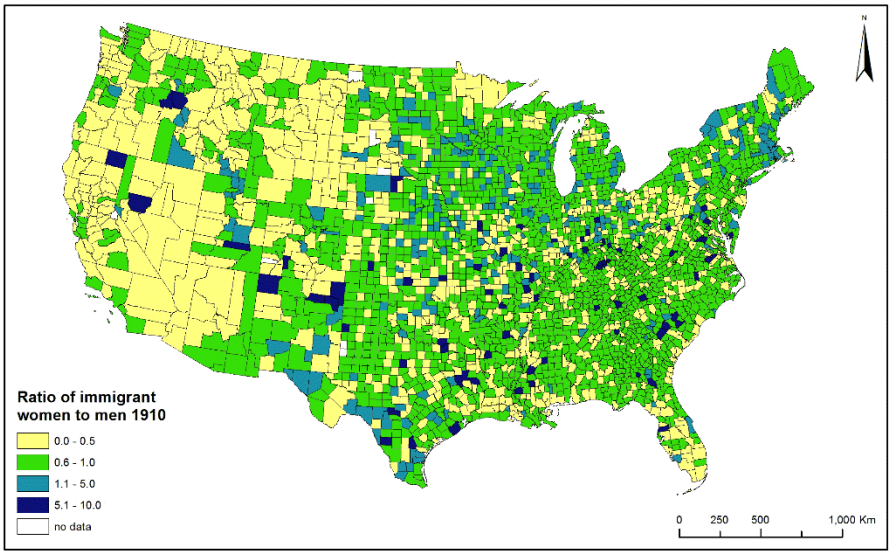
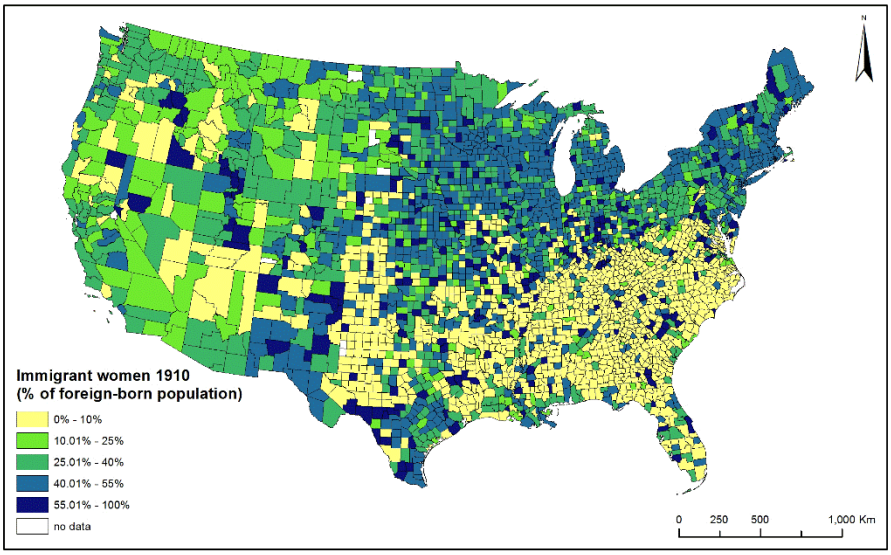


FIGURE 6-5 GENDER RATIO OF FEMALE TO MALE MIGRATION, 1880 AND 1910

Data source: Ruggles et al., 2015; own elaboration

6.9 APPENDIX 6B



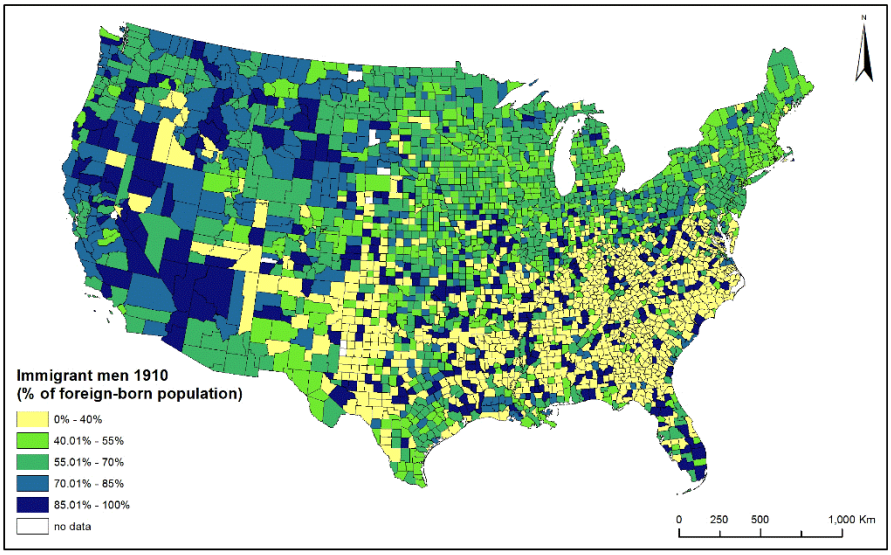


FIGURE 6-6 FEMALE & MALE MIGRANT SETTLEMENT PATTERN (% OF TOTAL FOREIGN-STOCK), 1910

Data source: Ruggles et al., 2015; own elaboration

6.10 APPENDIX 6C

TABLE 6-7 VARIABLE DESCRIPTIONS AND SOURCES

Variable	Description	Source
Main variables of interest		
Dependent variables:		
Income per capita 2010 (ln)	Income levels of county *i* in year *t* and t_0 respectively* • 2010: Income per capita data in current dollars - not adjusted for inflation, as natural log	US BEA
Mean income 1880/1910 (ln)	• 1880, 1910: Constructed mean income score on county level. Individual income levels assigned to occupational data on the basis of median total income per occupation in hundreds of 1950 dollars, as natural log in t_0. Basis for construction is total size of population in 1880/1910	IPUMS USA
Independent variables (Model 1):		
Femig Female migrants	Share of female foreign-born individuals of total population in county *i* in year t_0 constructed from individual data	IPUMS USA
Mig Total migrants	Share of total number of foreign-born individuals of total population in county *i* in year t_0 constructed from individual data	IPUMS USA
Ratio	Ratio of female to male foreign-born population in county *i* in year t_0 constructed from individual data	IPUMS USA
Independent variables (Model 2):		
Mforeign Children with foreign mother	Share of individuals below the age of 16 with foreign-born mother and native-born father in county *i* relative to the total number of children below the age of sixteen in the same county at time t_0	IPUMS USA
Fforeign Children with foreign father	Share of individuals below the age of 16 with foreign-born father and native-born mother in county *i* relative to the total number of children below the age of sixteen in the same county at time t_0	IPUMS USA
Aparents Children with American parents	Share of individuals below the age of 16 with both American-born parents in county *i* relative to the total number of children below the age of sixteen in the same county at time t_0	IPUMS USA

Variable	Description	Source
Instruments		
Married	Percentage of married population in county *i* relative to total county population in year t_0 constructed from individual data	IPUMS USA
Generations	Average number of distinct generations living in one household in county *i* in year t_0 constructed from individual data	IPUMS USA
Urban share	Share of population of county *i* living in urban areas in 1910	ICPSR
Controls included in *X* and *Z*		
Education	Percentage of population of county *i* with college degree in year *t-k**	ICPSR
Literacy	Literacy rate in county *i* in 1880, 1910 constructed from individual data	IPUMS USA
Manufacturing	Percentage of labour force employed in manufacturing in county *i;* for 1880 and 1910 constructed from individual data	US BLS and IPUMS USA
Black Population	Percentage of black population in county *i*	ICPSR and Census 2000 summary files
Female	Percentage of female population in county *i*	Census 2000 summary files
Female Participation	Female participation rate in the labour force in county *i;* for 1880 and 1910 constructed from individual data	ICPSR and IPUMS USA
Unemployment	Unemployment rate in county *i*; for 1880 and 1910 constructed from individual data ('Months unemployed last year' used as proxy in 1880 as unemployment rate not available)	IPUMS and US BLS
State Controls	State dummies	Own construction

6.11 APPENDIX 6D

TABLE 6-8 THE DIRECT IMPACT OF MIGRANT WOMEN IN THE LONG-TERM, 1910

Dep. Var.: Income per capita 2010 (ln)	OLS		IV			
	(1) Whole sample	(2) Female migration ≥ 10%	Whole sample		Female migration ≥ 10%	
			(3) Married	(4) Urban share	(5) Married	(6) Urban share
Female migrants ~	-0.00400	-0.0504*	-0.394	-0.453***	-1.344*	-0.648***
	(0.0201)	(0.0266)	(0.338)	(0.141)	(0.717)	(0.211)
Total migrants ~	0.217**	0.287***	0.426**	0.458***	0.349***	0.316***
	(0.0962)	(0.0817)	(0.209)	(0.102)	(0.0887)	(0.0728)
Ratio ~	-0.000231	0.00135	0.0238	0.0274***	0.0595*	0.0282***
	(0.00165)	(0.00152)	(0.0206)	(0.00952)	(0.0332)	(0.0105)
Manufacturing 2000	-0.147***	-0.108***	-0.139***	-0.138***	-0.0147	-0.0648
	(0.0285)	(0.0377)	(0.0318)	(0.0301)	(0.0890)	(0.0488)
Education 2000	0.0112***	0.0113***	0.0113***	0.0113***	0.0108***	0.0111***
	(0.000905)	(0.00120)	(0.000879)	(0.000893)	(0.00141)	(0.00122)
Female 2000	0.0161***	0.0151***	0.0180***	0.0183***	0.0156***	0.0153***
	(0.00159)	(0.00262)	(0.00239)	(0.00211)	(0.00350)	(0.00280)
Unemployment 2000	-0.0211***	-0.0221***	-0.0230***	-0.0233***	-0.0247***	-0.0233***
	(0.00437)	(0.00597)	(0.00498)	(0.00441)	(0.00714)	(0.00572)
Black Population 2000	0.000202	0.000309	0.000291	0.000304	0.000777	0.000525
	(0.000415)	(0.000725)	(0.000327)	(0.000303)	(0.000760)	(0.000674)
Fem. Part. 2000	0.00204*	0.00100	0.00326**	0.00344***	0.00177	0.00136
	(0.00110)	(0.00147)	(0.00141)	(0.00116)	(0.00164)	(0.00146)
Income ~	0.00151	-0.00246	-0.00128	-0.00169	-0.0243*	-0.0125***
	(0.00229)	(0.00349)	(0.00429)	(0.00324)	(0.0130)	(0.00402)
Manufacturing ~	-0.0931**	-0.0834**	-0.0903***	-0.0899***	-0.0649	-0.0749**
	(0.0374)	(0.0404)	(0.0286)	(0.0285)	(0.0461)	(0.0355)
Literacy ~	0.174*	0.281**	0.248**	0.259***	0.387***	0.330***
	(0.0866)	(0.115)	(0.115)	(0.0982)	(0.102)	(0.0987)
Unemployment ~	0.0364	-0.166	0.311	0.352**	0.130	-0.0294
	(0.131)	(0.128)	(0.246)	(0.176)	(0.276)	(0.145)
Black Population ~	0.147***	0.215	0.220***	0.231***	0.388*	0.295**
	(0.0436)	(0.129)	(0.0781)	(0.0454)	(0.203)	(0.125)
Fem. Participation ~	0.0667	0.0893	0.187	0.205**	0.558*	0.306***
	(0.0617)	(0.0794)	(0.153)	(0.0993)	(0.296)	(0.110)
State Controls	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	2,618	1,692	2,618	2,618	1,692	1,692
R-squared	0.672	0.644	–	–	–	–
First stage F-statistic	–	–	6.72	29.84	5.07	26.26

~ Variables date from respective year of migration, 1910

Robust standard errors in parentheses, clustered at state level | *** p<0.01, ** p<0.05, * p<0.1

7 INTERNAL MIGRATION AND ITS LONG-TERM IMPACT FOR ECONOMIC DEVELOPMENT

7.1 INTRODUCTION

"For we began as explorers, empire builders, pilgrims and refugees, and we have been moving, moving ever since." George Wilson Pierson (2011:91)

In 2015, the United Nations estimated that the number of international migrants was 244 million worldwide (United Nations, 2016). The size of global internal migration, however, was reckoned to be more than three times larger: around 763 million people or 11.7 percent of the world's population are assumed to live within the boundaries of their native country but outside their region of birth (Kuhn, 2015). Yet, despite this imbalance in size, our knowledge about the economic impact of domestic migration is more limited than that of international migrants. Many argue that measurement difficulties, coupled with the relative lack of popularity of internal compared to international movements, have kept the study of domestic migrants pinned to the bottom of the academic agenda.

Past research has focused mainly on the determinants of internal migration, the migrants' characteristics, the individual returns of the migration decision and the effect of remittances, but the macroeconomic impact of internal migrants on their receiving region still remains poorly understood. The few insights on this topic focus on the short- to medium-time frame, barely extending further than 10 to 20 years. To the best of our knowledge, the understanding of the long-lasting economic impact of domestic

migrants is virtually non-existent. Is the settlement pattern of historical internal migrants relevant for current disparities in development? Has domestic migration, taking place more than 100 years ago, left a long-lasting influence on the economic development of the places where migrants settled? And can this impact still be traced today?

The chapter intends to make up for these shortcomings in the literature by focusing on a country often considered as an exceptional case of internal migration – the United States (US). US citizens have traditionally displayed geographical mobility rates nearly double those of other advanced societies; almost one in three Americans lives outside their state of birth and over 30 million Americans change place of residence every year (Molloy et al., 2011; US Census Bureau, 2016). The annual geographical mobility rate in the US is about three times as large that of the EU15, EU27, or Canada and about one and a half times that of Australia (Gill and Raiser, 2012). These high internal migration rates are a legacy of the country's history. Already throughout the 19th century – a time of massive expansion towards the west – almost 60 percent of the US male population above the age of 30 had moved across county or state lines, often covering vast distances in their migration (Ferrie, 2005). By 1880, more than a third of the US population – a number previously unforeseen in US history – consisted of American-born internal migrants (Ruggels et al., 2015; own calculations).

Using a county level dataset for domestic migration in the late 19th and early 20th century covering the 48 US continental states, this chapter first assesses the effect of historical internal migration on long-term economic development. It evaluates whether large shares of domestic migrants have left a long-lasting trace on the territories where they settled. Second, it examines whether the distance covered by American-born migrants more than a century ago matters for current levels of development.

We assume that a vibrant domestic migrant population has the capacity to galvanise growth over the long-term, leaving a long-lasting economic impact that is largely determined by the geographical distance covered by internal migrants before settling down. We assume large numbers of risk-seeking and economically active internal migrants travelling over large distances to be able to transform counties of destination by increasing diversity levels, altering local institutions, and reshaping economic activity. The hypothesis driving this chapter is that the distance covered by internal migrants affects future local economic performance: the bigger the distance travelled by internal migrants more than 100 years ago, the larger the differences between new arrivals and locals, the higher the population diversity levels, the greater the local economic dynamism, and the stronger their long-term economic legacy.

To demonstrate whether this is the case, the chapter adopts the following structure: First, the historical background of internal migration in the US at the turn of the century is described in section 7.2. Section 7.3 summarises the literature on internal migration and economic development, while section 7.4 provides a description of the empirical approach and the data used in the analysis. The discussion of the results can be found in section 7.5. Section 7.6 concludes.

7.2 INTERNAL MIGRATION AT THE TURN OF THE 19TH TO THE 20TH CENTURY

Many of the facets that characterise the American population today are deeply rooted in the country's past. One decisive feature is exceptionally high geographical population mobility. Its origins can be traced back to the 19th century, an era in which the US grew not only in size and population but also economically and, ultimately, in global significance and power. Between 1774 and 1909, a 3.9 percent average annual growth led to a 175-fold increase of real gross national product. Rapid economic expansion was fundamentally the result of increased supply in two production factors: land and labour. Over the span of a few decades, the US territory expanded 3.5 times to encompass around 7.8 million km^2 by 1900. The population grew almost forty-fold over the same timeframe as a consequence of both natural increase and immigration (Gallman, 2000).

During this period, geographic mobility quickly increased to previously unforeseen levels. Ferrie (2005) estimates that nearly two-thirds of American men above the age of 30 migrated across county or state lines during their lifetime – sometimes even more than once. By 1850, the share of American-born population living outside their place of birth was almost 25 percent (Haines, 2000). By 1880, it had reached 33 percent (Ruggles et al., 2015; own calculations). "The American population was a restless one, continually uprooting and moving to a new location […] 'every day was moving day'" (Atack and Passell, 1994: 237).

The chief impetus behind rapid territorial mobility was finely engrained in the expansion of the country. Migration predominantly happened from east to west (Ferrie, 2005). Beginning with the Louisiana purchase in 1803, settlers from the

northeast started crossing the Appalachian Mountains and poured across the Mississippi River by 1850 (Atack et al., 2000). By 1860 they were pushing further westward. The northern plains were transformed into an enormous grain and feed producing area, while the southern plains became major cattle breeding grounds (Merck, 1978). News of mineral resources and animal stock in the far west set in motion thousands, rushing to places such as California, Colorado, Nevada, Arizona, Idaho, and Montana. By 1900, the US settlement stretched from the Atlantic to the Pacific (Gallman, 2000)[32].

Table 7-1 portrays the population redistribution over time. In 1790, no population (see footnote 32) is recorded in the Midwest. By 1830, 12.5 percent of the US population settled in this area and by 1860 one-third of all Americans lived in the region. The west coast population grew at a similar rate, but later. In 1860 only about half a million people lived in the Pacific and mountain states. By 1890 it was nearly 5 percent of the American population and, by 1910, the share had almost reached 10 percent. The geographical centre of the US population continuously shifted further to the west (see e.g. Plane and Rogerson, 2015). Alexis de Tocqueville wrote at the time, "[m]illions of men are marching at once towards the same horizon; […] their manners differ, their object is the same […] to the West they bend their course" (de Tocqueville, 1839: 292).

[32] The history of American geographical mobility ultimately starts with Native Americans. Historians estimate more than around 143 'language families' to have lived on the territory to the north of Mexico prior to the arrival of the Europeans (Spickard, 2007). Thus, the American continent was far from 'empty' and American expansion to the west was by no means a peaceful undertaking. Unfortunately, as data on settlement patterns or economic fundamentals of the native tribes is unavailable, we are unable to include this important part of the American population in the analysis.

TABLE 7-1 POPULATION BY REGION (IN THOUSANDS), 1790-1910

Region	1790		1830		1860		1890		1910	
	N	%	N	%	N	%	N	%	N	%
New England	1.009	25,7	1.955	15,2	3.135	10,0	4.701	7,5	6.553	7,1
Middle Atlantic	959	24,4	3.588	27,9	7.459	23,7	12.706	20,2	19.316	20,9
East North Central	–	–	1.470	11,4	6.927	22,0	13.478	21,4	18.251	19,8
West North Central	–	–	140	1,1	2.170	6,9	8.932	14,2	11.638	12,6
South Atlantic	1.852	47,1	3.646	28,3	5.365	17,1	8.858	14,1	12.195	13,2
East South Central	109	2,8	1.816	14,1	4.021	12,8	6.429	10,2	8.410	9,1
West South Central	–	–	246	1,9	1.748	5,6	4.741	7,5	8.785	9,5
Mountain	–	–	–	–	175	0,6	1.214	1,9	2.634	2,9
Pacific	–	–	–	–	444	1,4	1.920	3,0	4.449	4,8
Total	**3.929**	**100**	**12.861**	**100**	**31.444**	**100**	**62.979**	**100**	**92.228**	**100**

Data source: US Bureau of the Census, 1972

Settlement patterns were far from random. Topography, climate, and natural amenities, such as water supply, climate, soil quality, timber, and animal stock availability determined settler flows (Merck, 1978). Internal migration along the same latitude was the most frequent. Remaining within known climatic bands proved most rational as migrants sought to “maximise the value of their human and physical capital” (Atack et al., 2000: 324). Appendix 7A displays the settlement pattern of domestic migrants, measured as the share of a county’s total American-born resident population in 1880. A clear east-west divide emerges. The lowest numbers of internal migrants are found in the original thirteen states. The share of domestic migrants rises rapidly toward the west. In 1880, states in the Midwest and westerm mountain regions had the largest proportion of internal migrants.

In the late 19th century most domestic migrants travelled short distances: in 1880, 80 percent of all internal migrants settled less than 500km away from their place of birth. For example, in 1850, 67 percent of Pennsylvanian born migrants lived in Ohio, Indiana, or Illinois; 77 percent of South Carolina migrants settled in Georgia,

Alabama, Mississippi, and Tennessee (Haines, 2000). The remaining 20 percent covered much larger distances, in some cases up to 4,500 km (Ruggles et al., 2015; own calculations). Figure 7-1 displays the average distance travelled by domestic migrants in 1880. Not surprisingly, the further west, the larger the distance covered. Distances were shortest in the northeast, exceeded 500 km in the Midwest, and reached 1,500 km in the western part of the Great Plains. The average distance travelled by an internal migrant living on a Pacific coast county ranged between 2,500 and 4,000 km.

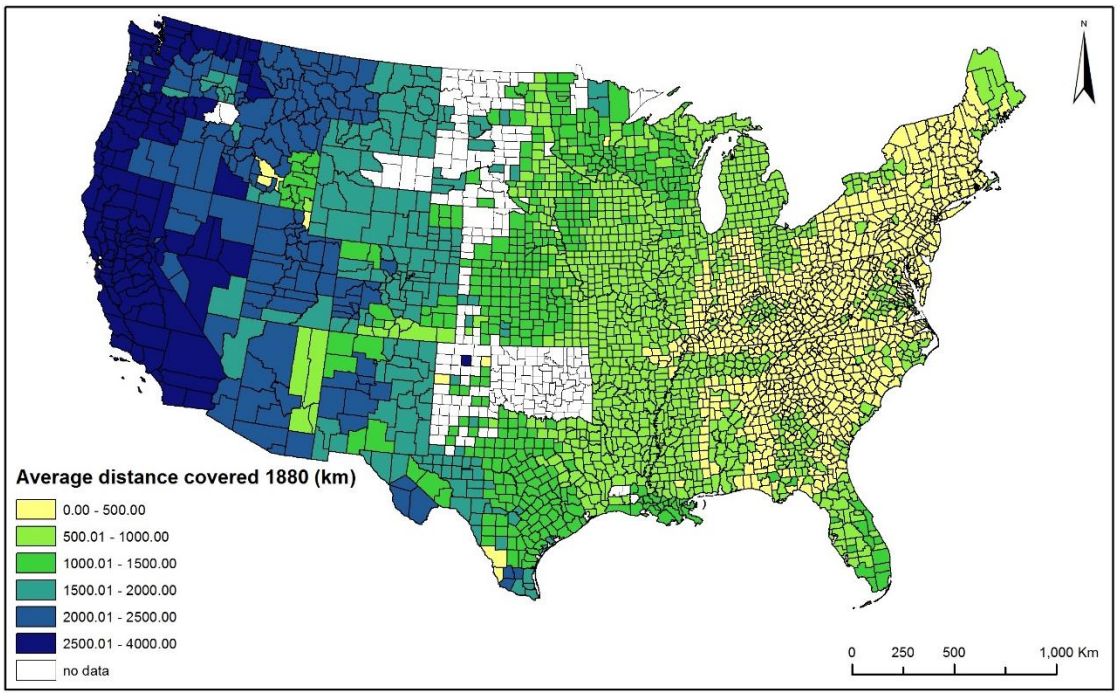


FIGURE 7-1 AVERAGE DISTANCE TRAVELLED BY AMERICAN-BORN MIGRANTS (IN KM), 1880

Data source: Ruggles, et al. 2015; own elaboration

The migrant catchment areas varied widely in size: in 1880 the average internal migrant living in New York City was born around 460 km away; an internal migrant living in Harris County (Houston, Texas) had travelled 1,200 km; while in San

Francisco City s/he had covered more than three times the distance – 3,700 km (Appendix 7B).

The profile of the domestic migrant changed during the period in question. At the outset of the 19th century, the probability to migrate was highest among Southerners. From 1870 onwards, it was those born in the Midwest that moved the most (Rosenbloom and Sundstrom, 2004). Throughout the century North-Easterners displayed the lowest inclination to migrate. Internal migrants came from all walks of life, although literate Americans had a higher propensity to migrate. Women were less likely to move, and whites moved more than blacks (Rosenbloom and Sundstrom, 2004). Internal migration rates were highest among the young, both for short- as well as long-distances (Table 7-2).

TABLE 7-2 19TH CENTURY GEOGRAPHIC MOBILITY OF WHITE, NATIVE-BORN MEN (% OF COHORT)

	Intercounty migrants	**Interstate migrants**
Age 20-29 in initial year		
1850-1860	49,5	26,2
1860-1870	38,2	17,8
1870-1880	54,7	30,1
Age 45-59 in initial year		
1850-1860	21,2	10,6
1860-1870	20,5	8,4
1870-1880	43,5	21,5

Source: Ferrie, 2005

7.3 INTERNAL MIGRATION AND ECONOMIC DEVELOPMENT

Uncovering the link between internal migration and development is difficult. Researchers analysing domestic migration face two challenges: the inconsistent and difficult conceptualisation and subsequent measurement of internal migration on the one hand, and its limited popularity on the other. Problems related to defining, measuring, and collecting data on domestic population flows have hindered research on internal migration. Many different conceptualisations of internal migration, employing a variety of temporal intervals and spatial entities, have been coined (i.e. Petersen, 1986; Rees et al., 2000; Rogers et al., 2003). However, "the size and shape of the spatial areas between which migration is measured are not uniform either within or between countries" (Greenwood, 2015: 443). Nor is there an internationally agreed standard for measuring internal migration, leading either to a count of migrations or a count of migrants depending on the data source. Drawing robust conclusions, identifying implications, or even calculating internal migration flows is hence a difficult undertaking. The field of internal migration studies is, as a consequence, limited in size (i.e. Kupiszewska and Nowok, 2008; Bell et al., 2015).

The second challenge relates to the popularity of migration flows across international borders, which has detracted attention from internal migration. Even though the bulk of global geographical mobility takes place within national borders, many might argue that the analysis of international migration has dominated both social sciences and the majority of policy discussions (Skeldon, 2006; Ellis, 2012). Studies on "population movements involving changes of residence within countries remain poorly developed" (Bell, et al., 2015: 33). The word 'migration' often seems to have been altered in meaning to refer almost exclusively to international migration. Internal migrants are

often subsumed under a ‘residential mobility’ or ‘population distribution’ category (Skeldon, 2006). According to many, the interest in internal migration is simply not proportional to the amount of people it concerns. Hence, the impact of domestic migration on economic development seems to remain “relatively understudied, undermeasured, and misunderstood” (Kuhn, 2015: 433).

Much of the substantive literature delving into internal migration dates from the 1980s and earlier (i.e. Thomas, 1936, 1941; Thomas and Kuznets, 1957; Åkerman, 1975; Masnick, 1986). This literature predominantly focuses on the patterns, streams, and changes across time in internal migration. Emphasising in particular the move between agricultural and urban as well as industrial areas, this research concentrates on social change. The insights we have gained on internal migrants over the more recent decades can be divided into three research strands – the individual level, the regional perspective, and migratory patterns. Most studies have converged on the individual, analysing the determinants of migration, the migrants’ characteristics, and individual returns. Individual features, such as age, schooling, marital status, health, job tenure, poverty or employment status, earnings, or retirement status have been analysed mainly in the context of individual utility maximisation models (i.e. Plane, 1993; Greenwood, 1997; Jung et al., 2004; Bell et al., 2015; Mohanty et al., 2016). Generally, internal migration is found to lead to positive individual returns in terms of earnings and employment opportunities (i.e. Yap, 1976; Lucas, 2004; Molloy et al., 2011) without necessarily improving subjective wellbeing (i.e. De Jong et al., 2002; Nowok et al., 2013; Sloan and Morrison, 2016).

A second group of studies adopts a regional perspective. Many focus on movements between rural and urban areas, evaluating social costs, brain drain, and integration issues (i.e. Huning and Huetl, 2012; Lerch, 2014; Eliasson et al., 2015; Rupasingha et

al., 2015). Place specific pull and push factors, such as the employment rate, wage differentials, tax rates, public services, local government quality, social capital, climate, and other local amenities are also covered (i.e. Hunter, 1998; Delisle and Shearmur, 2010; Biagi et al., 2011; Shen, 2013; Kuhn, 2015).

The final group evaluates the size, patterns, and trends of internal geographical mobility (i.e. Mueser, 1989; Borjas et al., 1992; Newbold and Bell, 2001; Molloy et al., 2011). Internal migration rates are calculated and compared across time and space, leading to policy recommendations (i.e. Shen, 2013; Aking and Dokmeci, 2015).

The macroeconomic consequences of internal migration and their economic impact, however, remain – to the best of our knowledge – largely overlooked by recent scientific research (White and Lindstrom, 2005). The few contemporary studies on domestic migration deliver inconsistent results. For some, internal migration leads to higher incomes, lower inequality, modernisation, and growth (e.g. Yap, 1976; IOM, 2005; Berker, 2011; Kuhn, 2015). Some specific internal migrant groups, such as, for example, college graduates, strongly improve macroeconomic outcomes (Moretti, 2012; Diamond, 2016). Other studies show that domestic population movements result in regional divergence and widening inequality, creating poverty traps, harming social ties, deteriorating the provision of public goods, and significantly lowering growth (i.e. Rodríguez Vignoli, 2008; Molloy et al., 2011). A third group fails to find any robust relationship between both factors (i.e. White and Lindstrom, 2005). In short, the limited research and wide range of findings in the literature on the macroeconomic impact of domestic migrants leaves us with no clear answers as to how internal migration shapes the economic growth trajectory of the receiving regions.

In order to establish a hypothesis about the potential link between internal migration and economic development, we therefore need to resort to the literature on international migration (Ellis, 2012). Pryor (1981), King and Skeldon (2010), and Ellis (2012) question the dichotomy in the field of migration studies between internal and international migration pushing towards an integrated system embracing the similarities: when analysing internal population flows in the simplest way, they could be defined, from a basic labour economics standpoint, as "a major mechanism through which labour resources are redistributed geographically" (Greenwood, 1997: 648). If we assume that the economic effect of internal migration mirrors that of international migration, domestic migrants will have a generally growth enhancing effect (i.e. Borjas, 1995; Card, 2005). The transmission channels identified by the international migration literature include the mere expansion of the labour force (i.e. Ortega and Peri, 2009), increasing returns to scale (i.e. Borjas, 1995), adjustments in the local market's skill and labour composition (i.e. Lundborg and Segerström, 2002), increases in wages (i.e. Carter and Sutch, 2006; Ottaviano and Peri, 2006), and the stimulation of productivity by means of innovation and skill set extensions (i.e. Alesina and La Ferrara, 2005; Gordon and McCann, 2005; Partridge and Furtan, 2008). Further growth potential can be derived from the increased population diversity as a result of labour inflows (Jacobs, 1969; Florida, 2002; Saxenian, 2006). Consequently, following Pryor (1981), King and Skeldon (2010), and Ellis (2012), internal migrants are expected – similarly to their international counterparts – to positively affect the growth trajectory of receiving territories.

Traditional migration literature tends to emphasise that one of the differentiating factors between international and internal migrants is the level of diversity in the receiving region. As migration research defines diversity mostly referring to birth

countries, ethnicities, or languages spoken (i.e. Alesina and La Ferrara, 2005), internal migrants are generally considered not to increase population diversity. Domestic migrants are ultimately native-born. They share the same birth country and often ethnicity and language with the local population in the receiving region. But these similarities do not necessarily mean that internal migrants do not enhance diversity. We beg to differ in this regard. Greater diversity within societies welcoming internal migrant inflows is related to the distance covered by internal migrants from point of origin to destination. Internal migrants come from the same country but often from faraway cities and regions, frequently with different habits and customs. While an Oregonian in Washington State will have had a short trip, a New Yorker in the same place would have covered more than 2,500 miles. This geographical distance and the diversity in places of origin – in spite of the fact that both domestic migrants speak the same language, share the same country of birth and possibly ethnic traits – distinguishes them. Hence the distance travelled represents an indicator of (dis)similarity in the institutional baggage internal migrants bring from home – individual traditions, customs, habits, and different mindsets are shaped by the place of birth. In his groundbreaking work on Italy, Putnam (1993) demonstrates that institutional constructs are highly place specific and vary greatly from one region to another. The closer two regions are, the greater the similarity between institutional constructs; the further away, the greater the difference (Arbia et al., 2010). A New Yorker in Washington State – especially in historical times – will have brought institutional constructs significantly different from those of the local Washingtonian. The institutional baggage brought by an Oregonian would have been less different. A New Yorker in Washington State would therefore have raised population diversity levels to a greater extent than an Oregonian. Hence, it could be hypothesised that the

geographical distance travelled by internal migrants affects the degree of (dis)similarity in a given place.

Population diversity itself is generally seen as central driver of innovation and creativity, positively influencing economic growth (Jacobs, 1961, 1969, Florida, 2002, 2012; Saxenian, 2006). A large variety of population groups stemming from a multitude of locations brings different skill sets, ideas, abilities, and experiences to their destinations. Population diversity thus becomes an important input factor in the process of economic growth: it promotes technological progress, productivity, and employment (Audretsch et al., 2010; Özgen et al., 2011; Kemeny and Cooke, 2017).

In brief, the geographical distance travelled by the internal migrant affects population diversity in the areas of destination, shaping their economic prospects. The greater the distance travelled by migrants, the greater the population diversity at destination, and, consequently, the higher the growth prospects of receiving territories. Large numbers of American-born settlers travelling over long distances would hence have contributed to transform the economic fortunes of receiving areas and, through their influence on local diversity levels, positively affected growth.

Geographical distance per se has only rarely been evaluated in migration research. Only a few studies calculate the distance travelled between place of origin and destination (e.g. White and Lindstrom, 2005). Data and definition inconsistencies, coupled with accuracy issues of distance measures, hamper once again the retrieval of robust results (Niedomysl et al., 2017). Apart from a few international comparisons (i.e. Long, et al., 1988), geographical distance has mostly been connected to migration when evaluating long distance in comparison to short distance migration. Dynamics, characteristics of migrants, and causes of migration were found to vary widely with

distance (i.e. Biagi et al., 2011; Pendakur and Young, 2013; Niedomysl and Fransson, 2014). Most studies report that distance is negatively correlated to the size of migration (i.e. Greenwood, 1997; Hipp and Boessen, 2016). Furthermore, while long distance moves are usually motivated by economic opportunities, short distance moves are more correlated with improvements in quality of life (Morrison and Clark 2011; Niedomysl, 2011). The effect of distance, however, on the macroeconomic impact of internal migration has been, in any case, neglected by the social science literature.

Moreover, past research has put the emphasis on the short- to medium-term impact of both internal and international migration. The focus has traditionally been on the immediate economic effects, covering a maximum of two decades since arrival of the migrant. Whether or how past migration affects regional economic performance after these initial years remains a black box. Rodríguez-Pose and von Berlepsch (2014) provide one of the few exceptions. They analyse 19th century international migrants and their impact on US economic development more than a century later, finding that migration improves the long-term development prospects of recipient areas.[33]

The persistence of growth enhancing effects of migration over the very long-term is often associated with the role of institutions in recipient areas (Acemoglu et al., 2001; Duranton et al., 2009; Tabellini, 2010). According to these studies, migrants convey their institutional constructs, "[…] in the way of culture, religion, social networks and links with the society of origin" (Joly, 2000: 30) from their place of origin to the destination region, preserving customs, traditions, habits, and mentality (Rice and Feldman, 1997). As Putnam (1993), Acemoglu et al. (2001), or Duranton et al. (2009) indicate, institutional frameworks persist in time, becoming engraved in the territory.

[33] Sequeira et al. (2017) support these findings.

Rodríguez-Pose and von Berlepsch (2014, 2015) theorise that the institutional frameworks derived from 19th century migration inflows into the US still shape current economic growth in the recipient areas. Whether the diversity of institutional constructs brought to places of destinations by short- and long-distance internal migrants has created a similarly positive and long-lasting economic effect remains an open question.

In this chapter, we tackle the aforementioned gaps in internal migration research by establishing a connection between shares of domestic migrants in a region's population, the distance covered by migrants, and economic development over the long-term. Two different research questions are examined: a) Do internal migrants shape long-term economic development in the same way as external ones? b) Does the distance covered by migrants matter for the influence they have on the subsequent growth of receiving regions?

7.4 EMPIRICAL APPROACH

7.4.1 THE MODEL

We estimate two different models to answer our research questions. Model 1, focusing on different migrant population shares, takes the following form:

$$y_{i,t} = \alpha + \beta Mig_{i,t_0} + \partial X_{i,t-k} + \theta Z_{i,t_0} + \mu state + \varepsilon_{i,t} \quad (1)$$

Where y is the natural log of income per capita of county i in period t (t=2010); *Mig* is our main variable of interest representing different specifications of the migrant population composition in a given county i at t_0 (t_0=1880 or 1910); X represents a vector of factors linked to income per capita levels of county i at time t-k (k=10) and Z corresponds to a vector of similar factors associated to the level of economic development of county i at time t_0, shaping the attractiveness of the county at the time of migration. Lastly, *state* represents state specific fixed effects controlling for arbitrary spatial correlations between counties within any given state and ε describes robust standard errors.

We estimate Model 1 in five different specifications, each using a variation of *Mig*. We first run (1) focusing on the share of total migrants – internal and external (foreign-born) – in a given county i at time t_0 (specification 1) measured as percentage of total county population. Subsequently, (1) is estimated distinguishing between the share of domestic migrants (*IntMig*), measured as the percentage of American-born residents having crossed state lines between their birthplace and current place of residence, and the share of external migrants (*ExtMig),* measured as the percentage of foreign-born in a county's population, as our two variables of interest (specification 2). Specifications 3, 4, and 5 focus on different American-born groups in county i's population at time t_0 while controlling for the share of external migrants. The American-born county population is divided into 'stayers', i.e. those born in the same state (specification 3), internal migrants from any neighbouring state (specification 4), and internal migrants from any other non-adjacent state (specification 5). Each group is measured as the percentage of total American-born population living in the county.

Model 2 estimates the direct relevance of the distance travelled by migrants on economic development 100 to 130 years later. Dependent and control variables remain the same as in Model 1. The independent variable of interest, however, changes to *Distance*, representing the average distance travelled by all migrants – both external and internal – of a given county *i*'s population at time t_0 between their birthplace and current county of residence (specification 1) and the distance covered by domestic migrants – American-born residents having crossed state lines – only (specification 2). The model takes the following form:

$$y_{i,t} = \alpha + \gamma Distance_{i,t_0} + \partial X_{i,t-k} + \theta Z_{i,t_0} + \mu state + \varepsilon_{i,t} \qquad (2)$$

7.4.2 THE DATA

The migration data used in the main variables of interest in Model 1 were constructed using the Integrated Public Use Microdata Series USA database (IPUMS) Version 6.0 (Ruggles et al., 2015). This database comprises representative population weighted samples covering US Censuses and American Community Surveys between 1850 and 2015.[34] Starting from 1850, the US Census included information on individual birthplaces, noted as either the country of birth for the foreign-born or the state of birth for the domestic population. Complete birthplace and current residence data for all continental US states are available only from 1880 onwards, allowing us to trace 'lifetime migration' (put simply, the migration between an individual's birthplace and his place of residence at the time of the Census). 1880 is thus selected as the main base year for the analysis – a 130-year timeframe between dependent and main independent

[34] The American Community Survey was only initiated in 2005.

variable of interest. The population data variable for 1910 is chosen to test the robustness of the 1880 results. As in 1910, the population and industry structure across the US already differed considerably from that of 1880, the 1910 sample represents an ideal candidate to validate the 1880 results, while keeping the long-term dimension. Unfortunately, neither inter-county or inter-state moves, nor the date of the location change, nor the duration of residence were recorded prior to 1935. These factors could unfortunately not be incorporated into our analysis.

In order to construct the main variables of interest, we use a sample of 5,791,531 individuals in 1880 representing 11.5 percent of the total US population at the time and 923,153 individuals in 1910 representing 1 percent of the population. All individual data were matched to the specific county of residence and aggregated at the county level. We define all US residents not born in the US as 'external migrants'. All American-born with a birthplace different from their state of residence are classified as 'internal or domestic migrants'.

Due to changes in size, geography and quantity of US counties over the period of analysis, counties in 1880 and 1910 were matched to their 2010 equivalent using cartographic boundary files provided by the US Census of the 48 continental states. All county boundaries were normalised to their 2010 borders, historical county averages were calculated and weighted by the population density at the time of the boundary change. This method allowed us to attribute historical county features to all counties of the 48 continental US states in 2010 (with the exception of 1880 values for Oklahoma).[35]

[35] Oklahoma only became organised territory in 1890.

The second set of independent variables of interest – the average distance covered by the migrant population living in a given county – was constructed using GIS software calculating the point-distance matrix between the centroids of all US counties of residence and the centroids of the 48 continental states (weighted by the population density at the time of migration), as well as all countries named as birthplaces by the foreign-born population. The individual distance travelled by each migrant was then allocated to the given county of residence and aggregated at county level. Two different specifications – average distance covered by all migrants (both external and internal) and average distance covered by internal migrants only – are calculated. As within state migration is not recorded in the data, all individuals within the population of a county who were born in the same state as their current county of residence are assumed to either having moved only across county lines or not at all.

The dependent variable (income per capita levels in 2010 in US dollars) as well as vector *X*, containing data for 2000, employ information extracted from the US Bureau of Economic Analysis (BEA) database, the Current Population Survey (CPS) tables of the US Bureau of Labor Statistics (BLS), and the 2000 Census Summary files. Vector *X* is measured 10 years before the dependent variable and is included in the model to control for county level characteristics directly affecting the level of income per capita of a given US county. *X* is included as a means of avoiding that recent county features could transfer their individual association to current levels of economic development onto the variable of interest. This would potentially lead to over- or underestimating its effect. The factors considered include educational attainment (share of people with college education), the unemployment rate, the share of black and female population, overall population size (as natural log), the share of the labour force employed in manufacturing, the infant mortality rate as a proxy for

levels of public health, and the women's participation rate in the labour force. All variables included in vector *X* are lagged by 10 years with respect to the period considered in our dependent variable to minimise problems of simultaneity between county features and income per capita. The descriptive statistics for all variables are provided in Appendix 7C.

Vector *Z*, the second vector of control variables, dates from 1880 and 1910, respectively. It consists of county features that may have influenced the county's level of economic prosperity at the time of migration. Simultaneously, these characteristics may have served as pull factors to migrants determining the level of attractiveness of a county in that period. The inclusion of *Z* follows a similar reasoning as X. By controlling for county features at the time of migration, we extract their potentially confounding influence on the regression results from the error term and include it explicitly in the regression. Their potential impact on economic development today can therefore not be transferred onto the internal migration variables. Vector *Z* includes, whenever possible,[36] the same variables as vector *X*. Educational attainment, however, is measured as the literacy rate. Furthermore, we control for the initial county level average income at the time of migration. As income per capita data were not collected at the time, a proxy is constructed with individual data on median total income per occupation in hundreds of 1950 dollars using the size of the total county population in 1880/1910 as base. All of the historical variables are constructed using the IPUMS USA and the Inter-University Consortium for Political and Social Research databases (ICPSR). All individual data were aggregated at the county level,

[36] Issues with data availability imply that not all variables can be reproduced exactly.

employing the same method used when constructing the migration variables. A description of all variables and sources is provided in Appendix 7D.

7.4.3 INSTRUMENTAL VARIABLE ESTIMATION

Any analysis of long-term migration data is prone to endogeneity issues. Potential spatial sorting, omitted variables, or reverse causality issues can pose threats to the internal validity and hence distort retrieved OLS results. In order to identify the true underlying linkage of internal migration and economic development, possible endogeneity biases in the least-squares estimates are addressed using an instrumental variable (IV) estimation. For both models, geography serves as an exogenous source of variation. Two instruments are selected: topography – as instrument for internal migration (Model 1) – and size of the water area in a county – instrumenting the average distance travelled (Model 2). Both instruments satisfy the relevance criterion of IV analysis as both retrieve sufficiently large first stage F-statistics based on the Staiger and Stock (1997) test for weak instruments in combination with the Stock and Yogo (2005) critical values.

In Model 1, topography is instrumented for internal migrants. The topography variable is extracted from the National Atlas of the USA (US Geological Survey), published as part of the US Natural Amenities Scale of the US Department of Agriculture (see detailed description in Appendix 7E). The 21-level scale categorises land surface forms at county level, ranging from flat plains and tablelands to hillsides and mountains. Topography is exogenous to income per capita in 2010 and hence highly likely to be fully uncorrelated with the error term. We assume that topography affects

economic development at county level via the settlement pattern of internal migrants. Topography was a crucial factor behind domestic migrant settlement patterns over 100 years ago. The interaction of topography with climate and vegetation affected soil quality and served as a crucial pull factor, raising incentives for migrants to pack up and move. Fresh good quality soils "[drew] migrants from worn-out lands in the East to virgin lands in the West" (Merk, 1978: 229). Cattle farmers, hunters, trappers, and miners were attracted by the vast lands of plateaus and hill sides. Grain, pork farmers, and wool producers settled in the plains, while fur traders saw high potential in the plateaus and basins of the northwest. The Gulf plains to the south attracted internal migrants focusing on cotton and sugar production. Rough mountains and sterile hills, by contrast, were avoided by migrants (Merk, 1978).

The variation in the average distance travelled by migrants (Model 2) is instrumented using the size of water areas within a county. The data, measured in square miles, were retrieved from the TIGER Geodatabases by the US Census Geography division. Water areas, again exogenous to income per capita levels in 2010, were a decisive element for the migrant settlement pattern in the 19th century as well as decisive in determining the distance travelled by migrants. Navigable rivers provided important infrastructure as well as vital access to drinking water and irrigation of nearby lands. The better the waterways and water supply, the farther the migrant travelled. Waterways shaped long distance migrant routes which, in turn, gave rise to migrant settlements along the way. Mountain ranges or rough landscapes such as the Appalachians or the Rocky Mountains could only be crossed following rivers, cutting deep canyons and valleys into the abrupt terrain. The main migrant trails across the country, such as the Oregon trail in the northwest crossing the northern part of the Rockies, followed tightly knit river systems. Even when the railroad network had considerably grown in size and

efficiency, waterways remained a fundamental element in the country's infrastructure for both transport as well as internal migrant settlement (Merk, 1978). Waterways facilitated the journey, enlarged the distances travelled, and made quick progress possible. A lack of water, however, shortened migration routes considerably.

7.5 ANALYSIS

The first part of the analysis focuses on whether internal migrants criss-crossing the US during the 19th century left a significant and long-lasting impact on their receiving territories. Model 1 is employed first using ordinary least squares (OLS), followed by an instrumental variable (IV) estimation as a robustness check of our results. Table 7-3 displays the OLS results of Model 1 specifications (1) through (5) for 1880, each time altering the main variable of interest.

In line with expectations, migration is positively associated with income per capita levels in specification (1). A large share of migrants in 1880, regardless of birthplace, is strongly and positively connected to county level GDP per capita in 2010. When splitting the migrant stock of a county into internal and external migrants (specification 2), the coefficients for both types of migrants remain positive and highly significant. Hence, counties that attracted large inflows of US and foreign settlers are significantly more prosperous today than those largely bypassed by migration. The coefficient for external migrants is, however, larger than that for domestic migrants. The more than four-fold difference between both coefficients, statistically significant below the 1% level, hints at foreign migrants as a more powerful influence on the long-term prosperity of US counties than their American-born counterparts.

Regressions (3) to (5) shed light on the connection between different American-born population groups within a given county in 1880 and GDP per capita in 2010, using the share of external migrants as control variable. We first focus on the share of the population living in their state of birth in 1880 (regression 3), including both individuals living in their county of birth and those who migrated within state borders. Data unavailability on intra-state moves means that these two groups have to be merged. As the group of stayers dominates over intra-state migrants, the results predominantly reflect the impact of a large group of locals on income per capita more than a century later. Regression (3) reveals that counties with a large percentage of 'stayers' in 1880 had lower GDP per capita levels 130 years later: counties bypassed by migration – other than intra-state – more than a century ago were significantly poorer in 2010. Hence, a largely homogenous county population in 1880 has represented a serious barrier for long-term development.

Regressions (4) and (5) in Table 7-3 focus on domestic inter-state migration, distinguishing between migration from neighbouring states on the one hand and from non-adjacent states on the other. The former is displayed in regression (4). No significant relation between the share of a county's population from neighbouring states and average income per capita in 2010 is found. One possible explanation for this result may be the similarities in institutional baggage, culture, traditions, customs and mentality between locals and migrants from adjacent states. One might argue, this type of migrant not be adding sufficient diversity, not to be sufficiently risk-taking to trigger additional economic dynamism.

Regression (5) focuses on interstate migration from non-adjacent states. The results point to a strong positive and significant relation between the current economic development of US counties and the presence of domestic migrants from more distant

locations in 1880. Compared to our main variables of interest in regressions (3) and (4), the significance level and the size of the coefficient are largest in (5). Domestic migrants travelling farther seem to have left a positive and enduring effect on the economic development of their receiving counties. A more heterogeneous and thus more diverse population composition in the receiving territory is connected to higher economic development levels. In this respect, American-born migrants moving over long distances seem to have left a similar economic imprint on local economic development than that of external migrants. The economic legacy of locals and migrants from neighbouring states is, in contrast, much more limited.

These results hint to the fact that the distance travelled by migrants has a considerable effect on long-term county economic growth. By dividing internal migrants into three groups – 'stayers' and within-state migrants, migrants from adjacent states, and long-distance migrants – domestic migrants have been implicitly grouped by distance travelled between birthplace and current residence. The results indicate that the bigger the distance travelled, the greater the long-term economic legacy of internal migration.

TABLE 7-3 THE LONG-TERM IMPACT OF INTERNAL MIGRATION, OLS 1880

Dep. Var.: income per capita 2010 (ln)	(1)	(2)	(3)	(4)	(5)
All migrants 1880	0.138***				
	(0.0251)				
Internal migrants 1880		0.0587**			
		(0.0282)			
Pop. same state 1880			-0.0587**		
			(0.0282)		
Pop. neighbour state 1880				-0.0489	
				(0.0360)	
Pop. rest of country 1880					0.134***
					(0.0388)
External migrants 1880		0.260***	0.260***	0.239***	0.279***
		(0.0401)	(0.0401)	(0.0400)	(0.0399)
Education 2000	0.0114***	0.0112***	0.0112***	0.0112***	0.0112***
	(0.000756)	(0.000750)	(0.000750)	(0.000751)	(0.000749)
Population 2000 (ln)	0.00388	0.00367	0.00367	0.00344	0.00340
	(0.00363)	(0.00366)	(0.00366)	(0.00367)	(0.00365)
Manufacturing 2000	-0.145***	-0.144***	-0.144***	-0.142***	-0.144***
	(0.0244)	(0.0242)	(0.0242)	(0.0242)	(0.0243)
Black population 2000	-0.00110***	-0.00101***	-0.00101***	-0.000950***	-0.00102***
	(0.000345)	(0.000340)	(0.000340)	(0.000339)	(0.000339)
Female 2000	0.0148***	0.0152***	0.0152***	0.0155***	0.0153***
	(0.00152)	(0.00153)	(0.00153)	(0.00153)	(0.00152)
Female part. 2000	0.00121	0.00112	0.00112	0.00106	0.00114
	(0.000786)	(0.000782)	(0.000782)	(0.000779)	(0.000781)
Unemployment 2000	-0.0203***	-0.0222***	-0.0222***	-0.0225***	-0.0220***
	(0.00320)	(0.00327)	(0.00327)	(0.00324)	(0.00327)
Infant mortality 2000	0.000201	0.000211	0.000211	0.000242	0.000214
	(0.000434)	(0.000444)	(0.000444)	(0.000449)	(0.000439)
Income 1880 (ln)	0.0133	0.0125	0.0125	0.0161	0.00295
	(0.0194)	(0.0192)	(0.0192)	(0.0190)	(0.0191)
Literacy 1880	0.133***	0.161***	0.161***	0.170***	0.142***
	(0.0293)	(0.0307)	(0.0307)	(0.0307)	(0.0311)
Population 1880 (ln)	-0.00996**	-0.0128***	-0.0128***	-0.0128***	-0.0137***
	(0.00440)	(0.00449)	(0.00449)	(0.00454)	(0.00447)
Manufacturing 1880	0.0105	-0.0189	-0.0189	-0.0175	-0.0114
	(0.0388)	(0.0374)	(0.0374)	(0.0375)	(0.0372)
Black population 1880	0.230***	0.241***	0.241***	0.227***	0.227***
	(0.0377)	(0.0378)	(0.0378)	(0.0378)	(0.0376)
Female 1880	0.101	0.113	0.113	0.0761	0.179
	(0.118)	(0.116)	(0.116)	(0.115)	(0.117)
Female part. 1880	-0.0667	-0.0621	-0.0621	-0.0462	-0.0482
	(0.0988)	(0.0974)	(0.0974)	(0.0964)	(0.0956)
Unemployment 1880	-0.00237	-0.00460	-0.00460	-0.00503	-0.00431
	(0.0115)	(0.0116)	(0.0116)	(0.0117)	(0.0115)
Observations	2,440	2,440	2,440	2,440	2,440
States	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
R-squared	0.682	0.686	0.686	0.686	0.688

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

The significance levels and signs of the control variables reinforce the validity of the results. All variables included in the two vectors controlling for wealth influencing factors both at the time of migration as well as today show the expected signs. The lagged controls indicate that counties with a more educated, gender-balanced population, less employment in manufacturing, a smaller share of black inhabitants, and a lower unemployment rate were richer in 2010 than those with a less educated, largely male population, with a high share of black people, large employment in manufacturing, and a high unemployment rate.

Most 1880 control factors are insignificant, meaning that whatever influence they had on economic development a century ago has waned or disappeared altogether. The coefficient for the initial income level – one of the potentially strongest pull factors at the time of migration – is insignificant. There are some exceptions, though. Literacy and the size of the black population of a county in 1880 are strongly positively connected to current levels of development. By contrast, the size of a county's population more than a century ago is associated with lower levels of development. We assume the black population variable to serve as proxy for the economic structure of the largely agrarian counties of the South, rather than indicating the effect of a polarised county population into black and white. The coefficient of the share of black population is highly likely to capture the path to convergence of the poor regions of the South to the richer regions in the North (see for example Barro and Sala-i-Martin, 1990, 1995).

The OLS results reported above could nevertheless be affected by endogeneity issues caused by reverse causality, sorting, or omitted variable bias. In order to address these

issues, two robustness checks are undertaken. We first shift the base migration year by 30 years to 1910, estimating again Model 1, including all five specifications of the independent variable of interest. The results are displayed in Appendix 7F. Apart from slightly weaker significance levels, the coefficients for the different groups of internal migrants largely coincide with those obtained when using 1880 as base year. Both a large share of internal as well as external migrants – taken together in regression (1) and apart in regression (2) – is significantly associated with higher levels of economic development one century later. A large percentage of Americans still living in their state of birth is negatively connected to long-term growth (regression (3)). The coefficient for the relationship between income per capita levels in 2010 and the share of internal migrants from neighbouring states in 1910 remains insignificant (regression 4). Again, the largest and most significant coefficient of the internal migrant subgroups pertains to the group of internal migrants from non-adjacent states (5).

A second robustness check, an IV estimation is performed, employing topography as an instrument for the five different migrant population subgroups in a given US county in 1880. Table 7-4 displays the results.[37]

When using topography as IV to retrieve the underlying effect of past internal migration flows on income per capita levels more than 100 years later, the large majority of the OLS results are validated. Internal as well as external migration are confirmed as growth enhancing factors over the very long-term – either taken together as in regression (1) or apart as displayed in regression (2) of Table 7-4. Consequently, counties which attracted large inflows of both foreign and American-born migrants at the end of the 19th century display significantly larger income per capita levels in 2010

[37] The following discussion of the results is equally valid for the IV analysis shifting our base year to 1910. Similar results for 1910 to those displayed in Table 7-4 can be found in Appendix 7G.

than those that lacked a strong migrant pool at the time. Large numbers of migrants, regardless of origin, have become an essential factor behind county level economic development for more than a century.

The results of the OLS analysis are, once again, largely validated when focusing on the different American-born population groups in regressions (3) to (5) in Table 7-4. Counties bypassed by out-of-state domestic migration more than a century ago are poorer today than those that attracted migrants in considerable numbers (regression (3)). Having a large percentage of locals, and therefore low population diversity levels, has been detrimental for long-term economic growth. By contrast, any type of cross-border internal migration has resulted in greater long-term economic dynamism and wealth at county level. This applies for both migrants from neighbouring states (regression 4) and from farther away (regression (5)). The sway of domestic migrants crossing at least two state lines in their journey to a better future remains in any case considerably higher, as indicated by the coefficient in regression (5). Distance thus seems to largely explain the relation between internal migrants and long-term economic performance. The larger the distance travelled by internal migrants, the stronger their long-term impact on county income per capita levels.

TABLE 7-4 THE LONG-TERM IMPACT OF INTERNAL MIGRATION, IV 1880

Dep. Var.: income per capita 2010 (ln)	(1)	(2)	(3)	(4)	(5)
All migrants 1880	1.123***				
	(0.258)				
Internal migrants 1880		1.497***			
		(0.433)			
Pop. same state 1880			-1.497***		
			(0.433)		
Pop. neighbour state 1880				2.299***	
				(0.787)	
Pop. rest of country 1880					4.291*
					(2.297)
External migrants 1880		0.730***	0.730***	0.341***	1.456**
		(0.151)	(0.151)	(0.0743)	(0.647)
Education 2000	0.0109***	0.0114***	0.0114***	0.0121***	0.0102***
	(0.000820)	(0.000895)	(0.000895)	(0.00102)	(0.00154)
Population 2000 (ln)	0.00550	0.00622	0.00622	0.0107	-0.00214
	(0.00493)	(0.00520)	(0.00520)	(0.00695)	(0.0114)
Manufacturing 2000	-0.163***	-0.169***	-0.169***	-0.168***	-0.170**
	(0.0346)	(0.0393)	(0.0393)	(0.0441)	(0.0750)
Black population 2000	-0.00170***	-0.00212***	-0.00212***	-0.00173***	-0.00284**
	(0.000472)	(0.000589)	(0.000589)	(0.000578)	(0.00125)
Female 2000	0.0128***	0.0110***	0.0110***	0.00990***	0.0131***
	(0.00192)	(0.00238)	(0.00238)	(0.00286)	(0.00409)
Female part. 2000	0.00172	0.00209*	0.00209*	0.00145	0.00330
	(0.00108)	(0.00121)	(0.00121)	(0.00127)	(0.00247)
Unemployment 2000	-0.0231***	-0.0162***	-0.0162***	-0.0197***	-0.00985
	(0.00525)	(0.00569)	(0.00569)	(0.00507)	(0.0115)
Infant mortality 2000	-0.000213	-0.000283	-0.000283	-0.000255	-0.000335
	(0.000601)	(0.000654)	(0.000654)	(0.000752)	(0.00138)
Income 1880 (ln)	-0.0995**	-0.104**	-0.104**	0.0760	-0.441*
	(0.0401)	(0.0463)	(0.0463)	(0.0552)	(0.258)
Literacy 1880	0.0225	-0.0934	-0.0934	0.276***	-0.784
	(0.0603)	(0.0942)	(0.0942)	(0.0729)	(0.542)
Population 1880 (ln)	-0.0285***	-0.0188***	-0.0188***	-0.00272	-0.0487*
	(0.00774)	(0.00687)	(0.00687)	(0.0106)	(0.0256)
Manufacturing 1880	-0.103	0.000230	0.000230	-0.132	0.247
	(0.0700)	(0.0662)	(0.0662)	(0.0849)	(0.193)
Black population 1880	0.417***	0.390***	0.390***	0.596***	0.00586
	(0.0672)	(0.0675)	(0.0675)	(0.130)	(0.172)
Female 1880	1.106***	1.135***	1.135***	-0.128	3.493*
	(0.321)	(0.370)	(0.370)	(0.251)	(1.888)
Female part. 1880	-0.235*	-0.258*	-0.258*	-0.473**	0.145
	(0.134)	(0.154)	(0.154)	(0.237)	(0.319)
Unemployment 1880	-0.00383	0.00460	0.00460	-0.00180	0.0165
	(0.0167)	(0.0184)	(0.0184)	(0.0192)	(0.0426)
Observations	2,435	2,435	2,435	2,435	2,435
States	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
First stage F-stat	36.29	20.45	20.45	15.00	8.16

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

In order to assess in a more accurate way the extent to which the average distance covered by internal migrants affects the link between migration and long-term regional economic growth, Model 2 is run in its two different specifications focusing on, first, the average distance travelled by the entire migrant population – both foreign and American-born – of a given county in 1880 and, second, on the average distance covered exclusively by domestic migrants who have crossed state lines between their state of birth and destination. Table 7-5 reports the results for the OLS (regressions 1 and 2) as well as IV estimations (regressions (3) and (4)) for the base year 1880.

Regression (1) shows that the average distance travelled by the entire migrant population in a county has a positive and highly significant impact on long-term economic development: the farther a migrant travelled before settling down, the stronger the impact. As this positive effect can be driven by the presence of international migrants having travelled for thousands of kilometres, in a second step the analysis is performed focusing exclusively on American-born migrants having crossed state lines. Regression (2) displays the results.[38] Again, the connection between the average distance travelled by internal migrants and the long-term economic performance of the receiving counties is positive and strongly significant (regression (2)). Consequently, the more a county managed to attract large shares of internal migrants from faraway places, the higher the diversity in its population and the higher its income per capita 130 years later.

The coefficients for both sets of controls are in line with those reported for the previous analysis.

[38] As birthplace data are not available at county level, within state migration is not considered in the analysis as a separate category. In those cases, the migration distance between birth state and destination is assumed to be zero.

TABLE 7-5 THE LONG-TERM IMPACT OF AVERAGE MIGRATION DISTANCE, 1880

Dep. Var.: income per capita 2010 (ln)	OLS		IV	
	(1) All Migrants	(2) Internal migrants	(3) All migrants	(4) Internal migrants
Distance all mig (ln) 1880	0.0376***		0.199***	
	(0.00550)		(0.0509)	
Distance int. mig (ln) 1880		0.0312***		0.485***
		(0.0102)		(0.122)
Education 2000	0.0114***	0.0114***	0.0114***	0.0102***
	(0.000758)	(0.000764)	(0.000790)	(0.000946)
Population 2000 (ln)	0.00261	0.00324	-0.00346	-0.00745
	(0.00361)	(0.00364)	(0.00447)	(0.00528)
Manufacturing 2000	-0.137***	-0.139***	-0.112***	-0.0910**
	(0.0246)	(0.0246)	(0.0332)	(0.0389)
Black population 2000	-0.000986***	-0.00101***	-0.000748*	-0.000841*
	(0.000343)	(0.000344)	(0.000423)	(0.000489)
Female 2000	0.0157***	0.0153***	0.0176***	0.0171***
	(0.00153)	(0.00153)	(0.00192)	(0.00206)
Female participation 2000	0.000834	0.00116	-0.000292	0.00220**
	(0.000778)	(0.000786)	(0.000929)	(0.00104)
Unemployment 2000	-0.0195***	-0.0197***	-0.0178***	-0.0166***
	(0.00312)	(0.00313)	(0.00375)	(0.00440)
Infant mortality 2000	0.000229	0.000268	0.000115	0.000478
	(0.000442)	(0.000436)	(0.000538)	(0.000546)
Income 1880 (ln)	0.0150	0.0226	-0.0467	-0.0788**
	(0.0190)	(0.0195)	(0.0285)	(0.0367)
Literacy 1880	0.120***	0.137***	0.0108	0.0111
	(0.0292)	(0.0294)	(0.0487)	(0.0547)
Population 1880 (ln)	-0.00864*	-0.00669	-0.0127***	0.00555
	(0.00442)	(0.00451)	(0.00488)	(0.00633)
Manufacturing 1880	0.00203	0.0326	-0.105**	0.116**
	(0.0373)	(0.0402)	(0.0487)	(0.0560)
Black population 1880	0.186***	0.192***	0.107**	0.0257
	(0.0377)	(0.0377)	(0.0517)	(0.0686)
Female 1880	-0.0468	-0.0277	-0.0698	0.164
	(0.112)	(0.115)	(0.113)	(0.135)
Female participation 1880	-0.0509	-0.0441	-0.0935	-0.0130
	(0.0975)	(0.0984)	(0.111)	(0.130)
Unemployment 1880	-0.00598	-0.00391	-0.0144	-0.0192
	(0.0116)	(0.0118)	(0.0143)	(0.0161)
Observations	2,444	2,441	2,444	2,441
States	*Yes*	*Yes*	*Yes*	*Yes*
R-squared	0.684	0.678	-	-
First stage F-stat	-	-	24.35	28.27

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

Two robustness checks are undertaken to validate these results. Firstly, as in the analysis of Model 1, we shift our base year by 30 years and run Model 2 based on 1910 migrants and average distances travelled. The results are displayed in Appendix

7H and present an almost carbon copy of those reported for 1880. Average distance travelled by the migrant population of any given US county in 1910 is in both cases highly significant and positively associated with income per capita levels in 2010. No matter which base year, it seems that long distance migrants leave a greater economic trace than short distance ones.

Secondly, to address potential endogeneity issues, an IV estimation is conducted. In Model 2, we use the size of water areas in a given county as instrument for the average distance travelled by migrants. The results, displayed in regressions (3) and (4) of Table 7-5, reconfirm the findings of the OLS analysis: the average distance travelled by the migrant population of any given county in 1880 between birthplace and current residence has a positive impact on long-term economic growth. Both the average distance travelled by the total migrant population as well as by the internal migrant subgroup display positive and highly significant coefficients. The presence of a higher share of long-distance domestic migrants in 1880 has therefore been strongly beneficial for long-term economic growth at the county level.

In short, US domestic 19th century migrants have been highly beneficial for the long-term economic development of their receiving counties. A larger share of internal migrants more than 100 years ago has determined the economic fortunes of US counties in a way that is still evident today. Distance is shown to act as crucial factor in this relationship. The larger the average distance travelled by the average internal migrant, the longer lasting the positive economic effect of migration. Simply crossing county lines does not do the trick. The biggest benefits were reaped by counties that attracted migrants travelling over long distances, precisely those who, on the one hand, are more likely to bring a different institutional baggage from that of the local population and, on the other hand, are more risk-taking and economically dynamic. A

more heterogeneous population, marked by the adventurous and entrepreneurial spirit of the long distance traveller seeking a new life in a faraway place, seems to have generated the seed of long-term economic development.

7.6 CONCLUSION

With international migration high on the priority list of current policy discussions, the big majority of migration research has bypassed within-country migration. The complexity in the definition of what constitutes internal migration in conjunction with measurement issues has further contributed to the small number of contemporary studies on the macroeconomic effect of domestic migrants. The limited knowledge gained has generally focused on the individual, analysing the migrants' characteristics, the determinants, and the returns of the migration decision. Regional-level research has emphasised push and pull factors, patterns, and trends of internal population movements, but the macroeconomic impact of internal migrants still remains poorly understood. Results – mainly focusing on the short-term impact – are far from conclusive: some have stressed that internal migrants promote economic prosperity, lower inequality, modernisation, and increased growth; others argue that migration fosters divergence in wages, income and employment, generating poverty traps and lowering growth. The study of the long-term impact of internal migration on regional economic development as well as the analysis of geographical distance covered as a crucial factor determining the long-term economic legacy of internal migration have, however, been neglected by past research. This chapter has addressed these important gaps in the literature, examining the impact of domestic US migration and the average distance covered by migrants in 1880 and 1910 on the wealth of US counties in 2010.

Two main research questions have driven the analysis: (1) Did late 19th century internal migrants transform the economy of their counties of destination in a way that can still be felt today? Was the settlement pattern of internal migrants more than 100 years ago decisive for ensuing economic performance? and (2) Does geographical distance matter for the impact of internal migration on long-term economic development? Do internal migrants travelling longer distances have a different impact on regional economic development over the long-term than those from nearby counties and states?

The findings confirm that internal migrants having crossed state lines between their birth state and destination exert a significant and positive long-term impact on the economic performance of their receiving regions. They leave a trace which is still evident more than 100 years after the settlement took place. Counties that attracted a large share of domestic migrants around the turn of the 20th century became and remain more prosperous in 2010 than those largely bypassed by internal out-of-state migration streams. Similar to their foreign-born counterparts (Rodríguez-Pose and Berlepsch, 2014; Sequeira et al., 2017), 19th century internal migrants have been a powerful force in the subsequent economic dynamism of US counties.

The geographical distance travelled by migrants is shown to have played a decisive role in this relationship. When analysing the average distance covered by internal migrants, the bigger the distance travelled, the greater the positive long-term economic legacy. Counties which attracted a large number of long distance migrants around the turn of the 20th century have been more dynamic over the next century. Large shares of population either born locally or in the same state have resulted in significantly lower regional economic development over the long-term.

Over the past century, internal migrants – and especially those having crossed vast distances to arrive at their final destination – have acted as decisive force for regional economic growth in the US. Unfortunately, we can only make assumptions about the exact mechanisms as to why and how this is the case. Bringing their habits, customs, and institutions from far-flung areas of the country to their receiving regions more than 100 years ago, we speculate that they have increased regional diversity levels creating new heterogeneous societies. The venturesome, hard-working and risk-seeking spirit associated with the long-distance migrant brought novel ideas, experiences and abilities of the east to the otherwise quite homogenous local born populations of the western territories. Therefore, places which did not manage to attract internal long-distance migrants – those which remained largely homogenous in their population composition – did not achieve a similar economic dynamism as they lacked the enlarged skillset, the different ideas, experiences and abilities long distance migrants brought along with them. They could not benefit from the economic boost elevated population diversity entails.

Thus, internal long-distance migrants travelling over vast distances played a major role in raising local diversity levels, creating a fruitful soil for economic dynamism which lies at the root of a territorial prosperity that is much longer lived than could have been imagined. Consequently, historic internal migrants represented much more than pure increases in the local supply of labour or a mere population redistribution at one point in time. They planted the seeds of a remarkably resilient long-term prosperity – a legacy which has determined the economic dynamism and vibrancy of places in America for more than a century. However, further research following a case study approach will be needed to prove our assumptions and add further detail to properly evaluate these transfer mechanisms across time.

Ignoring this important legacy of worker mobility is dangerous. In a country made by migrants, short sighted and short-term migration policies may stymie the possibility of very important long-term economic impacts associated with facilitating the mobility of people looking for jobs and a better life for themselves and their children. Limiting this type of migration may consequently mean missing out on a huge economic potential and on a force for sustainable economic development which will be felt not just for years, but for decades and even centuries to come.

7.7 REFERENCES

Acemoglu, D., Johnson, S., and Robinson, J. (2001): The colonial origins of comparative development: An empirical investigation. *American Economic Review*, *91*(5), 1369–1401.

Alesina, A., and La Ferrara, E. (2005): Ethnic diversity and economic performance. *Journal of Economic Literature*, *43*(3), 762–800.

Åkerman, S. (1975): Internal migration, industrialisation and urbanisation (1895–1930): a summary of the Västmanland study. *Scandinavian Economic History Review*, *23*(2), 149–158.

Akın, D., and Dökmeci, V. (2015): Cluster analysis of interregional migration in Turkey. *Journal of Urban Planning and Development*, *141*(3).

Arbia, G., Battisti, M. and Di Vaio, G. (2010): Institutions and geography: Empirical test of spatial growth models for European regions. *Economic Modelling*, *27*(1), 12-21.

Atack, J. and Passell, P. (1994): *A new economic view of American history: From colonial times to 1940*. P. Passell and S. Lee (Eds.). Second edition, New York: Norton.

Atack, J., Bateman, F., and Parker, N. (2000): Northern agriculture and the westward movement. In R. Engerman and R. Gallman (Eds.), *The Cambridge Economic History of the United States: The Long Nineteenth Century*, 285–328. Cambridge: Cambridge University Press.

Audretsch, D., Dohse, D. and Niebuhr, A. (2010): Cultural diversity and entrepreneurship: a regional analysis for Germany, *An International Journal of Urban, Regional and Environmental Research and Policy, 45*(1), 55–85.

Barro, R and X. Sala-i-Martin (1990): Economic growth and convergence across the United States. *NBER Working Paper Series*, No. 3419.

Barro, R and X. Sala-i-Martin (1995): Capital mobility in neoclassical models of growth. *The American Economic Review, 85*(1), 103-115.

Bell, M., Charles-Edwards, E., Kupiszewska, D., Kupiszewski, M., Stillwell, J., and Zhu, Y. (2015): Internal migration data around the world: Assessing contemporary practice. *Population, Space and Place*, *21*(1), 1–17.

Berker, A. (2011): Labor-market consequences of internal migration in Turkey. *Economic Development and Cultural Change*, *60*(1), 197–239.

Biagi, B., Faggian, A., and McCann, P. (2011): Long and short distance migration in Italy: The role of economic, social and environmental characteristics. *Spatial Economic Analysis*, *6*(1), 111–131.

Borjas, G., Bronars, S., and Trejo, S. (1992): Self-selection and internal migration in the United States. *Journal of Urban Economics*, *32*(2), 159–185.

Borjas, G. (1995): The economic benefits from immigration. *Journal of Economic Perspectives*, *9*(2), 3–22.

Card, D. (2005): Is the new immigration really so bad? *Economic Journal*, *115*(507), 300–323.

Carter, S. and R. Sutch (2006): Resident impacts of immigration: Perspectives from America's Age of Mass Migration. *Working Paper No 200808*, University of California at Riverside.

Delisle, F., and Shearmur, R. (2010): Where does all the talent flow? Migration of young graduates and nongraduates, Canada 1996-2001. *Canadian Geographer-Géographe Canadien*, *54*(3), 305–323.

DeTocqueville, A. (1839): *Democracy in America*. Third edition, Henry Reeve (trans.), New York: George Adlard.

Diamond, R. (2016): The determinants and welfare implications of US workers; Diverging location choices by skill: 1980–2000. *American Economic Review*, *106*(3), 479–524.

Duranton, G., Rodríguez-Pose, A and Sandall, R (2009): Family types and the persistence of regional disparities in Europe. *Economic Geography, 85*(1), 23–47.

Eliasson, K., Westlund, H., and Johansson, M. (2015): Determinants of net migration to rural areas, and the impacts of migration on rural labour markets and self-employment in rural Sweden. *European Planning Studies*, *23*(4, SI), 693–709.

Ellis, M. (2012): Reinventing US internal migration studies in the age of international migration. *Population, Space and Place*, *18*(2), 196–208.

Ferrie, J. (2005): History lessons: The end of American exceptionalism? Mobility in the United States Since 1850. *Journal of Economic Perspectives*, *19*(3), 199–215.

Florida, R. (2002): *The rise of the creative class and how it's transforming work, leisure, community and everyday life*. New York, NY: Basic Books.

Florida, R. (2012): *The rise of the creative class: revisited*. New York, NY: Basic Books.

Gallman, R. (2000): Economic growth and structural change in the long nineteenth century. In R. Engerman and R. Gallman (Eds.), *The Cambridge Economic History of the United States: The long Nineteenth Century*, 1–56. Cambridge: Cambridge University Press.

Gill, I., and Raiser, M. (2012): *Golden growth: Restoring the lustre of the European economic model*. The World Bank, Washington, DC. Retrieved 14.10.2017 under http://documents.worldbank.org/curated/en/539371468036253854/Main-report

Gordon, I. and McCann, P. (2005): Innovation, agglomeration, and regional development. *Journal of Economic Geography*, *5*, 523–543.

Greenwood, M. (1997): Chapter 12 Internal migration in developed countries. *Handbook of Population and Family Economics*, *1*, 647–720.

Greenwood, M. (2015): Internal migration in industrialized countries. In *International Encyclopaedia of the Social and Behavioural Sciences,* 443–446.

Haines, M. (2000): The population of the United States, 1790–1920. In S. Engerman and R. Gallman (Eds.), *The Cambridge Economic History of the United States:*

The long Nineteenth Century, 143–206. Cambridge: Cambridge University Press.

Hipp, J. and Boessen, A. (2016): The shape of mobility: Measuring the distance decay function of household mobility. *The Professional Geographer*, 1–13.

Huning, S., Bens, O., and Huettl, R. (2012): Demographic change beyond the urban-rural divide: Re-framing spatial differentiation in the context of migration flows and social networks. *Erde*, *143*(1–2), 153–172.

Hunter, L. (1998): The association between environmental risk and internal migration flows. *Population and Environment*, *19*(3), 247–277.

IOM (2005): Internal migration and development: A global perspective. In: *United Nations, Overcoming barriers – 2009*. IOM Migration Research Series. International Organization for Migration. Geneva, Switzerland.

Jacobs, J. (1961): *The death and life of great American cities*. New York, NY: Random House.

Jacobs, J. (1969): *The economy of cities*. New York, NY: Random House.

Joly, D. (2000): Some structural effects of migration on receiving and sending countries. *International Migration*, *38*(5), 25–40.

Jung, N., Lee, J., Heinemann, P., Kim, D., and Kim, H. (2004): Development of an areal elderly migration model considering spatial interaction. *Journal of Urban Planning and Development*, *130*(4), 175–183.

Kemeny, T., and Cooke, A. (2017): Urban immigrant diversity and inclusive institutions. *Economic Geography*, *93*(3), 267-291.

Kuhn, R. (2015): Internal migration: Developing countries. In *International Encyclopedia of the Social and Behavioral Sciences*, 433–442.

Kupiszewska, D. and Nowok, B. (2008): Comparability of statistics on international migration flows in the European Union. In J. Raymer and F. Willekens (Eds.), *International Migration in Europe: Data, Models and Estimates*, 41–71. Chichester: Chichester.

Lerch, M. (2014): The role of migration in the urban transition: A demonstration from Albania. *Demography*, *51*(4), 1527–1550.

Long, L., Tucker, C., and Urton, W. (1988): Migration distances: An international comparison. *Demography*, *25*(4), 633–640.

Lucas, R. (2004): Life earnings and rural-urban migration. *The Journal of Political Economy*, *112*(1), S29–S59.

Lundborg, P. and Segerström, P. (2002): The growth and welfare effects of international mass migration. *Journal of International Economics*, *56*, 177–204.

Masnick, G. (1986): *Changing patterns of interregional and interestate migrations in the United States, 1970 to 1985*. Cambridge, MA: Joint Center for Housing Studies of Massachusetts Institute of Technology and Harvard University.

Merk, F. (1978): *History of the westward movement*. New York: Alfred A. Knopf.

Mohanty, S., Mohapatra, S., Kastor, A., Singh, A. and Mahapatra, B. (2016): Does Employment-Related Migration Reduce Poverty in India? *Journal of International Migration and Integration*, *17*(3), 761–784.

Molloy, R., Smith, C. and Wozniak, A. (2011): Internal migration in the United States. *Journal of Economic Perspectives*, *25*(3), 173–196.

Moretti, E. (2012): *The new geography of jobs*. Boston, MA: Houghton Mifflin Harcourt.

Morrison, P., and Clark, W. (2011): Internal migration and employment: Macro flows and micro motives. *Environment and Planning A*, *43*(8), 1948–1964.

Mueser, P. (1989): The spatial structure of migration: an analysis of flows between states in the USA over three decades. *Regional Studies*, *23*(3), 185–200.

Newbold, K. and Bell, M. (2001): Return and onwards migration in Canada and Australia: Evidence from fixed interval data. *International Migration Review*, *35*(4), 1157–1184.

Niedomysl, T. and Fransson, U. (2014): On distance and the spatial dimension in the definition of internal migration. *Annals of the Association of American Geographers*, *104*(2), 357–372.

Niedomysl, T., Ernstson, U. and Fransson, U. (2017): The accuracy of migration distance measures. *Population, Space and Place*, *23*(1).

Niedomysl, T. (2011): How migration motives change over migration distance: Evidence on variation across socio-economic and demographic groups. *Regional Studies*, *45*(6), 843–855.

Nowok, B., van Ham, M., Findlay, A. and Gayle, V. (2013): does migration make you happy? A longitudinal study of internal migration and subjective well-being. *Environment and Planning A*, *45*(4), 986–1002.

Nunn, N., Qian, N., and Sequeira, S. (2017): Migrants and the making of America: The short and long run effects of immigration during the Age of Mass Migration (March 2017). *CEPR Discussion Paper,* No. DP11899.

Ortega, F., and Peri, G. (2009): The causes and effects of international migrations: Evidence from OECD countries 1980-2005. *NBER Working Paper Series*, No. 14833.

Ottaviano, G. and Peri, G. (2006): The economic value of cultural diversity: Evidence from US cities. *Journal of Economic Geography*, *6*, 9–44.

Özgen, C., Nijkamp, P. and Poot, J. (2011): Immigration and innovation in European regions. *IZA Discussion Papers*, Institute for the Study of Labor (IZA), No. 5676.

Partridge, J. and Furtan, H. (2008): Increasing Canada's international competitiveness: Is there a link between skilled immigrants and innovation? Paper presented at the *American Agricultural Economics Association annual meeting*, Orlando, FL.

Pendakur, R. and Young, N. (2013): Putting on the moves: Individual, household, and community-level determinants of residential mobility in Canada. *Demographic research, 29*, 767–795.

Petersen, W. (1986): *Dictionary of demography: terms, concepts and institutions*. R. Petersen (Ed.). New York: Greenwood Press.

Pierson, G. (2011): The moving American. *Yale Review*, *99*(3), 90–103.

Plane, D. (1993): Demographic influences on migration. *Regional Studies, 27*(4), 375–383.

Plane, D. and Rogerson, P. (2015): On tracking and disaggregating center points of population. *Annals of the Association of American Geographers, 105*(5), 968–986.

Poot, J. (2008): Demographic change and regional competitiveness: the effects of immigration and ageing. *International Journal of Foresight and Innovation Policy*, *4*(1/2), 129–145.

Putnam, R. (1993): *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.

Rees, P., Bell, M., Duke-Williams, O. and Blake, M. (2000): Problems and solutions in the measurement of migration intensities: Australia and Britain compared. *Population Studies, 54*(2), 207–222.

Rice, T. and J. Feldman (1997): Civic culture and democracy from Europe to America. *Journal of Politics, 59*, 1143–1172.

Rodríguez-Pose, A. and von Berlepsch, V. (2014): The legacy of mass migration on economic development in the United States. *Annals of the Association of American Geographers*, *104*(3), 628–651.

Rodríguez-Pose, A. and von Berlepsch, V. (2015): European migration, national origin and long-term economic development in the United States. *Economic Geography*, *91*(4), 393–424.

Rodríguez Vignoli, J. (2008): Spatial distribution, internal migration and development in Latin America and the Caribbean. *CEPAL Review*, *96*, 137–157.

Rogers, A., Raymer, J. and Newbold, K. (2003): Reconciling and translating migration data collected over time intervals of differing widths. *The Annals of Regional Science*, *37*(4), 581–601.

Ruggles, S., Genadek, K., Goeken, R., Grover, J. and Sobek, M. (2015): *Integrated Public Use Microdata Series: Version 6.0*. Minneapolis: University of Minnesota.

Rupasingha, A., Liu, Y. and Partridge, M. (2015): Rural bound: Determinants of metro to non-metro migration in the United States. *American Journal of Agricultural Economics*, *97*(3), 680–700.

Rosenbloom, J. and Sundstrom, W. (2004): The decline and rise of interstate migration in the United States: Evidence from the IPUMS, 1850-1990. In *Research in Economic History*, *22*, 289–325.

Saxenian, A. (2006): *The new Agronauts: Regional advantage in a global economy*. Cambridge, MA: Harvard University Press.

Shen, J. (2013): Increasing internal migration in China from 1985 to 2005: Institutional versus economic drivers. *Habitat International*, *39*, 1–7.

Skeldon, R. (2006): Interlinkages between internal and international migration and development in the Asian region. *Population, Space and Place*, *12*(1), 15–30.

Sloan, M. and Morrison, P. (2016): The post-move satisfaction of internal migrants. *Migration Studies*, *4*(1), 97–115.

Spickard, P. (2007): *Almost all aliens: Immigration, race, and colonialism in American history and identity*. New York: Routledge.

Staiger, D. and Stock, J. (1997): Instrumental Variables Regression with Weak Instruments. *Econometrica*, *65*(3), 557–586.

Stock, J. and M. Yogo (2005): Testing for weak instruments in linear IV regression. In: Donald, A. and Stock, J. (Eds.), *Identification and Inference for Econometric Models*. New York, NY: Cambridge University Press.

Tabellini, G. (2010): Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association*, *8*(4), 677–716.

Thomas, D. (1936): Internal migrations in Sweden: A note on their extensiveness as compared with net migration gain or loss. *American Journal of Sociology*, *42*(3), 345–357.

Thomas, D. (1941): *Social and economic aspects of Swedish population movements, 1750-1933*. New York: The Macmillan Company.

Thomas, D. and Kuznets, S. (1957): *Population redistribution and economic growth: United States, 1870-1950*. Philadelphia: the American Philosophical Society.

United Nations (2016): *International Migration Report 2015*. Department of Economic and Social Affairs, United Nations, New York.

U.S. Census Bureau (2016): *Annual geographical mobility rates, by type of movement: 1948-2016*. US Census Bureau, Current Population Survey. Retrieved 09.02.2017 under www.census.gov/data/tables/time-series/demo/geographic-mobility/historic.html

U.S. Bureau of the Census (1972): *U.S. Census of Population 1970*. Department of Commerce, Bureau of the Census, Government Printing Office, Washington, D.C.

White, M. and Lindstrom, D. (2005): Chapter 11: Internal Migration. In *Handbook of Population*, 311–346, Springer Science and Business Media B.V.

Yap, L. (1976): Internal migration and economic development in Brazil. *The Quarterly Journal of Economics*, 90(1), 119–137.

7.8 APPENDIX 7A

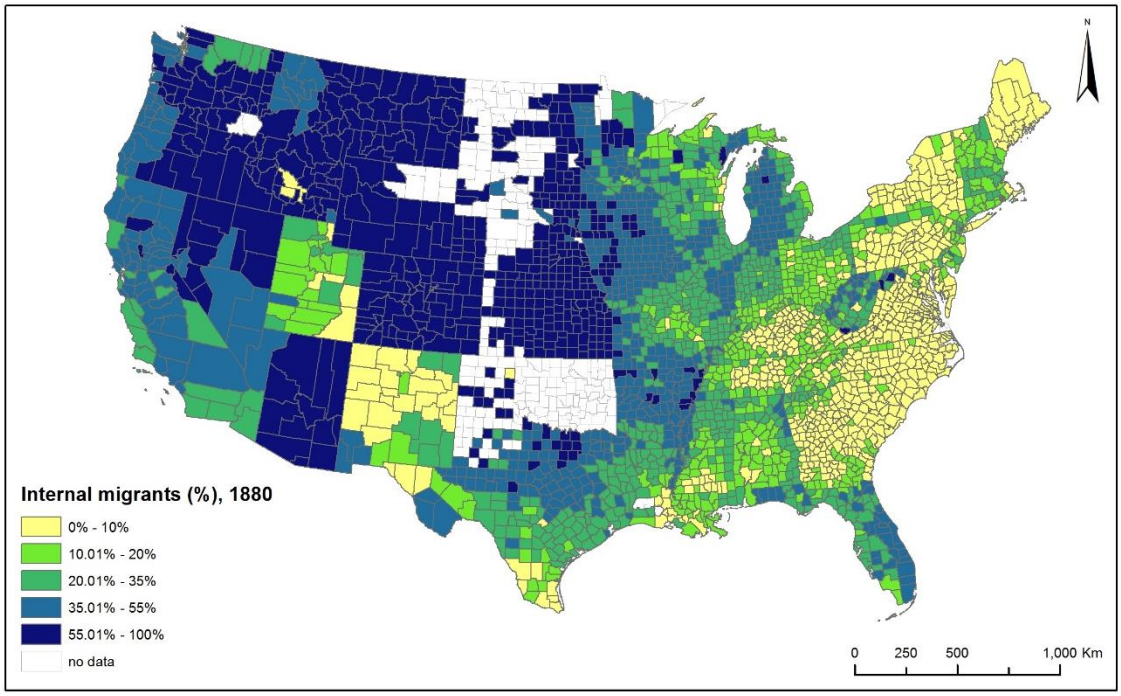


FIGURE 7-2 AMERICAN-BORN INTERNAL MIGRANTS BY COUNTY, 1880
(% OF TOTAL AMERICAN-BORN POPULATION)

Source: Ruggles, et al. 2015; own elaboration

7.9 APPENDIX 7B

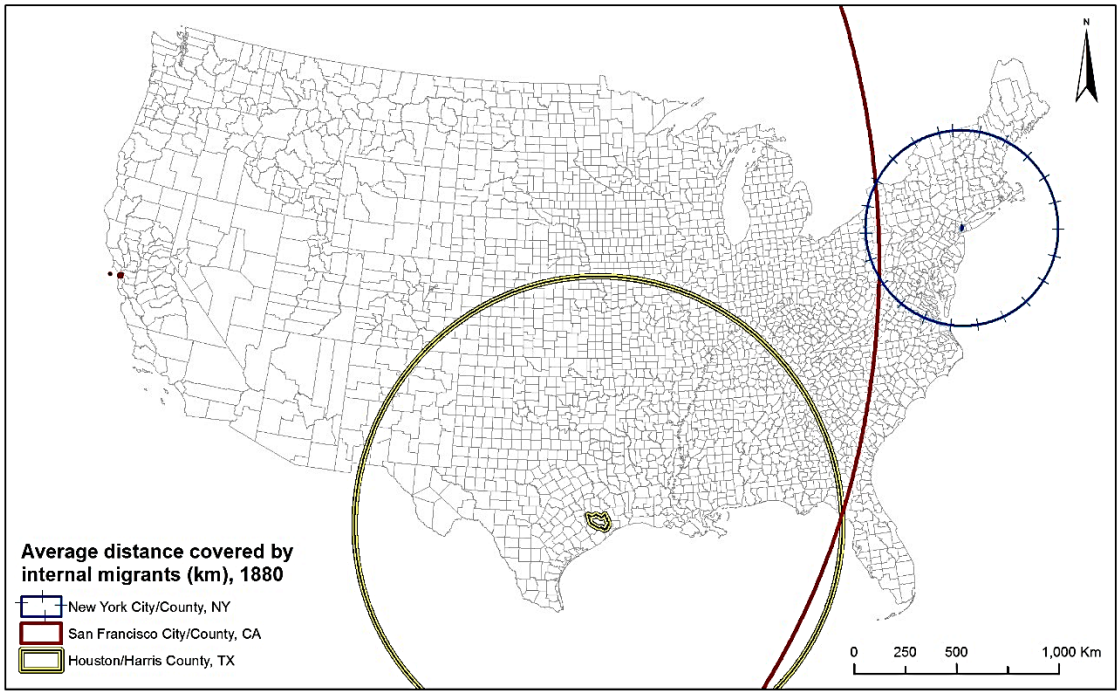


FIGURE 7-3 AVERAGE MIGRATION DISTANCE RADIUS FROM HOUSTON, NEW YORK, SAN FRANCISCO, 1880

Source: Ruggles, et al. 2015; own elaboration

7.10 APPENDIX 7C

TABLE 7-6 DESCRIPTIVE STATISTICS DEPENDENT AND MAIN INDEPENDENT VARIABLES OF INTEREST

Variable	Obs.	Mean	Std. dev.	Min	Max
Income per capita 2010	3142	34072,8	7764,75	16023	110956
All migrants 1880	2858	0,37	0,2770218	0,002884	1
Internal migrants 1880	2921	0,33	0,2653268	0	1
Population same state 1880	2921	0,67	0,2653268	0	1
Population neighbour state 1880	2921	0,10	0,1005804	0	0,8552632
Population rest of country 1880	2921	0,23	0,2465148	0	1
Av. distance int. migrants 1880 (km)	2921	898,51	669,41	0	3855,5
Av. distance all migrants 1880 (km)	2921	2.140,74	1.527,82	0	8875,2
All migrants 1910	3131	0,33	0,2560091	0,000223	1
Internal migrants 1910	3135	0,26	0,2274818	0	1
Population same state 1910	3135	0,74	0,2274818	0	1
Population neighbour state 1910	3135	0,11	0,1028442	0	0,88
Population rest of country 1910	3135	0,16	0,1847167	0	1
Av. distance int. migrants 1910 (km)	3135	780, 45	560,64	0	3362,7
Av. distance all migrants 1910 (km)	3135	2.138,44	1.710,03	0	9208,0

TABLE 7-7 DESCRIPTIVE STATISTICS CONTROL VARIABLES

Variable	Obs.	Mean	Std. dev.	Min	Max
Education 2000	3143	0,17	0,08	0,05	0,64
Population 2000	3144	89920,43	292369,90	67	9519338
Black population 2000	3120	0,09	0,145282	0	0,87
Female 2000	3144	0,50	0,019063	0,33	0,57
Female participation 2000	3142	0,55	0,065496	0,27	0,81
Unemployment 2000	3141	0,04	0,016573	0,01	0,18
Infant mortality 2000	3142	7	7,495610	0	93,8
Manufacturing 2000	3049	0,14	0,104576	0,003	1
Mean income 1880	2877	6,36	2,872315	1,75	30,27
Literacy 1880	2921	0,74	0,227872	0	1
Population 1880	3045	17799,95	39283,61	0	1206299
Black population 1880	2994	0,14	0,215203	0	0,92
Female 1880	2994	0,45	0,091157	0	0,88
Female participation 1880	2921	0,06	0,060511	0	0,39
Unemployment 1880	2918	0,16	0,295097	0	5
Manufacturing 1880	2921	0,06	0,085532	0	0,91
Mean income 1910	3128	6,82	2,132763	0	21,74
Literacy 1910	3125	0,89	0,1183161	0,333333	1
Population 1910	3138	31277,43	103795,6	0	2762522
Black population 1910	3135	0,02	0,0827582	0	0,717431
Female 1910	3135	0,47	0,0331056	0,278877	0,553718
Female participation 1910	3125	0,10	0,0721319	0	0,449153
Unemployment 1910	3128	0,06	0,0346171	0	0,484472
Manufacturing 1910	3122	0,097	0,115119	0	0,75

7.11 APPENDIX 7D

TABLE 7-8 VARIABLE DESCRIPTIONS AND SOURCES

Variable	Description	Source
Main variables of interest		
Inc per cap 2010 (ln)	Natural log of average income per capita levels in current US dollars on county level for county *i* in year *t* = 2010 (unadjusted for inflation)	US BEA
All migrants~	Share of total migrants, both foreign-born and American-born, relative to total population in county *i* in year t_0	IPUMS USA
Internal migrants~	Share of American-born internal migrants relative to total population in county *i* in year t_0 (internal migrants refer to all current residents having crossed state-lines between birth state and current state of residence)	IPUMS USA
Population same state~	Fraction of American-born population in county *i* in year t_0 relative to total American-born population with birth state equivalent to current state of residence	IPUMS USA
Population neighbour state~	Fraction of American-born population in county *i* in year t_0 relative to total American-born population with birth state equivalent to any adjacent state of current state of residence	IPUMS USA
Population rest of country~	Fraction of American-born population in county *i* in year t_0 relative to total American-born population with birth state equivalent to any non-adjacent state of current state of residence	IPUMS USA
External migrants~	Share of foreign-born migrants relative to total population in county *i* in year t_0	IPUMS USA
Distance all mig (ln)~	Natural log of average distance travelled between birth state/birth country and current county of residence by all migrants, both foreign-born and American-born, living in county *i* in year t_0	IPUMS USA
Distance int. mig (ln)~	Natural log of average distance travelled between birth state and current county of residence by American-born internal migrants living in county *i* in year t_0	IPUMS USA

Variable	Description	Source
Instruments		
Topography	Scale variable extracted from the National Atlas of the United States of America of the U.S. Geological Survey (1970) published as part of the Natural Amenities Scale corresponding to topography type of county *i* with range [1-21]	U.S. Department of Interior
Water area	Natural log of size of water areas within a county measured in square miles retrieved from the TIGER Geodatabases	U.S. Census Geography division
Control Variables		
Education	Percentage of population of county *i* with college degree in *t-k*	ICPSR
Literacy~	Literacy rate in county *i* in t_0	IPUMS USA
Income~	Constructed mean income score on county level. Individual income levels assigned to occupational data on the basis of median total income per occupation in hundreds of 1950 dollars, as natural log in t_0. Basis for construction is total size of county population in 1880/1910	IPUMS USA
Population (ln)~	Natural log of total population of county *i* in t_0 and *t-k*	ICPSR
Manufacturing~	Percentage of labour force employed in manufacturing in county *i* in t_0 and *t-k*	US BLS and IPUMS USA
Black Population~	Percentage of black population in county *i* in t_0 and *t-k*	ICPSR
Female~	Percentage of female population in county *i* in t_0 and *t-k*	Census 2000 summary files and IPUMS USA
Female Participation~	Female participation rate in the labour force in county *i* in t_0 and *t-k*	ICPSR and IPUMS USA
Unemployment~	Unemployment rate in county *i* in t_0 and *t-k* Proxy in 1880 (unemployment rate not available): Months unemployed last year	IPUMS USA and US BLS
Infant mortality~	Infant mortality rate in county *i* in *t-k* measured as number of deaths among infants aged <1 year per 1,000 live births	CDC
State Controls	State dummies	Own construction

~ refers to respective year in question: 1880/1910 for variables of interest and 1880/1910/2000 for control variable

7.12 APPENDIX 7E

Description of topography scale retrieved from The National Atlas of the United States of America

TABLE 7-9 US LAND SURFACE TOPOGRAPHY CODES

Category	Code	Land surface
Plains	1	Flat plains
	2	Smooth plains
	3	Irregular plains, slight relief
	4	Irregular plains
Tablelands	5	Tablelands, moderate relief
	6	Tablelands, considerable relief
	7	Tablelands, high relief
	8	Tablelands, very high relief
Plains with Hills or Mountains	9	Plains with hills
	10	Plains with high hills
	11	Plains with low mountains
	12	Plains with high mountains
Open Hills and Mountains	13	Open low hills
	14	Open hills
	15	Open high hills
	16	Open low mountains
Hills and Mountains	18	Hills
	19	High hills
	20	Low mountains
	21	High mountains

Source: U.S. Department of Interior, U.S. Geological Survey, Washington, DC., 1970

7.13 APPENDIX 7F

TABLE 7-10 THE LONG-TERM IMPACT OF INTERNAL MIGRATION, OLS 1910

Dep. Var.: income per capita 2010 (ln)	(1)	(2)	(3)	(4)	(5)
All migrants 1910	0.0721***				
	(0.0220)				
Internal migrants 1910		0.0375*			
		(0.0215)			
Pop. same state 1910			-0.0375*		
			(0.0215)		
Pop. neighbour state 1910				-0.0149	
				(0.0337)	
Pop. rest of country 1910					0.0768**
					(0.0327)
External migrants 1910		0.198***	0.198***	0.186***	0.205***
		(0.0494)	(0.0494)	(0.0495)	(0.0493)
Education 2000	0.0116***	0.0114***	0.0114***	0.0114***	0.0114***
	(0.000759)	(0.000749)	(0.000749)	(0.000749)	(0.000747)
Population 2000 (ln)	-0.000998	-0.000598	-0.000598	0.000321	-0.00113
	(0.00468)	(0.00468)	(0.00468)	(0.00461)	(0.00465)
Manufacturing 2000	-0.130***	-0.131***	-0.131***	-0.131***	-0.131***
	(0.0242)	(0.0241)	(0.0241)	(0.0240)	(0.0240)
Black population 2000	0.000196	0.000278	0.000278	0.000299	0.000261
	(0.000273)	(0.000273)	(0.000273)	(0.000273)	(0.000273)
Female 2000	0.0163***	0.0165***	0.0165***	0.0168***	0.0166***
	(0.00152)	(0.00152)	(0.00152)	(0.00152)	(0.00152)
Female part. 2000	0.00214***	0.00218***	0.00218***	0.00208***	0.00217***
	(0.000761)	(0.000763)	(0.000763)	(0.000764)	(0.000761)
Unemployment 2000	-0.0221***	-0.0223***	-0.0223***	-0.0222***	-0.0225***
	(0.00297)	(0.00302)	(0.00302)	(0.00303)	(0.00299)
Infant mortality 2000	8.43e-05	7.11e-05	7.11e-05	6.59e-05	7.68e-05
	(0.000419)	(0.000418)	(0.000418)	(0.000421)	(0.000418)
Income 1910 (ln)	-0.000805	7.27e-05	7.27e-05	0.00507	-0.00441
	(0.0143)	(0.0143)	(0.0143)	(0.0137)	(0.0142)
Literacy 1910	0.123***	0.162***	0.162***	0.175***	0.157***
	(0.0388)	(0.0397)	(0.0397)	(0.0396)	(0.0398)
Population 1910 (ln)	7.61e-05	-0.00291	-0.00291	-0.00325	-0.00273
	(0.00587)	(0.00579)	(0.00579)	(0.00576)	(0.00579)
Manufacturing 1910	-0.0778***	-0.0918***	-0.0918***	-0.0929***	-0.0855***
	(0.0263)	(0.0267)	(0.0267)	(0.0267)	(0.0270)
Black population 1910	0.127**	0.134**	0.134**	0.135**	0.135**
	(0.0580)	(0.0582)	(0.0582)	(0.0583)	(0.0581)
Female 1910	-0.609***	-0.513***	-0.513***	-0.570***	-0.500***
	(0.157)	(0.164)	(0.164)	(0.165)	(0.162)
Female part. 1910	0.122**	0.122**	0.122**	0.119**	0.125**
	(0.0529)	(0.0527)	(0.0527)	(0.0526)	(0.0526)
Unemployment 1910	0.0598	0.0515	0.0515	0.0560	0.0544
	(0.0838)	(0.0837)	(0.0837)	(0.0837)	(0.0835)
Observations	2,617	2,617	2,617	2,617	2,617
States	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
R-squared	0.673	0.674	0.674	0.674	0.675

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

7.14 APPENDIX 7G

TABLE 7-11 THE LONG-TERM IMPACT OF INTERNAL MIGRATION, IV 1910

Dep. Var.: income per capita 2010 (ln)	(1)	(2)	(3)	(4)	(5)
All migrants 1910	1.420***				
	(0.331)				
Internal migrants 1910		1.653***			
		(0.421)			
Pop. same state 1910			-1.653***		
			(0.421)		
Pop. neighbour state 1910				3.042***	
				(0.965)	
Pop. rest of country 1910					3.622**
					(1.428)
External migrants 1910		0.661***	0.661***	0.374***	1.003***
		(0.142)	(0.142)	(0.105)	(0.343)
Education 2000	0.0120***	0.0131***	0.0131***	0.0123***	0.0140***
	(0.000898)	(0.00107)	(0.00107)	(0.00113)	(0.00166)
Population 2000 (ln)	-0.0341***	-0.0371***	-0.0371***	-0.0136	-0.0651**
	(0.0107)	(0.0122)	(0.0122)	(0.0114)	(0.0278)
Manufacturing 2000	-0.147***	-0.136***	-0.136***	-0.136**	-0.135***
	(0.0384)	(0.0410)	(0.0410)	(0.0538)	(0.0525)
Black population 2000	-5.67e-05	-0.000573	-0.000573	0.000128	-0.00141
	(0.000484)	(0.000533)	(0.000533)	(0.000716)	(0.000867)
Female 2000	0.00910***	0.00726**	0.00726**	0.00699	0.00759
	(0.00306)	(0.00358)	(0.00358)	(0.00443)	(0.00512)
Female part. 2000	0.00596***	0.00575***	0.00575***	0.00593***	0.00553**
	(0.00148)	(0.00153)	(0.00153)	(0.00199)	(0.00219)
Unemployment 2000	-0.0260***	-0.0252***	-0.0252***	-0.0157***	-0.0364***
	(0.00465)	(0.00453)	(0.00453)	(0.00598)	(0.00950)
Infant mortality 2000	0.000211	0.000294	0.000294	5.97e-05	0.000573
	(0.000620)	(0.000671)	(0.000671)	(0.000845)	(0.00108)
Income 1910 (ln)	-0.203***	-0.211***	-0.211***	-0.0182	-0.440**
	(0.0605)	(0.0666)	(0.0666)	(0.0421)	(0.190)
Literacy 1910	-0.0662	-0.313**	-0.313**	-0.0790	-0.592*
	(0.0902)	(0.136)	(0.136)	(0.120)	(0.304)
Population 1910 (ln)	-0.00724	0.0111	0.0111	0.00241	0.0215
	(0.0111)	(0.0121)	(0.0121)	(0.0176)	(0.0205)
Manufacturing 1910	-0.0950*	-0.00892	-0.00892	-0.262***	0.292*
	(0.0509)	(0.0566)	(0.0566)	(0.0829)	(0.169)
Black population 1910	0.129	0.0901	0.0901	0.0812	0.101
	(0.0790)	(0.0848)	(0.0848)	(0.120)	(0.0950)
Female 1910	2.138***	1.587**	1.587**	0.947	2.350*
	(0.754)	(0.643)	(0.643)	(0.643)	(1.292)
Female part. 1910	0.271***	0.272***	0.272***	0.153	0.413**
	(0.0978)	(0.104)	(0.104)	(0.117)	(0.183)
Unemployment 1910	-0.142	-0.0923	-0.0923	-0.196	0.0312
	(0.165)	(0.177)	(0.177)	(0.248)	(0.291)
Observations	2,612	2,612	2,612	2,612	2,612
States	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
First stage F-stat	27.32	22.35	22.35	12.83	7.85

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

7.15 APPENDIX 7H

TABLE 7-12 THE LONG-TERM IMPACT OF AVERAGE MIGRATION DISTANCE, 1910

	OLS		IV	
Dep. Var.: income per capita 2010 (ln)	(1) All migrants	(2) Internal migrants	(3) All migrants	(4) Internal migrants
Distance all mig (ln) 1910	0.0185***		0.230***	
	(0.00360)		(0.0559)	
Distance int. mig (ln) 1910		0.0411***		0.563***
		(0.0100)		(0.135)
Education 2000	0.0116***	0.0115***	0.0103***	0.00986***
	(0.000763)	(0.000775)	(0.000991)	(0.00102)
Population 2000 (ln)	0.000480	-0.000660	-0.00289	-0.0104
	(0.00459)	(0.00484)	(0.00594)	(0.00678)
Manufacturing 2000	-0.121***	-0.124***	-0.123***	-0.0664
	(0.0237)	(0.0244)	(0.0372)	(0.0423)
Black population 2000	0.000268	6.54e-06	0.000334	-0.00156***
	(0.000273)	(0.000282)	(0.000464)	(0.000595)
Female 2000	0.0169***	0.0168***	0.0177***	0.0175***
	(0.00155)	(0.00153)	(0.00261)	(0.00222)
Female participation 2000	0.00181**	0.00210***	0.00146	0.00209**
	(0.000755)	(0.000765)	(0.00103)	(0.00102)
Unemployment 2000	-0.0213***	-0.0211***	-0.0184***	-0.0155***
	(0.00298)	(0.00312)	(0.00415)	(0.00477)
Infant mortality 2000	0.000141	0.000158	9.41e-05	0.000490
	(0.000425)	(0.000435)	(0.000604)	(0.000582)
Income 1910 (ln)	0.00532	0.0161	-0.0538**	-0.0484*
	(0.0138)	(0.0145)	(0.0263)	(0.0266)
Literacy 1910	0.149***	0.111***	0.205***	0.101
	(0.0392)	(0.0395)	(0.0681)	(0.0678)
Population 1910 (ln)	-0.00214	0.00288	-0.0210**	0.0215**
	(0.00590)	(0.00653)	(0.00952)	(0.00964)
Manufacturing 1910	-0.0805***	-0.0773***	-0.154***	0.00390
	(0.0255)	(0.0265)	(0.0426)	(0.0446)
Black population 1910	0.116**	0.126**	0.0901	0.106
	(0.0583)	(0.0585)	(0.0927)	(0.104)
Female 1910	-0.681***	-0.747***	-0.0815	-0.124
	(0.154)	(0.161)	(0.261)	(0.279)
Female participation 1910	0.118**	0.0988*	0.137	0.185**
	(0.0516)	(0.0537)	(0.0842)	(0.0864)
Unemployment 1910	0.0516	0.0611	-0.191	-0.0810
	(0.0857)	(0.0860)	(0.149)	(0.127)
Observations	2,588	2,501	2,588	2,501
States	*Yes*	*Yes*	*Yes*	*Yes*
R-squared	0.677	0.678	-	-
First stage F-stat	-	-	23.51	25.23

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

8 MIGRATION-PRONE AND MIGRATION-AVERSE PLACES: PATH DEPENDENCE IN LONG-TERM MIGRATION

8.1 INTRODUCTION

The beginning of the 21st century marked a very important shift in the population composition of the United States (US). African Americans were replaced by Hispanics (or Latinos)[39] as the largest ethnic minority. Today, Latinos make 18 percent of the US population and their economic, cultural, and political clout is growing rapidly. By 2050, Latinos are expected to represent around 30 percent of the total US population (Bergad and Klein, 2010; U.S. Census Bureau, 2016). Despite a sizable Spanish-speaking community traditionally established in what were former Mexican territories in the South West, the massive migration of Hispanics is a relatively new phenomenon in most parts of the US. It was not until the post-second World War years that Hispanics became the largest migrant group. In earlier migration waves they were only a small fraction of the incoming population. Migration to the US in the 19th and early 20th centuries was dominated by Europeans.

There is no shortage of literature aiming to understand the reasons why people migrate and the factors determining not only the volume but also the direction of migration

[39] There are intense debates regarding whether it is more appropriate to use the term Hispanic or Latino (Taylor *et al.*, 2012). However, following Sáenz and Morales (2015), the two terms are used indistinctively to refer to the population originating from Mexico, Spanish-speaking Caribbean islands, and Central and South America.

flows and the final settlement patterns. Much of this literature has revolved around socio-demographic, climatic, and cultural aspects influencing individual migration decisions (i.e. Pissarides and McMaster, 1990; Simon, 2004; Massey *et al.*, 2005; Haapanen and Ritsilä, 2007; Partridge and Rickman, 2008; Rodríguez-Pose, Ketterer and Castells-Quintana, 2015).

One migration-pull factor that has attracted considerable scientific attention is the presence of kinship and migrant networks at destination. Migrants go to where their relatives, friends, and co-nationals have previously settled. A large migrant network at destination decreases the costs of moving and raises the potential net gains. Ethnic, village, or even family ties are important attractions for migrants and lead to the establishment of distinctive migrant settlement patterns (i.e. Tassinopoulos and Kristensen, 1998; Epstein, 2002; Epstein and Gang, 2006; Radu, 2008; Jewell and Molina, 2009; Bodvarsson, Simpson and Sparber, 2014). But do migrant networks expand beyond kinship, ethnic, local, and national origins? Do migrants flock to places that have, over time, become 'migration-prone'? And, more importantly, do migrants of different national origins and generations end up in the same places that welcomed historical migration once the diverse factors that make a place attractive to different migrant generations are controlled for? These are questions that have attracted little attention in migration research, which has predominantly focused on micro-personal ties and networks rather than long-term factors that make particular places migration-prone or migration-averse.

This chapter aims to answer these questions, using a dataset comprising county level data of foreign-born population shares during two major migration waves – the Age of Mass Migration at the turn of the 20th century and the period of Latin American immigration to the US from 1950/60 onwards. The objective of this research is to

ascertain if completely unrelated migration waves follow similar geographical patterns once the socio-economic factors affecting migration are controlled for. We assume that large numbers of migrants settling in a region at a given point in time leave an imprint on the territory – a 'migration buzz'[40] – which acts as pull factor for ensuing migrant waves, irrespective of ethnicity or origin. It will be argued that the migrants' character, in particular their willingness to take risks, their entrepreneurialism, and their dynamism becomes engraved in particular territories, contributing to the creation of a long-lasting division between migration-prone and migration-averse places.

The chapter adopts the following structure: Section 8.2 summarises main facts and presents descriptive data on the two migration waves in question. Section 8.3 reviews the relevant literature before presenting a discussion of the model and data in Section 8.4. In Section 8.5 the regression results are examined, using a variety of different estimation techniques. Section 8.6 concludes and presents some preliminary policy implications.

8.2 US MIGRATION PATTERNS – A BRIEF SKETCH

The history of the US has been shaped by huge inflows of people seeking a better life for themselves and their families. Millions of people from all over the world have moved to the US, transforming the country into a 'nation of immigrants'. According to Spickard (2007: 4), "More than 99 percent of the current US population can at least theoretically trace its ancestry back to people who came […] from somewhere else".

[40] We use the word 'buzz' here in analogy to the path-breaking work by M. Storper and A. Venables (2004)

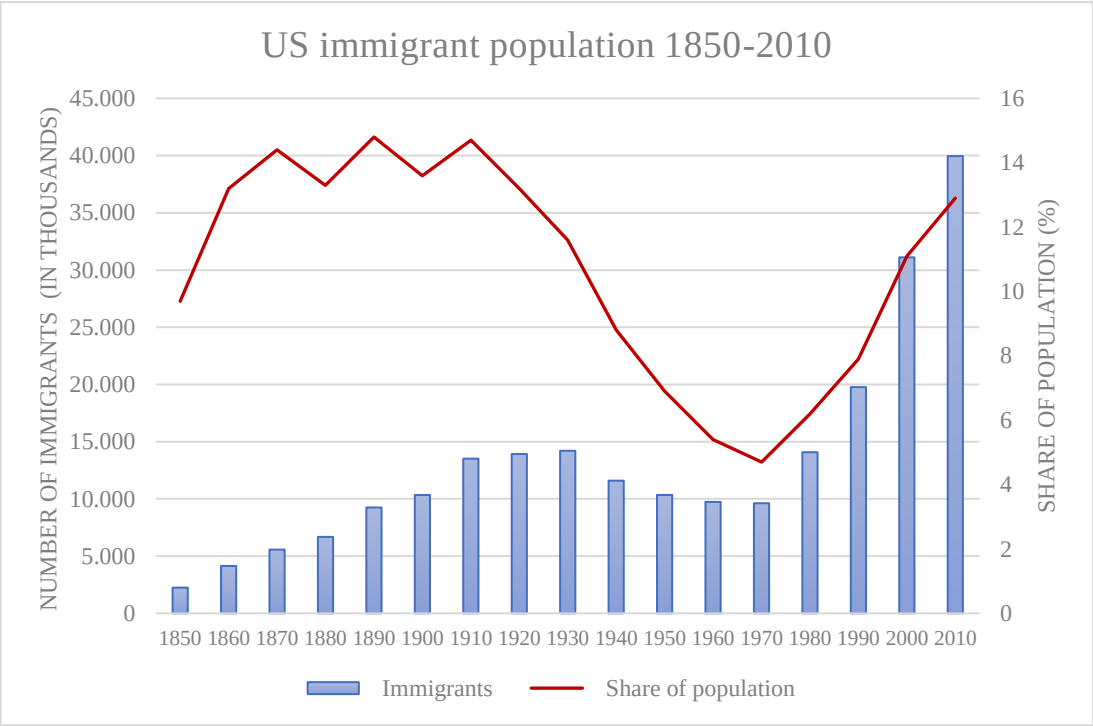


FIGURE 8-1 US MIGRANT POPULATION (ABSOLUTE AND SHARE OF TOTAL POPULATION), 1850-2010

Data source: Migration Policy Institute (2017); own elaboration

Two episodes in US immigration history stand out: the Age of Mass Migration around the turn of the 20th century and the period between 1970 and 2010. During these periods immigrant numbers rose rapidly year-on-year just as immigrant population growth rates exceeded formerly known levels. Between 1850 and 1920, the absolute number of immigrants – defined as those people having a non-US birthplace but residing in the US – increased from 2 to 14 million (Figure 8-1). Between 1970 and 2010, the foreign population in the US rose from nearly 10 million to 40 million. Such numbers represented close to 15 percent of the total population in historical and 13 percent in modern times (Figure 8-1). Both migration waves differ, however, in composition and settlement pattern. Figure 8-2 depicts the shift in major sending

regions between 1880 and 1999.

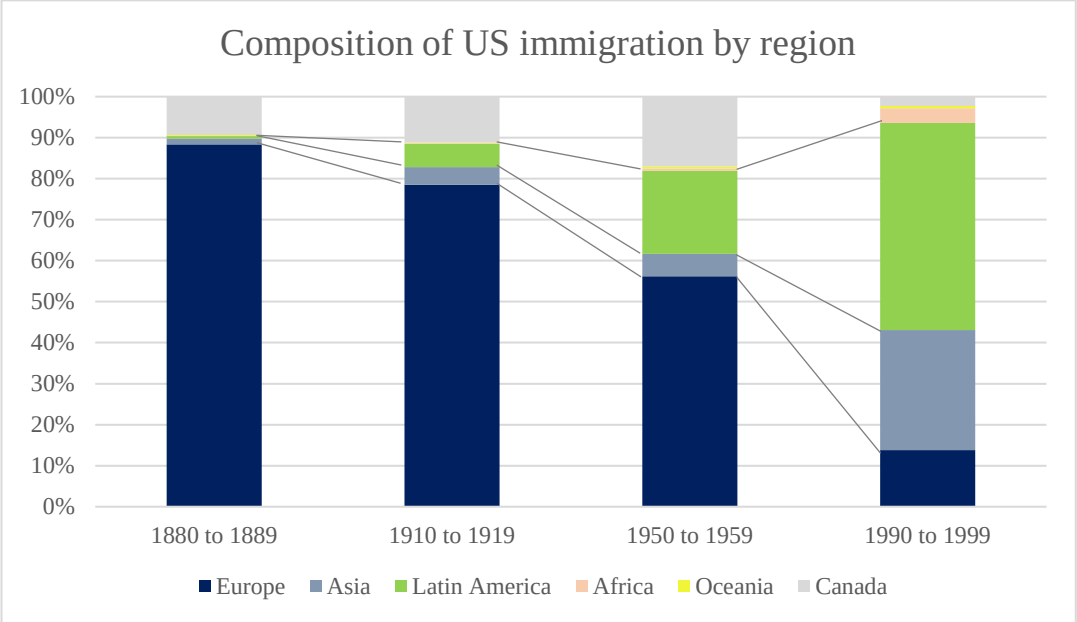


FIGURE 8-2 PERSONS OBTAINING LEGAL PERMANENT RESIDENT STATUS, 1880–1999

Data source: US Department of Homeland Security (2011); own elaboration

The first two columns of Figure 8-2 depict the US migrant population during the two peaks of the Age of Mass Migration. Between 1880 and 1889 close to 90 percent of the foreign-born population in the US was of European stock (Bertocchi and Strozzi, 2006). Early migrant contingents involved Northern and Western Europeans, primarily from England, Ireland, Germany, or the Scandinavian countries. Later migrants had Southern and Eastern European roots: Italians, Poles, Russians, Greeks, or Portuguese (Alexander, 2007; Rodríguez-Pose and von Berlepsch, 2015). Migrants tended to be male, young, single, relatively poor, and unskilled, with limited knowledge of English. After entering the US mostly via the major ports of the eastern seaboard, migrants quickly followed in the footsteps of friends and relatives, forming a distinct migrant settlement pattern and creating marked migrant communities across

the country (Bruhn, 2005). Most migrants settled in what were then sparsely populated regions of the north and west of the country. Most southern states, by contrast, were avoided by migrants (Figure 8-3).

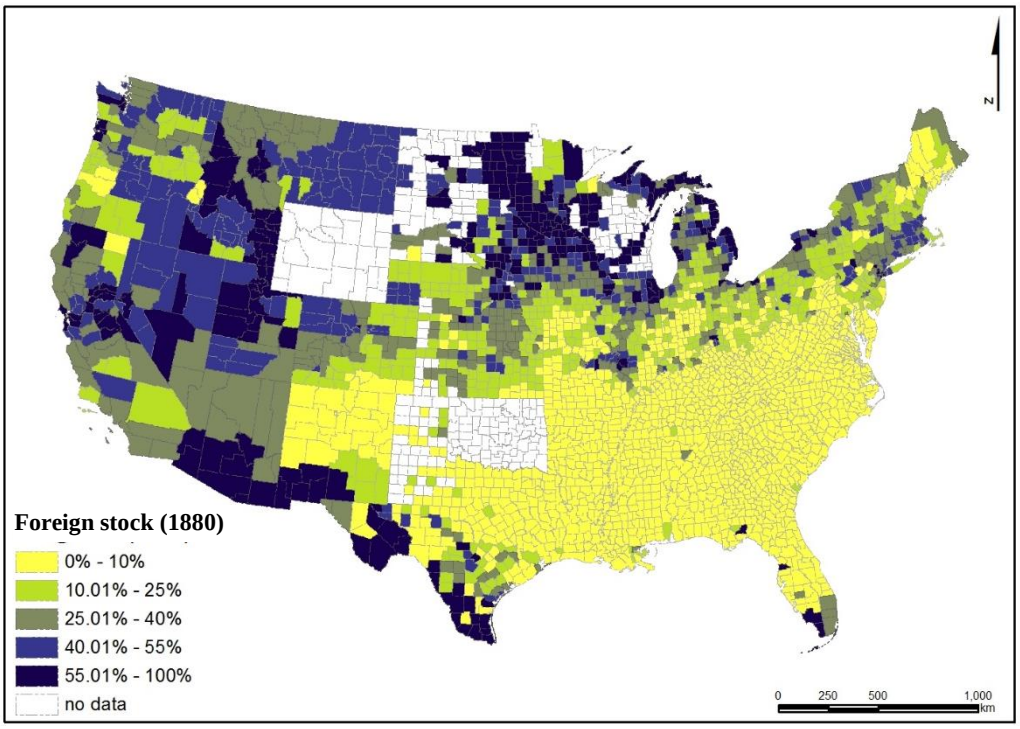


FIGURE 8-3 FOREIGN-BORN POPULATION AS SHARE OF TOTAL POPULATION, 1880

Source: Rodríguez-Pose and von Berlepsch (2014)

Columns 3 and 4 in Figure 8-2 display a clear shift in demographic makeup of the foreign-born population in the US. Over the course of the second half of the 20th century, the proportion of Europeans drastically shrank and two 'new' major sending regions emerged – Asia and, to an even greater degree, Latin America. By 2000, migrants from Mexico, the Spanish-speaking Caribbean, and Central and South America made up more than 50 percent of the foreign-born population, Asians represented around one third, while the European share had decreased to just 14 percent (Figure 8-2).

The foreign-born of the late 20th century adopted a somewhat different settlement pattern than their predecessors one hundred years earlier (see Appendix 8A). Late 20th and early 21st century migrants fundamentally resided in western states, such as Washington, Oregon, parts of Idaho, California, or Nevada and along the border with Mexico in Arizona, New Mexico, and Texas. Historically attractive states around the Great Lakes and the Mid-West appealed to limited numbers of migrants. Additional concentrations were found in Florida and the north east. The states of the 'Old South' remained, at least until 1990, relatively migration free.

These migration patterns have been mainly shaped by the largest migrant group: the Latinos (see Figure 8-4). Although migration from the 'new world' is by no means a new phenomenon, migrant flows from Latin America rapidly increased after WWII. The influx of Latinos was boosted by initiatives targeting labour shortages, such as the *bracero* program with Mexico (Daniels, 1990); by changing regulations in immigration law – replacing a system of national origin quotas by one giving preference to skills, occupations, refugee status, or facilitating family reunifications (Sáenz and Morales, 2015); and by political unrest in many Latin American countries, such as Castro's ascent to power in Cuba or the later *Mariel Crisis* (Daniels, 1990; Bergad and Klein, 2010).

TABLE 8-1 LATIN AMERICANS OBTAINING LEGAL PERMANENT RESIDENT STATUS, 1960-69, 2000-09, BY COUNTRY OF LAST RESIDENCE

Region and country of last residence	1960 to 1969		2000 to 2009	
	Total number	Share	Total number	Share
Latin America	1.241.044	100%	4.205.877	100%
Mexico	441.824	36%	1.704.166	41%
Caribbean	427.235	34%	1.053.969	25%
Cuba	202.030	16%	271.742	6%
Dominican Republic	83.552	7%	291.492	7%
Haiti	28.992	2%	203.827	5%
Jamaica	62.218	5%	172.523	4%
Other Caribbean	50.443	4%	114.385	3%
Central America	98.560	8%	591.130	14%
Belize	4.185	0%	9.682	0%
Costa Rica	17.975	1%	21.571	1%
El Salvador	14.405	1%	251.237	6%
Guatemala	14.357	1%	156.992	4%
Honduras	15.078	1%	63.513	2%
Nicaragua	10.383	1%	70.015	2%
Panama [1]	22.177	2%	18.120	0%
Other Central America	-	-	-	-
South America	250.754	20%	856.593	20%
Argentina	49.384	4%	47.955	1%
Bolivia	6.205	0%	21.921	1%
Brazil	29.238	2%	115.404	3%
Chile	12.384	1%	19.792	0%
Colombia	68.371	6%	236.570	6%
Ecuador	34.107	3%	107.977	3%
Guyana	4.546	0%	70.373	2%
Paraguay	1.249	0%	4.623	0%
Peru	19.783	2%	137.614	3%
Suriname	612	0%	2.363	0%
Uruguay	4.089	0%	9.827	0%
Venezuela	20.758	2%	82.087	2%
Other South America	28	0%	87	0%
Other America	22.671	2%	19	0%

Source: own elaboration using data from the US Department of Homeland Security (2011)

While in the early 1960s, the incoming Latino immigrants were overwhelmingly from Mexico, with relatively large Cuban and Puerto Rican[41] contingents, by 2000, the Latino community grew massively in diversity (see Table 8-1).[42] Caribbeans – both Spanish- as well as French- and English-speaking migrants – and Central and South

[41] The Puerto Ricans are not shown in Table 8-1, as they are considered American-born since the Jones Act in 1917 (Bergad and Klein, 2010).

[42] Table 8-1 only captures the official extent of legal immigration. Undocumented migrants to the US are not included in these numbers, even though they became increasingly relevant from the 1960s onwards.

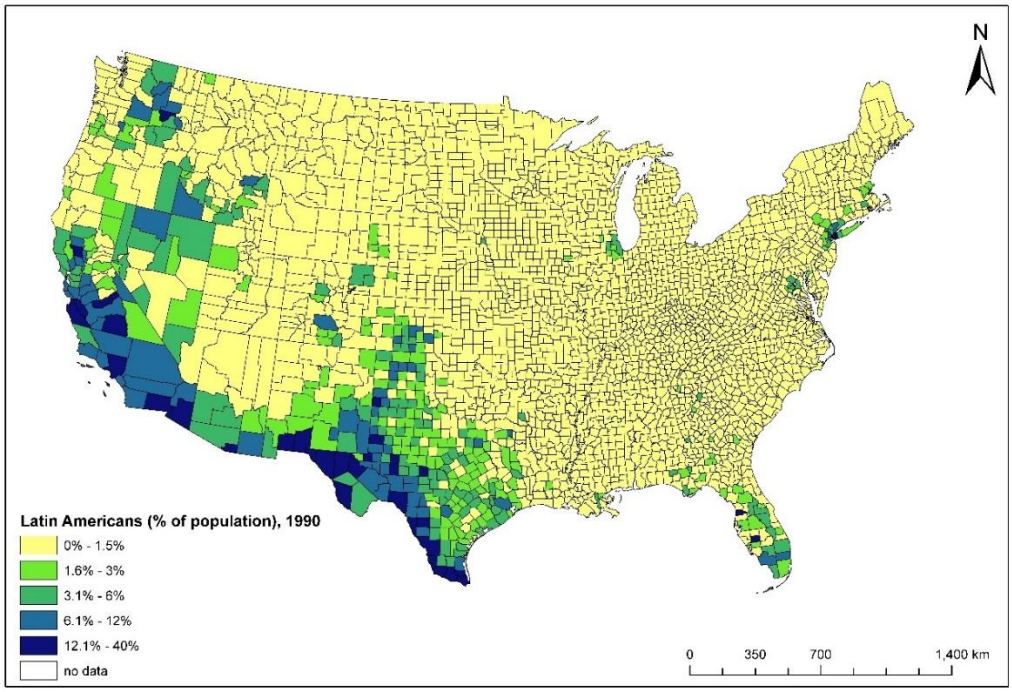
Americans – mainly from El Salvador, Honduras, Guatemala, Nicaragua, Colombia, Peru, and Ecuador – accounted for almost 40 percent of immigration from the 'new world'. Despite the differences in countries of origin and in traditions, backgrounds, and customs, the late 20th century Latin American migrant did not greatly differ from the late 19th century European migrant: young, poor, unskilled, pious, and from rural background. There was, however, a more balanced gender-ratio compared to their European precursors. Women immigrants, especially from South and Central America, were strongly represented in this migration wave (Bergad and Klein, 2010).

Once in the US, Latino immigrants generally occupied the lowest levels of the social and economic scale. They tended to work in low-paying jobs within the "Latino immigrant occupational niche" (Sáenz and Morales, 2015: 109), such as "agriculture; meat, poultry, and seafood processing; construction; waiters/waitresses; cooks; maids and housekeeping cleaners; and janitors and building cleaners […]" (Douglas and Sáenz, 2008: 169). Although still overrepresented in agriculture, by 1990 over 90 percent of Latinos lived in metropolitan areas. New York, Los Angeles, and Miami are the hubs of Latino culture in the US (Cafferty and Engstrom, 2000). Hispanic immigrants and their children have the highest high school dropout and the lowest tertiary education rates in the US population and, as a consequence, cluster at the lower end of the median annual income scale (Daniels, 1990).

Prior to 2000, Latinos were concentrated in three states: Florida, California, and Texas. New York, New Mexico, Arizona, and parts of Nevada and Washington also had sizeable Latino populations (Figure 8-4a). From the 2000s onwards and following processes of industrial restructuring, variations in the geography of labour demand, and changes to immigration legislation and border policies, Latino immigrants fanned out beyond their traditional destinations (Donato *et al.*, 2008). As depicted in Figure

8-4b, states such as Colorado, Utah, and Wyoming began to receive large Hispanic contingents. Areas without a history of immigration of any sort, such as North Carolina, Virginia, or Georgia, as well as some states in the Midwest lured growing numbers of Latino migrants, (i.e. Bailey, 2005; Bump, 2005; Hansen, 2005). Figure 8-4 displays the geographic dispersion of Hispanic immigrants, which transformed Latin American population flows "from a regional to a national phenomenon" (Massey and Capoferro, 2008: 47).

a)



b)

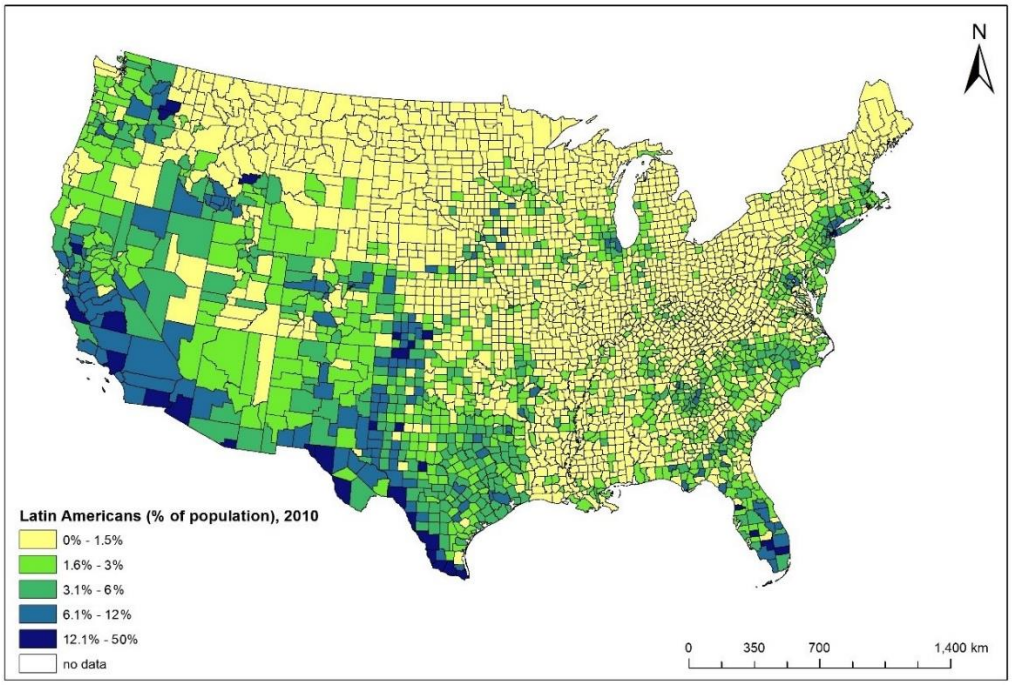


FIGURE 8-4 LATINO FOREIGN-BORN POPULATION AS SHARE OF TOTAL POPULATION, 1990 AND 2010[43]

Data source: Minnesota Population Center (2016); own elaboration

[43] As the relative size of the foreign-born Latino county population in 1990 or 2010 is of a different magnitude relative to that of the foreign-born population during the Age of Mass Migration, the map intervals are different from those in Figure 8-3.

8.3 WHY DO MIGRANTS END UP IN PARTICULAR PLACES AND NOT IN OTHERS?

Many studies have aimed to understand why people migrate and what determines the volume of migration and settlement patterns. Traditional theory, such as the neoclassical economic framework, regards migration as a logical consequence of factor price differentials across geographic units generated by varying endowments in the supply and demand of labour. Assuming perfect competition, complete information, and prefect mobility of capital and labour, differences in factor prices represent the fundamental trigger for population mobility. Migrants thus move from low- to high-wage regions (i.e. Hicks, 1932; Ranis and Fei, 1961; Harris and Todaro, 1970; Ödberg, 1997). To the individual, the decision to migrate boils down to a simple income maximisation problem in which a rational individual moves if the calculation of relocation costs and benefits – mostly regarded solely in financial terms – generates a positive net return. Where a migrant decides to settle hinges on the maximisation of individual productivity, i.e. where the highest expected net income level can be attained relative to the acquired skill set and expected time horizon (Sjaastad, 1962; Todaro, 1969; Borjas, 1990). Migration-prone regions are, therefore, according to this theory, high productivity areas offering high wages (Bauer and Zimmermann, 1997; Greenwood, 1997; Ödberg, 1997).

Many of the assumptions of traditional migration theories have, however, been challenged (e.g. Roy, 1951; Borjas, 1987; Borjas et al. 1992) . Greater attention has been paid to, among many others, how labour markets, globalisation, history, households, quality of life, absolute vs. relative income positioning and social capital influence the migration decision (see overviews in i.e. Massey *et al.*, 1993, 2005;

Bodvarsson, Simpson and Sparber, 2014). Two strands have become particularly prominent – one analysing the regional characteristics behind a territory's attractiveness to potential immigrants and another evaluating individual factors influencing the probability to migrate. A combination of both goes far in explaining not only the magnitude of migration but also the composition, direction, and final settlement pattern (Massey *et al.*, 1993).[44]

Key push and pull factors influencing population movements have been identified. On top of high wages, job availability, the possibility of job progress, and low unemployment and long-term unemployment ratios (Pissarides and McMaster, 1990; Ritsilä and Ovaskainen, 2001; Faggian and McCann, 2008; Biagi, Faggian and McCann, 2011), the size and composition of the economy are also a draw for immigrants (Piore, 1979; Partridge and Rickman, 1996; Simon, 2004). Regional market potential linked to agglomeration patterns can increase a place's attractiveness (Ottaviano and Puga, 1998). Other relevant push factors are quasi-financial forms of income, such as social welfare spending and public amenities, as well as re-distributional transfer mechanisms, providing insurance against income losses, and the availability of public goods and regional institutional quality (Day, 1992; Haapanen and Ritsilä, 2007; Ketterer and Rodríguez-Pose, 2015). In short, economically dynamic regions with high levels of regional development, adequate financial incentives, and widely available employment opportunities generally act as magnets for immigrants.

Beyond the purely economic factors, a raft of social aspects also captured the attention of researchers. A good education system and favourable human capital endowments

[44] As this chapter deals with regional aspects, the focus in this theoretical section will predominantly lie on the characteristics of the region rather than on individual ones.

at destination have been shown to facilitate increases in migrant productivity as well as easing transitions into the job market. Highly educated individuals are also more likely to migrate (Greenwood, 1997; i.e. Chiquiar and Hanson, 2005; Zimmermann, 2005), as "higher education reduces the risks of migration through a higher ability to collect and process information" (Zimmermann, 2005: 429). The demographic composition of the population in the receiving region further shapes migration flows. A predominantly young population, for example, is more likely to lead to larger population out-flows, once labour market conditions tighten (Cairns and Menz, 2009). Expected gains in lifetime income are significantly lower for older people than for the young, making the latter group more mobile (i.e. Bowles, 1970; Burda and Wyplosz, 1992; Zimmermann, 2005). Urban and natural amenities, such as a region's cultural life, its history, climate and natural environment, and quality of life, have also attracted the attention of migration research. Boosting a region's amenities is increasingly regarded as an important pull factor for incoming skills and talent (Ferguson *et al.*, 2007; Rappaport, 2007; Partridge, 2010; Rodríguez-Pose and Ketterer, 2012).

Lastly, the territory's existing stock of previous migrants within the same 'reference or peer group' of incoming population – those linked by kinship, ethnicity, or friendship who arrived earlier and established themselves in the receiving territory – plays a central part in the size and the origin of migrants. Research in economics, sociology, history, and political science has placed increasing emphasis on community ties stretching from place of origin to host region as a fundamental determinant of individual migration decisions and overall level of migration (i.e. Tassinopoulos and Kristensen, 1998; Epstein, 2002; Epstein and Gang, 2006; McGovern, 2007; Radu, 2008; Jewell and Molina, 2009; Bodvarsson, Simpson and Sparber, 2014). Migration decisions therefore depend not only on the individual's own actions and

characteristics, but also on the specific migration choices of the individual's peer group. A large community of friends and family at destination (a so-called kinship network) reduces relocation costs and increases the expected return associated with migration. Having relatives in a foreign destination generates a self-perpetuating element as, "each act of migration creates additional social ties for future migrants, who in turn extend the range of social capital for further migrants" (McGovern, 2007: 220). Similarly, large and well established groups of earlier migrants, sharing a common origin (known as a migrant network), generates an equally positive externality (Winters, De Janvry and Sadoulet, 2001; Bodvarsson, Simpson and Sparber, 2014). Once the kinship and migrant network in a territory reaches a critical mass, it affects new migrant settlement patterns by attracting more and more population from the shared region of origin. The result is a geographical clustering of immigrants from specific local, national, or ethnic backgrounds (Bauer, Epstein and Gang, 2007, 2009; Jaeger, 2007).

Access to these networks lowers the psychological and information costs for the arriving individual and hence significantly affects the volume of migrants and their choice of destination (i.e. Yap, 1977; Hugo, 1981; Massey and España, 1987; Radu, 2008). A large stock of migrants from the same origin generates social capital which reduces employment and housing search costs, lowers language barriers, offers protection from crime and income loss, provides temporary credit and lodging, and eases the individual's settlement process. It therefore facilitates integration into what is initially an alien environment, administration, culture, and society (Marks, 1989; Daniels, 1990; Carrington, Detragiache and Vishwanath, 1996; Massey *et al.*, 2005; Bodvarsson, Simpson and Sparber, 2014). This phenomenon implies that total moving costs are endogenous to the volume of previous migrants. "Once started, migration

develops momentum, as current migration reduces the cost of future migration […] [it] continues or even accelerates […] [and] is channelled in that migratory paths emerge" (Carrington, Detragiache and Vishwanath, 1996: 910). Many examples can be found to support these findings. Massey *et al.* (1987), for example, reported that nearly 40 percent of Mexicans found employment in the US via friends or relatives. Munshi (2003) calculated that established community ties led to significantly higher employment rates and income levels among Mexican immigrants in the US. 'Migration clubs' managed by kith and kin among black southerners in the north, as well as soccer clubs involving Mexican immigrants in Los Angeles, provided the linguistic and cultural cushion for the new arrivals (Massey *et al.*, 1987; Carrington, Detragiache and Vishwanath, 1996). Similarly, European migrants nearly a century earlier sent letters to family and friends with information on employment and the housing market in the US, often asking to join them and attaching money to pay for transatlantic trips (Daniels, 1990; Joly, 2000).

Migrant as well as kinship networks can trigger path dependence or 'herd behaviour' (Bikhchandani, Hirshleifer and Welch, 1998; Epstein, 2002) between place of origin and destination, thus affecting the volume and direction of migration flows. "Once the number of network connections in an origin area reaches a critical threshold, migration becomes self-perpetuating because each act of migration itself creates the social structure needed to sustain it" (Massey *et al.*, 1993: 449). In short, a large, pre-established migrant network impacts self-selection, decreases the costs of moving, raises the potential net gain for the would-be migrant, stimulates mobility, and steers migration flows into migration-prone regions, following ethnic, village, or even family ties (Shah and Menon, 1999; Clark, Hatton and Williamson, 2007; McKenzie and Rapoport, 2010; Simpson and Sparber, 2013).

One factor of particular importance in the migration network literature is the common birthplace. A shared geographical origin, ethnic bond, or common background is crucial for migrants to reap the benefits of networks and in determining a region's appeal. The literature assumes that a common origin, a shared community-belonging, or family ties are the sources for the establishment of migrant networks and the development of path dependence. Membership of a shared community therefore acts as a pull factor for incoming migrants, determining the final migrant settlement pattern. However, one important question has not been explored: what if the migrant networks are not formed solely on the basis of a common ethnicity or birthplace, but simply on the presence of previous migration groups, regardless of origin? What if a large community of migrants in a given place creates a favourable environment for migrants which endures over time? Could the institutional setting established by previous generations of migrants transform a locality or region into a migration-prone area, welcoming to migrants for generations, regardless of origin? What if past migration begets future migration? In their work on the long-term impact of migration and its implications for regional development in the US, Rodríguez-Pose and von Berlepsch (2014, 2015) assume that the character of the migrant – more risk-seeking, entrepreneurial, and dynamic – becomes engraved in the territories where migrants settle. They posit that the presence of large groups of immigrants can generate a migration vibe in the receiving areas which transforms the institutional framework in ways that sends signals to future migrant generations.

There is, however, limited empirical evidence to support or refute the claim that past migration patterns transform territories into migration-prone or migration-averse areas over long timeframes. This is precisely what this chapter aims to demonstrate: Comparing the settlement patterns of two migration waves of a very different nature

and separated by almost one century, this chapter analyses whether past migration waves in the US are determining the settlement pattern of current migration waves, irrespective of origin and ethnicity of migrant groups.

8.4 MODEL AND DATA

8.4.1 ECONOMETRIC SPECIFICATION

The model estimated to assess whether past migration waves shape the settlement pattern of completely unrelated later waves of migrants takes the following form:

$$Latinos_{i,t} = \alpha + \beta\, Migration_{i,t_0} + \partial\, X_{i,t-k} + \theta\, Z_{i,t_0} + \mu\, distMX_i + \vartheta\, state + \varepsilon_{it}$$

where *Latinos* is the share of total population born in a Latin American country[45] in county i in period t (t=1970, 1980….2010); *Migration*, the variable of interest, represents the share of foreign-born population in county i at t_0 (t_0=1880 or 1910); X is a vector of economic and socio-demographic characteristics of county i which are assumed to serve as pull factors to recent migrants (at time t-k; k=10 years before the migrant account is taken) and Z corresponds to a vector at t_0 which includes the same county characteristics as vector X that would have determined the appeal of the county to migrants more than a century ago. *DistMX* stands for the distance of any given county i to the Mexican-American border, while *state* represents state-specific fixed-

[45] The focus of the analysis lies only on the foreign-born Hispanic population. Other indicators recorded by the census, such as a Spanish last name, the usage of Spanish at home, or Hispanic origin were not consistently included in the questionnaires across the years and are therefore discarded.

effects controlling for potential spatial correlation between counties within a given state. ε describes robust standard errors.

The model is run consecutively per decade between 1970 and 2010 for the 3,109 continental counties of the US in 2010, covering a 40-year timeframe and shifting the dependent variable in each regression by 10 years. As any analysis involving economic and migration data is prone to endogeneity issues, introducing the dependent (*Latinos*) and explanatory variables (vector *X*) within the same time structure is highly problematic (Treyz *et al.*, 1993; Özgen, Nijkamp and Poot, 2011). Hence, all explanatory variables determining current migration flows are lagged by 10 years. This way, the risk of reverse causality is mitigated with, on the one hand, migration flows impacting regional economic features and, on the other, regional characteristics simultaneously directing migratory settlement patterns.

Further endogeneity issues connected to omitted variable bias or spatial sorting are treated by means of three different robustness checks. First, the consecutive estimations are rerun, shifting the base migration year by 30 years, from 1880 to 1910. Both 1880 and 1910 depict peak years within the period of mass migration to the US: 1880 represents the peak of the first wave of migrants, while 1910 that of the second wave. Two different migrant compositions are thus considered. As the main source of migration shifted around the turn of the century, the 1880 specification comprises mostly northern and western Europeans, predominantly from the British Isles, Germany, or Scandinavia. Using 1910 as base year, an entirely different composition of the foreign-born population is represented: mostly southern and eastern Europeans from Italy, Poland, the former Soviet Union, Portugal, or Greece (Hatton and Williamson, 1998). The shift in the composition of migrant groups between 1880 and

1910 minimises the risk of the results being driven by omitted variables connected to the character of the specific type of migrant.

Second, an instrumental variable estimation is used in order to extract the underlying effect of historical on current migration waves. By capturing the exogenous variation in the migration variable of 1880, the potential endogeneity bias in the least-squares estimates is reduced. The discussion of the instruments used in the analysis is presented in section 8.4.3.

Last, a third robustness check consists of reshaping the dataset into a quasi-panel structure. As the use of traditional FE-models is impossible, given that the analysis is built around a time-invariant variable of interest (*Migration*) and the probability of a high correlation between some of the time-varying variables and region-specific fixed-effects, an alternative Hausman and Taylor (HT) (1981) instrumental variable estimator is used. This estimator allows for the inclusion of time-invariant variables such as *Z* or *distMX* in a panel setting using the within transformation of time-varying variables to estimate consistent coefficients. It also accepts the potentiality of a correlation between some of the independent variables and individual specific effects. Individual means of the uncorrelated regressors, on the other hand, are employed to instrument for the endogenous variables (Baltagi, Bresson and Pirotte, 2003; Baltagi, 2008). Vector *X* contains the endogenous variables in our model.

8.4.2 DATA

The migration data employed to construct the independent variable of interest *Migration* was extracted from the Integrated Public Use Microdata Series USA database (IPUMS) Version 6.0 (Ruggles *et al.*, 2015). This database provides US microdata covering censuses and American Community Surveys between 1790 and 2010.[46] Using a weighted sample of the US population including 5,791,531 individuals (11 percent of the total population in 1880) and 923,153 individuals (1 percent of the population in 1910), individual data points are matched to the respective county of residence at the time of the census and aggregated at county level. Following changes in county numbers, size, and boundaries between 1880 and 2010, cartographic boundary files of the 48 continental states (excluding Alaska and Hawaii) provided by the US Census are used for every decade of analysis. In order to normalise borders to 2010 county boundaries, averages weighted by population density at the time of the boundary change were calculated for each individual county in historical years. With the exception of Oklahoma in 1880,[47] this method allowed to match historical county features to their 2010 equivalent.[48]

The dependent variable, *Latinos*, was constructed using the National Historical Geographic Information System (NHGIS) Version 11.0 (Minnesota Population Center, 2016), which consists of data on US geographic units covering the population, housing, agriculture, and the economy from 1790 until today. The data on the Latino population at county level was extracted for the years between 1970 and 2010. As not

[46] The American Community Survey was only initiated in 2005.
[47] Oklahoma only became an organized territory in 1890.
[48] The US territory comprised 2,875 counties or equivalent territorial units in the 48 contiguous states in 1880; 3,123 in 1910; and 3,109 in 2010.

all Latin American countries are listed independently within the NHGIS database, the share of Latin American population refers to Cubans, Mexicans, and all other people born in the Americas, excluding the US and Canada.

Three different data sources were used to construct the county level database of control variables included in vector *X*. All county data for the years 1960 to 2000 was extracted from the US Bureau of Economic Analysis (BEA) database, the Social Explorer data collection, and the Inter-University Consortium for Political and Social Research (ICPSR) historical databases. Vector *X* is included in the model to control for recent county level economic and socio-demographic characteristics, covering the relevant pull factors for migrant volume and settlement discussed in the literature. These are assumed to directly affect the settlement pattern of Latino migrants and the share of Latinos in the population of any given US county. Control factors include income per capita (as natural log), the share of population living in urban areas, the education level measured as the share of adults with tertiary education, the unemployment rate, the share of employment in agriculture as proxy for industry structure, female participation rates in the labour force, and the share of the black and female populations. All income variables were adjusted for inflation and converted to 2010 dollars using the CPI inflation converter of the Bureau of Labor Statistics (BLS).

The second vector of control variables, vector *Z*, dates from the time of the historical migration: 1880 and 1910 respectively. Vector *Z* comprises the historical equivalent of all variables included in vector *X*. It is used to control for historical economic and socio-demographic county features which may have influenced not only economic development levels but also the settlement pattern of migrants in 1880 and 1910 respectively. These historic variables are either extracted from the ICPSR database or constructed using the IPUMS USA weighted population samples. In the latter case,

the same method as with the independent variable of interest is repeated, aggregating all individual data at county level. As income per capita data were not collected at the turn of the century, income proxies for 1880 and 1910 are constructed using the median total income per occupation category in 1950 dollars. These occupational income equivalents were allocated to an individual's occupation in 1880 and 1910 and then aggregated at county level using the 1880/1910 total size of the county population as basis.

The distance to Mexico (as natural log), included as further control, was constructed using GIS software, calculating the point-distance matrix between a county's centroid and the nearest point on the continental border with Mexico. An exact description of all variables, including sources, is presented in Appendix 8B.

8.4.3 INSTRUMENTATION STRATEGY

Potential endogeneity bias in the least-squares estimates is tackled by means of instrumental variable (IV) estimations. Two instruments are used to assess the direction of causation and certify the exogenous variation of the variable of interest, *Migration*. The first is population density in 1880, extracted from the ICPSR database. The second is a measure of the endowment in public services, proxied by share of employment in the health and education sector as well as in public administration in any given county *i* in 1880 (or 1910). Data for the endowment in public services stem from the IPUMS USA individual data points, aggregated at the county level.

The justification for the choice of both instruments is related to their role in determining migrant settlement patterns during the Age of Mass migration. Density

played a role as the newly arrived generally moved to places where their predecessors had already settled and where basic forms of infrastructure were already established. Public and often cultural organisations, such as churches, schools, newspapers, and meeting houses, became especially important in the settlement process, as they offered channels for information and knowledge exchange. German meeting houses, Swedish schools, or Irish churches facilitated settling down in previously unknown places (Daniels, 1990). Moreover, migrants valued access to markets and nearby settlements for basic supplies as well as the presence of community networks, schooling for their children, and the possibility of sustaining connections to their home countries. The presence of a post office was crucial in maintaining ties both with family and friends in the country of origin and with other migrant settlements across the US. Finally, high-density and an efficient public service increased security and reduced the perception of danger. Hence, population density and public services influenced migrant settlement decisions. As population density patterns and employment features drastically changed between 1880 (1910) and 2010, the exogeneity condition separating out the uncorrelated component of our endogenous variable, *Migration*, in the first stage regression is fulfilled. There is no significant correlation between 1880 population density and public service endowments and the location pattern of Latinos more than 90 years later. This reinforces the exogeneity of the chosen instruments.

From an econometric perspective, the essential relevance property for IV analysis is satisfied as the combination of both instruments retrieves sufficiently large first-stage F-statistics based on the Staiger and Stock (1997) test for weak instruments in combination with the Stock and Yogo (2005) critical values. A further validation test of the quality of our chosen instruments is undertaken when testing for overidentifying

restrictions employing the Anderson-Rubin test. This partial test of instrument exogeneity confirms, yet again, the quality of the instruments.

8.5 ANALYSIS

Table 8-2 reports the results of estimating the model for the settlement pattern of Latinos in the US between 1970 and 2010, employing a cross-sectional structure. Specifications (1) through (5) present the coefficients and significance levels, shifting the dependent variable (share of Latinos in a given county *i*) by 10 years in each regression.

In line with our assumption, historical migration is positively associated with the dependent variable, *Latinos*, across all five decades of analysis ranging from 1970 to 2010. A large share of migrants in 1880 is strongly and positively connected to the share of foreign-born Latino migrants in any given US county 90 to 130 years later. Hence, once other factors are controlled for, counties that attracted large inflows of European settlers at the end of the 19th century remain significantly more appealing to incoming Latino migrants over a century later, than those which were largely bypassed by European migration. Despite differences in time periods, backgrounds, and cultures, both migration waves are connected by place of settlement at significance levels of 5% and 1% respectively. Hence, the results of the analysis uncover a strong territorial dependency in US migration over the last century (Table 8-2).

The coefficients of the lagged recent (*t*-10) control variables indicate that, for much of the late 20th century, Latino migration to the US did not follow traditional pull factors. Prior to 2000, economic factors mattered little for Latino migration. Latinos generally settled in areas of the US not particularly well off in terms of income or education

levels. These pull factors do not display any significant relation to the dependent variable of interest. Even more surprisingly, local unemployment levels were positively associated with the share of Latino immigrants in a county. Hispanic settlement areas had, on average, higher levels of unemployment than the rest of the country. Socio-demographic factors, such as ethnic and gender composition, also display non-significant coefficients. Before 2000, Latinos mainly settled in two types of counties: either in highly urbanised areas in and around Los Angeles, the Bay Area, El Paso, San Antonio, Miami, or Houston, or in rural areas with large employment in agriculture. A large participation of women in the labour force also lured Latino migrants. These findings are in line with the relevant literature emphasising the appeal of the cities with largest Latino concentrations, the rising contribution of Latinos to agricultural employment, and the gender balance in migration flows (Sáenz and Morales, 2015).

In more recent decades (2000 and 2010), Latino migrant settlement increasingly responds to more traditional factors. Regional income levels are, in line with the relevant migration literature, positively and significantly associated with Latino population shares. This implies that the location of Latinos in the US has become more geographically diversified in recent decades (see Figure 8-4), moving towards more prosperous areas without eroding the relevance of distance to the Mexican border as a key marker for the location of Latino immigrants (Table 8-2).

As could be expected, most 1880 controls are completely irrelevant to Latino settlement patterns 90 to 130 years later. There are two exceptions: regional income levels in 1880 are positively correlated to present day shares of Latino population, albeit with fluctuating significance levels. A higher female participation in the labour

force more than a century ago is, by contrast, associated with lower levels of Latino presence today.

As the OLS results presented in Table 8-2 could be prone to potential endogeneity issues caused by omitted variable bias, spatial sorting, or reversed causality, a range of robustness checks is conducted. These imply, first, shifting the base year, second, resorting to IV estimation techniques and, last, using the dataset as a quasi-panel structure and estimating the model by means of a HT-estimator.

TABLE 8-2 THE IMPACT OF HISTORICAL MIGRATION ON LATINO SETTLEMENT PATTERNS, OLS 1880

Dep. Var: % of Latinos in county	(1) 1970	(2) 1980	(3) 1990	(4) 2000	(5) 2010
Migration 1880	0.0109***	0.0165***	0.0198***	0.0213**	0.0146**
	(0.00315)	(0.00524)	(0.00726)	(0.00944)	(0.00696)
Income pc (ln)~	0.000664	0.00418	0.00272	0.0409***	0.0462***
	(0.00134)	(0.00273)	(0.00374)	(0.00807)	(0.00487)
Urban share~	0.00411***	0.00658***	0.0118***	0.0298***	0.0449***
	(0.00128)	(0.00179)	(0.00231)	(0.00338)	(0.00272)
Education~	-0.00294	-0.00263	-0.00629	-0.0604***	-0.122***
	(0.00282)	(0.00571)	(0.00726)	(0.0112)	(0.00983)
Unemployment~	0.0293**	0.0356**	0.0102	0.105***	0.196***
	(0.0131)	(0.0140)	(0.0157)	(0.0330)	(0.0295)
Empl. in agri~	0.0102***	0.0172***	0.0231***	0.0800***	0.156***
	(0.00230)	(0.00396)	(0.00718)	(0.0150)	(0.0135)
Fem. labforce~	0.00808*	0.0141**	0.0233**	-0.000482	0.0779***
	(0.00424)	(0.00692)	(0.0101)	(0.00827)	(0.0145)
Female pop~	-0.0144	0.00253	-0.0354	-0.185***	-0.229***
	(0.0125)	(0.0135)	(0.0260)	(0.0424)	(0.0320)
Black pop~	0.000216	0.00472	0.0117**	0.0124	-0.0119
	(0.00193)	(0.00342)	(0.00553)	(0.00837)	(0.00802)
Distance to MX (ln)	-0.0251***	-0.0386***	-0.0440***	-0.0465***	-0.0377***
	(0.00263)	(0.00403)	(0.00438)	(0.00504)	(0.00215)
Income (ln) 1880	0.00525**	0.00811**	0.00859*	0.00461	0.0115***
	(0.00225)	(0.00373)	(0.00513)	(0.00500)	(0.00398)
Urban share 1880	0.00382	0.00467	0.00657	0.0117*	0.0115**
	(0.00257)	(0.00420)	(0.00561)	(0.00666)	(0.00480)
Education 1880	-0.00107	-0.00395	0.00185	0.00634	0.00698
	(0.00293)	(0.00477)	(0.00616)	(0.00744)	(0.00614)
Unempl. 1880	-0.000149	0.00185	0.00162	0.00174	-0.000253
	(0.00109)	(0.00190)	(0.00195)	(0.00301)	(0.00202)
Empl. in agri 1880	0.00341**	0.00497*	0.00567	0.00627	0.00653
	(0.00161)	(0.00283)	(0.00389)	(0.00542)	(0.00457)
Fem. labforce 1880	-0.0182**	-0.0355***	-0.0376**	-0.0414**	-0.0737***
	(0.00849)	(0.0137)	(0.0189)	(0.0173)	(0.0193)
Female pop 1880	0.0172**	0.0275**	0.0260*	0.0301	0.0407**
	(0.00717)	(0.0122)	(0.0158)	(0.0217)	(0.0164)
Black pop 1880	0.000327	-0.000676	-0.00228	-0.00772	-0.00136
	(0.00282)	(0.00448)	(0.00561)	(0.00822)	(0.00782)
States	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	2,846	2,851	2,851	2,851	2,850
R-squared	0.640	0.625	0.583	0.515	0.527

~ Controls from 10 years prior to dependent variable
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Appendix 8C displays the results for the regressions shifting the base year to 1910. Apart from temporal changes to the control variables (vector *Z*) and the *Migration* variable, all factors remain as before. The results are almost a carbon copy of those

presented in Table 8-2. *Migration 1910* is positively and significantly associated with the presence of Latino immigrants 60 to 100 years later. Constant significance levels of below 1% underline the robustness of this link. This means that the settlement pattern of southern and eastern European migrants in 1910 is connected, in the same way as that of their northern and western European predecessors, to Hispanic migration to the US. Latino immigrants have followed in the footsteps of past migration waves, drawn into seemingly migrant-prone regions and bypassing areas of the country that developed a historical aversion to migration.

The large majority of control variables reproduce the results of the regressions in Table 8-2. A high income per capita, a large urban share, a strong agricultural sector, a large participation of women in the labour force, and close proximity to the Mexican border are, once again key for Latino immigrant settlement (Appendix 8C). The coefficients of traditional pull factors, such as high general education or low regional unemployment rates, yet again counter expectations.

The control variables that affected migration waves in historical times are mostly insignificant. The only regional characteristic dating back to the Age of Mass Migration that is significantly connected to the settlement decisions of Latino migrants since the 1970s is the share of foreign-born population in 1910.

In the second robustness check, IV estimation techniques are employed to extract the underlying effect of historical migration on current Latino migration waves. As mentioned earlier, *Migration* is instrumented using the combination of historical population density and public service endowment. Table 8-3 displays the results referring to both base years 1880 and 1910.

The use of IV estimation techniques – which minimise the potential influence of omitted variable bias on the regression coefficients – validates the results of the previous OLS regressions in Table 8-2. Nine out of ten regression outcomes, incorporating shifts between both decades of recent migration inflow and historical base years, confirm the territorial connection between past European and recent Latino migration flows. With significance levels mostly below 5% for 1880 and below 1% for 1910, counties that attracted migrants in 1880 have again become the chosen areas of settlement for Latino immigrants 60 to 130 later. The geography of US migration is therefore shaped by a strong path dependency, with different generations of migrants, regardless of national and/or ethnic origin, discriminating between areas that are more welcoming (migration-prone) and those that are more inimical to migrants (migration-averse). The estimated sign and significance of coefficients for both sets of controls are in line with those reported for the previous analysis (Table 8-3).

TABLE 8-3 THE IMPACT OF HISTORICAL MIGRATION ON LATINO SETTLEMENT PATTERNS, IV 1880, 1910

Dep. Var: % of Latinos	1880					1910				
in county	1970	1980	1990	2000	2010	1970	1980	1990	2000	2010
Migration #	0.0501**	0.0903**	0.0993**	0.109*	0.0562	0.127***	0.231***	0.382***	0.470***	0.451***
	(0.0237)	(0.0377)	(0.0488)	(0.0571)	(0.0606)	(0.0299)	(0.0536)	(0.0708)	(0.109)	(0.129)
Income pc (ln)~	-0.000423	0.00273	-0.000217	0.0353***	0.0436***	-0.000842	0.000631	-0.00623	0.0122	0.0181
	(0.00148)	(0.00272)	(0.00390)	(0.00902)	(0.00786)	(0.00188)	(0.00370)	(0.00502)	(0.0141)	(0.0146)
Urban share~	0.00369***	0.00704***	0.0122***	0.0302***	0.0455***	0.00282*	0.00598**	0.0150***	0.0378***	0.0539***
	(0.00137)	(0.00199)	(0.00246)	(0.00353)	(0.00365)	(0.00169)	(0.00263)	(0.00387)	(0.00534)	(0.00533)
Education~	-0.000232	0.00106	-0.00269	-0.0512***	-0.118***	0.00910*	0.0112	0.00837	-0.0266	-0.113***
	(0.00365)	(0.00679)	(0.00812)	(0.0124)	(0.0122)	(0.00475)	(0.00889)	(0.0116)	(0.0180)	(0.0167)
Unemployment~	0.0314**	0.0355**	0.00229	0.101***	0.191***	0.0458***	0.0724***	0.0502*	0.152***	0.219***
	(0.0135)	(0.0164)	(0.0188)	(0.0343)	(0.0385)	(0.0152)	(0.0222)	(0.0286)	(0.0424)	(0.0530)
Empl. in agri~	0.00870***	0.0161***	0.0192**	0.0816***	0.156***	0.0117***	0.0244***	0.0375***	0.117***	0.175***
	(0.00240)	(0.00437)	(0.00764)	(0.0151)	(0.0201)	(0.00296)	(0.00590)	(0.0108)	(0.0203)	(0.0265)
Fem. labforce~	0.0159**	0.0173**	0.0240**	0.00172	0.0768***	0.0126*	0.0187*	0.0192	0.00566	0.102***
	(0.00668)	(0.00774)	(0.0108)	(0.00855)	(0.0149)	(0.00643)	(0.00996)	(0.0166)	(0.0123)	(0.0215)
Female pop~	-9.44e-05	0.0176	-0.00761	-0.153***	-0.220***	-0.000259	0.0370*	0.0215	-0.165***	-0.209***
	(0.0154)	(0.0170)	(0.0327)	(0.0446)	(0.0600)	(0.0135)	(0.0211)	(0.0388)	(0.0568)	(0.0594)
Black pop~	-0.000105	0.00426	0.0115**	0.0138*	-0.00998	0.000972	0.00350	0.00813	0.00536	-0.0140*
	(0.00189)	(0.00334)	(0.00535)	(0.00828)	(0.00879)	(0.00227)	(0.00368)	(0.00533)	(0.00747)	(0.00835)
Dist. to MX (ln)	-0.0215***	-0.0316***	-0.0364***	-0.0379***	-0.0336***	-0.0183***	-0.0264***	-0.0228***	-0.0192**	-0.0122
	(0.00367)	(0.00573)	(0.00665)	(0.00762)	(0.00777)	(0.00329)	(0.00561)	(0.00683)	(0.00949)	(0.0100)
Income (ln) #	0.00558**	0.00892**	0.00929*	0.00552	0.0120*	-0.00127	-0.00405	-0.00280	-0.000364	0.000301
	(0.00233)	(0.00397)	(0.00527)	(0.00511)	(0.00698)	(0.00143)	(0.00273)	(0.00432)	(0.00563)	(0.00603)
Urban share #	-0.000467	-0.00310	-0.00184	0.00353	0.00785	-0.00667**	-0.0116**	-0.0233***	-0.0290***	-0.0319***
	(0.00387)	(0.00618)	(0.00822)	(0.00897)	(0.00977)	(0.00305)	(0.00505)	(0.00703)	(0.0101)	(0.0113)
Education #	0.00247	0.00345	0.00935	0.0151	0.0111	0.0180**	0.0433***	0.0812***	0.109***	0.127***
	(0.00346)	(0.00576)	(0.00678)	(0.00955)	(0.00925)	(0.00780)	(0.0141)	(0.0184)	(0.0276)	(0.0318)
Unemployment #	0.000384	0.00275	0.00270	0.00276	0.000195	-0.00854	-0.00791	-0.0134	-0.0260	-0.0391
	(0.00120)	(0.00226)	(0.00236)	(0.00332)	(0.00303)	(0.00837)	(0.0136)	(0.0203)	(0.0257)	(0.0299)
Empl. in agri #	0.00857**	0.0145**	0.0155**	0.0173*	0.0117	0.00260	0.00787*	0.0195***	0.0280***	0.0258**
	(0.00346)	(0.00572)	(0.00693)	(0.00896)	(0.00873)	(0.00271)	(0.00478)	(0.00735)	(0.00937)	(0.0102)
Fem. labforce #	-0.0196**	-0.0367***	-0.0385**	-0.0432**	-0.0740***	-0.00557	-0.00406	-0.00709	-0.0246	-0.0119
	(0.00877)	(0.0141)	(0.0191)	(0.0174)	(0.0251)	(0.00444)	(0.00845)	(0.0129)	(0.0165)	(0.0175)
Female pop #	0.0279**	0.0495***	0.0481**	0.0571**	0.0536*	0.124***	0.209***	0.372***	0.503***	0.450***
	(0.0111)	(0.0192)	(0.0235)	(0.0279)	(0.0315)	(0.0307)	(0.0591)	(0.0862)	(0.127)	(0.148)
Black pop #	0.00279	0.00455	0.00266	-0.00310	0.000262	0.00947***	0.0150***	0.0218***	0.0149	0.0231*
	(0.00295)	(0.00476)	(0.00590)	(0.00900)	(0.00891)	(0.00218)	(0.00434)	(0.00715)	(0.0118)	(0.0139)
States	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	2,844	2,849	2,849	2,849	2,848	3,118	3,120	3,120	3,120	3,119
First stage F-stat	18.99	17.14	16.04	12.21	11.69	25.59	33.39	35.75	25.82	18.16

~ Controls from 10 years prior to dependent variable | # respective base year 1880 or 1910 | Robust standard errors in parentheses | *** p<0.01, ** p<0.05, * p<0.

After reshaping the database to a panel structure, the model is run again using a HT-estimation technique. Exploiting the panel data structure while using time and regional fixed effects allows to control for omitted factors that do not change over time but could potentially cause biased regression coefficients. The results are presented in Table 8-4.

For the third time, historical migration is confirmed as fundamental pull factor for ensuing migration waves. There is strong evidence, as indicated by coefficients significant at the 1 percent level, that past US migration patterns shape the geography of future migration in a path dependent way irrespective of origin or ethnicity. The result holds for both base years in question, 1880 and 1910 respectively (Table 8-4). Regardless of whether the focus falls on western and northern European migrants, forming the majority of the first migration wave, or on eastern and southern European migration inflows culminating in 1910, Latin American immigrants settling in the US in the second half of the 20th century and early 21st century were drawn – once other factors influencing recent migration are controlled for – to the exact same regions where their European predecessors settled over a century earlier.

The control variables largely reproduce the results of the OLS and IV regressions. The only differences relate to income per capita which, once the characteristics of the panel structure controlling for regional and time fixed effects are exploited, turns out to be negatively associated to the presence of a large Latino community in any given US county from 1970 onwards. In contrast to the importance of historical migration settlement patterns, income and employment opportunities cannot be considered a key driver for the recent settlement of Hispanic populations in the US.

In short, all three robustness checks have confirmed the initial OLS results. Irrespective of the method, there is a consistent, positive, and strongly significant geographical link between past and current migrant stocks. Settlement patterns of predominantly European migrants around the turn of the 19th to the 20th century have had a strong influence on those of Latin American migrants a century later. The assumed transformation of territories by past migration seems to have turned these regions into magnets for subsequent migrant waves, serving as a long-lasting pull factor – a more important one than any other historical county characteristic. Entirely different backgrounds, ethnicities, and origins, fundamental differences in customs, traditions, and migration incentives, and a time gap of three to five generations have not managed to break the invisible cord linking both migration waves. One might argue that past migration waves have left an imprint that makes particular areas of the US attractive to new generations of migrants. This legacy, generated more than 100 years ago, seems to have become engraved in the receiving territories and represents a fundamental but often ignored pull factor for current migrant inflows, creating a long-lasting division between migration-prone and migration-averse areas.

TABLE 8-4 HT ESTIMATION EXPLOITING QUASI-PANEL STRUCTURE, 1880 AND 1910

Dep. Var. Share of Latino population (1970-2010)	(1) Migrants 1880	(2) Migrants 1910
Migration #	0.0202***	0.0341***
	(0.00626)	(0.00879)
Income per capita (t-10)	-0.0136***	-0.0122***
	(0.00214)	(0.00189)
Urban share (t-10)	0.0258***	0.0264***
	(0.00283)	(0.00277)
Education (t-10)	0.0105**	-0.000224
	(0.00529)	(0.00553)
Unemployment (t-10)	0.0787***	0.0770***
	(0.0111)	(0.0104)
Employment in agriculture (t-10)	0.0257***	0.0277***
	(0.00424)	(0.00400)
Fem. part. in the lab.force (t-10)	0.00250	0.00212
	(0.00385)	(0.00368)
Female population (t-10)	-0.188***	-0.190***
	(0.0361)	(0.0347)
Black population (t-10)	-0.00139	-0.00238
	(0.0110)	(0.0108)
Distance to Mexico (ln)	-0.0365***	-0.0359***
	(0.00378)	(0.00359)
Income (ln) (t_0)	0.00910**	0.00334
	(0.00422)	(0.00224)
Urban share (t_0)	0.00642	0.00114
	(0.00500)	(0.00342)
Education (t_0)	0.0114**	0.0240***
	(0.00495)	(0.00662)
Unemployment (t_0)	0.00146	-0.0176
	(0.00203)	(0.0119)
Employment in agriculture (t_0)	0.00776**	0.00806***
	(0.00349)	(0.00306)
Fem. part in the lab. force (t_0)	-0.0423***	-0.0158*
	(0.0153)	(0.00931)
Female population (t_0)	0.0292**	0.0296
	(0.0147)	(0.0258)
Black population (t_0)	0.00356	0.00274
	(0.00781)	(0.00813)
State fixed effects	*Yes*	*Yes*
Year fixed effects	*Yes*	*Yes*
Observations	14,249	15,597
Number of cnty2	2,851	3,120

Robust standard errors in parentheses | *** p<0.01, ** p<0.05, * p<0.1 |
respective base year 1880 or 1910

8.6 CONCLUSION

Over the course of two and a half centuries, the US has attracted millions of migrants. Those migrants have come from every part of the globe with the aim of building a better life for themselves and their families. Ninety-nine percent of the current US population can trace their ancestry to former migrant stock (Spickard, 2007). Whilst settlement patterns of migrants in the US have been highly uneven, and despite differences from migration wave to migration wave, certain areas of the country have proven more attractive to the foreign-born population than others.

Why is this the case? Traditional migration theory has aimed to explain why migrants flock to particular areas using economic and socio-demographic characteristics as well as natural and cultural aspects as explanations. More recently, the focus has been on the presence of kinship and migrant networks at destination. The presence of a large migrant network, a shared geographical origin, ethnic connection, or common background between already established and newly arriving migrants in a particular place has been identified as a particularly crucial pull factor for new arrivals.

Nevertheless, there is virtually no research on whether past migration serves – once networks based on nationality, ethnic origin, village or family ties have been accounted for – as significant pull factor for new migrants across centuries. This chapter has aimed to cover this gap in the literature by assessing the presence of path dependency in migration flows in the US in absence of a common ethnicity or birthplace. The research intended to ascertain if completely unrelated migration waves to the US, involving very different groups, separated by several generations, have followed similar geographical patterns, creating a division between migration-prone and migration-averse areas. The research has involved putting together a dataset

comprising county data on foreign-born individuals for two of the largest waves of migration to the US – the Age of Mass migration at the turn of the 19th to the 20th century and the large Latin American immigration to the US from the 1960s onwards.

The regression analysis has led to a very clear conclusion: the settlement pattern of the predominantly European migrants at the turn of the 19th to the 20th century strongly influenced the geography of Latino migration to the US a century later. Counties which received large population inflows of European settlers at the end of the 19th century remain significantly more appealing to incoming population from Latin America today than those which were largely bypassed by migrant streams in the 1900s. This result is robust to introducing a large number of controls and estimating the analysis with different econometric techniques. The differences in the composition of migrants between both waves and a considerable time gap has not altered what is a highly persistent geography of migration. Latino migrants have followed in their European predecessors' footsteps 60 to 130 years later, creating a path dependency in migration that is stronger than any historical factor that may have shaped the prosperity of a given county in the US. Historical migration stocks act as important pull factor for subsequent generations of migrants in a more consistent way than many of the standard regional pull factors. Past income levels, historic unemployment rates or levels of education have not left a trace shaping future migration; past migration has. A county's historical foreign-born migrant stock acts as a long-lasting magnet for future generations of migrants.

The results stress that past migration in itself, beyond national or kinship networks, represents an important and long-lasting pull factor for future migrants. One possible explanation for this phenomenon might be that migrants indeed rework the territory where they settle in large numbers. Late 19th century European migrants seem to have

left a legacy in specific parts of the US which survives until today and still appeals to new generations of migrants. The migration buzz generated in historic times by Germans, Italians, Poles, or Irish settlers seems to have become engraved in the receiving territories and to persist well after the 'original' migrants are long gone and their descendants have long been assimilated and become American.

The results of the analysis point to factors that have hitherto not been considered in migration policy and provide interesting food for thought for future migration and development policy design. Policies crafted to attract migrants into a specific region need to take into account not only the region's economic and socio-demographic characteristics but also the territory's migration history. Areas with a limited background in immigration history may lack an important pull factor for new migrants: 'the buzz of past migration'. A lack of 'migrant buzz' would therefore have to be overcome by other means – i.e. better local amenities or additional economic incentives – in order to attract a similar magnitude of incoming migration.

However, the mechanisms through which path dependent migration mechanisms are created remain a mystery. Further research is needed in this respect. In-depth analyses of migration-prone areas, such as New England, California, or the north west of the US, can provide crucial insights about the mechanisms that make regions consistently migration-prone. Similarly, studies of new destination areas without an important history of immigration, such as the South of the US, could lead to eye-opening results as to what factors can help overcome a lack of history in migration. Furthermore, there is a need to pinpoint what is exactly behind migration path dependence and a 'regional migration buzz'. In-depth, more anthropological case studies can play an important role in understanding how this buzz is created, how it reproduces itself, and how exactly it affects the attractiveness of a place to newly arriving migrants.

8.7 REFERENCES

Alexander, J. (2007): *Daily life in immigrant America: 1870–1920*. London: Greenwood Press.

Bailey, R. (2005): New Immigrant communities in the North Carolina Piedmont Triad: Integration issues and challenges, in Goździak, E. M. and Martin, S. F. (eds) *Beyond the Gateway: Immigrants in a Changing America*, 57–86. Lanham: Lexington Books.

Baltagi, B. (2008): *Econometric analysis of panel data*. 4th edn. Chichester, UK: John Wiley & Sons.

Baltagi, B., Bresson, G. and Pirotte, A. (2003): Fixed effects, random effects or Hausman–Taylor? A pretest estimator, *Economics Letters*, *79*(3), 361–369.

Bauer, T., Epstein, G. and Gang, I. (2007): The influence of stocks and flows on migrants' location choices, *Research in Labor Economics*, *26*, 199–229.

Bauer, T., Epstein, G. and Gang, I. (2009): Measuring ethnic linkages among migrants, *International Journal of Manpower*, *30*(1/2), 56–69.

Bauer, T. and Zimmermann, K. (1997): Network migration of ethnic Germans, *The International Migration Review*, *31*(1), 143–149.

Bergad, L. and Klein, H. (2010): *Hispanics in the United States: a demographic, social, and economic history, 1980-2005*. Cambridge: Cambridge University Press.

Bertocchi, G. and Strozzi, C. (2006): *The age of mass migration: Economic and institutional determinants*, Discussion paper Series, IZA Institute for the Study of Labour, 2499. Bonn, Germany.

Biagi, B., Faggian, A. and McCann, P. (2011): Long and short distance migration in Italy: The role of economic, social and environmental characteristics, *Spatial Economic Analysis, 6*(1), 111–131.

Bikhchandani, S., Hirshleifer, D. and Welch, I. (1998): Learning from the behaviour of others: Conformity, fads, and informational cascades, *The Journal of Economic Perspectives, 12*(3), 151–170.

Bodvarsson, Ö., Simpson, N. and Sparber, C. (2014): Migration Theory, in Chiswick, B. and Miller, P. (eds) *Handbook of the Economics of International Migration*, 3–51. Burlington: Elsevier Science.

Borjas, G. (1987): Self-selection and the earnings of immigrants. *American Economic Review, 77*, 531-553.

Borjas, G. (1990): *Friends or strangers: The impact of immigrants on the U.S. economy*. New York: Basic Books.

Borjas, G., Bronars, S. and Trejo, S. (1992): Self-selection and internal migration in the United States, *Journal of Urban Economics*, *32*(2), 159-185.

Bowles, S. (1970): Migration as investment: Empirical tests of the human investment approach to geographical mobility, *The Review of Economics and Statistics*, *52*(4), 356–362.

Bruhn, J. (2005): *The sociology of community connections*. Boston, MA: Springer US.

Bump, M. (2005): From temporary picking to permanent plucking: Hispanic newcomers, integration, and change in the Shenandoah Valley, in Goździak, E. M. and Martin, S. F. (eds) *Beyond the Gateway: Immigrants in a Changing America*, 137–177. Lanham: Lexington Books.

Burda, M. and Wyplosz, C. (1992): Human capital, investment and migration in an integrated Europe, *European Economic Review*, *36*(2), 677–684.

Cafferty, P. and Engstrom, D. (2000): *Hispanics in the United States: an agenda for the twenty-first century*. New Brunswick: Transaction Publishers.

Cairns, D. and Menz, S. (2009): Youth on the move? Exploring youth migrations in Eastern Germany and Northern Ireland, in Geisen, T. and Riegel, C. (eds) *Jugend, Partizipation und Migration: Orientierungen im Kontext von Integration und Ausgrenzung*, 325–339. Wiesbaden: VS Verlag für Sozialwissenschaften.

Carrington, W., Detragiache, E. and Vishwanath, T. (1996): Migration with endogenous moving costs, *American Economic Review*, *86*(4), 909–930.

Chiquiar, D. and Hanson, G.H. (2005): International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States, *Journal of Political Economy*, *113*(2), 239–281.

Clark, X., Hatton, T. and Williamson, J. (2007): Explaining U.S. immigration, 1971–1998, *The Review of Economics and Statistics*, *89*(2), 359–373.

Daniels, R. (1990): *Coming to America: A history of immigration and ethnicity in American life*. New York: Harper Perennial.

Day, K. (1992): Interprovincial migration and local public goods, *Canadian Journal of Economics*, *25*(1), 123–144.

Donato, K., Tolbert, C., Nucci, A. and Kawano, Y. (2008): Changing faces, changing places: The emergence of new nonmetropolitan immigrant gateways, in Massey, D. S. (ed.) *New Faces in New Places*, 75–98. New York, NY: Russell Sage Foundation.

Douglas, K. and Sáenz, R. (2008): No phone, no vehicle, no English, and no citizenship: The vulnerability of Mexican immigrants in the United States, in Hattery, A., Embrick, D. G., and Smith, E. (eds) *Race, Human Rights and Inequality*, 161–180. Lanham: Rowman and Littlefield.

Epstein, G. (2002): Informational cascades and decision to migrate, *IZA Discussion paper series*, No. 445, Bonn, Germany.

Epstein, G. and Gang, I. (2006): The influence of others on migration plans, *Review of Development Economics*, *10*(4), 652–666.

Faggian, A. and McCann, P. (2008): Human capital, graduate migration and innovation in British regions, *Cambridge Journal of Economics*, *33*(2), 317–333.

Ferguson, M., Ali, K., Olfert, M. R. and Partridge, M. (2007): Voting with their feet: Jobs versus amenities, *Growth and Change*, *38*(1), 77–110.

Greenwood, M. (1997): Chapter 12 Internal migration in developed countries, *Handbook of Population and Family Economics*, *1*, 647–720.

Haapanen, M. and Ritsilä, J. (2007): Can migration decisions be affected by income policy interventions? Evidence from Finland, *Regional Studies*, *41*(3), 339–348.

Hansen, A. (2005): Black and white and the other: International immigration and change in metropolitan Atlanta, in Goździak, E. M. and Martin, S. F. (eds) *Beyond the Gateway: Immigrants in a Changing America*, 87–110. Lanham: Lexington Books.

Harris, J. and Todaro, M. (1970): Migration, Unemployment and Development: A Two-Sector Analysis, *The American Economic Review*, *60*(1), 126–142.

Hatton, T. and Williamson, J. (1998): *The Age of Mass Migration: Causes and economic impact*. New York: Oxford University Press.

Hicks, J. (1932): *The theory of wages*. London: Macmillan.

Hugo, G. (1981): Village-community ties, village norms, and ethnic and social networks: A review of evidence from the third world, in Dejong, G. and Gardner, R. (eds) *Migration Decision Making: Multidisciplinary Approaches to Micro-Level Studies in Developed and Developing Countries*, 186–224. New York: Pergamon Press.

Jaeger, D. (2007): Green Cards and the Location Choices of Immigrants in the United States, 1971–2000, *Research in Labor Economics*, *27*, 131–183.

Jewell, R. and Molina, D. (2009): Mexican migration to the US: A Comparison of income and network effects, *Eastern Economic Journal*, *35*(2), 144–159.

Joly, D. (2000): Some structural effects of migration on receiving and sending countries, *International Migration*, *38*(5), 25–40.

Ketterer, T. and Rodríguez-Pose, A. (2015): Local quality of government and voting with one's feet, *The Annals of Regional Science*, *55*(2), 501–532.

Marks, C. (1989): *Farewell - we're good and gone: the great Black migration*. Bloomington: Indiana University Press.

Massey, D., Alarcon, R., Durand, J., and Gonzalez, H. (1987): *Return to Aztlan: The social process of international migration from western Mexico*. Berkeley: University of California Press.

Massey, D., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J. (1993): Theories of International Migration: A Review and Appraisal, *Population and Development Review*, *19*(3), 431–466.

Massey, D., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A. and Taylor, J. (2005): *Worlds in motion: Understanding international migration at the end of the millennium*. Oxford: Oxford University Press.

Massey, D. and Capoferro, C. (2008): The geographic diversification of American immigration, in Massey, D. S. (ed.) *New Faces in New Places*, 25–50. New York: Russell Sage Foundation.

Massey, D. and España, F. (1987): The Social Process of International Migration, *Science*, *237*(4816), 733–738.

McGovern, P. (2007): Immigration, labour markets and employment relations: Problems and prospects. *British Journal of Industrial Relations*, *45*(2), 217–235.

McKenzie, D. and Rapoport, H. (2010): Self-Selection patterns in Mexico-U.S. migration: The Role of migration networks, *The Review of Economics and Statistics*, *92*(4), 811–821.

Migration Policy Institute (2017): *U.S. immigrant population and share over time, 1850-Present*. Available at: http://www.migrationpolicy.org/programs/data-hub/us-immigration-trends#history (Accessed: 28 June 2017).

Minnesota Population Center (2016): *National Historical Geographic Information System*, Version 11.0 [Database], Minneapolis: University of Minnesota.

Munshi, K. (2003): Networks in the modern economy: Mexican migrants in the US labor market, *The Quarterly Journal of Economics*, *118*(2), 549–599.

Ödberg, S. (1997): Theories in inter-regional migration: An overview, in Blotevogel, H. and Fielding, A. (eds) *People, jobs and mobility in the new Europe*, 3–22. Chichester: Wiley.

Ottaviano, G. and Puga, D. (1998): Agglomeration in the global economy: A survey of the "New Economic Geography", *World Economy*, *21*(6), 707–731.

Özgen, C., Nijkamp, P. and Poot, J. (2011): Immigration and innovation in European regions, *IZA Discussion Papers*, Institute for the Study of Labor (IZA), No. 5676. Bonn, Germany.

Partridge, M. (2010): The duelling models: NEG vs amenity migration in explaining US engines of growth, *Papers in Regional Science*, *89*(3), 513–536.

Partridge, M. and Rickman, D. (2008): Distance from urban agglomeration economies and rural poverty, *Journal of Regional Science, 48*(2), 285–310.

Partridge, M. and Rickman, D. (1996): The role of industry structure, costs, and economic spill-overs in determining state employment growth rates, *The Review of Regional Studies*, *26*(3), 235–264.

Piore, M. (1979): *Birds of passage: Migrant labor and industrial societies*. Cambridge: Cambridge University Press.

Pissarides, C.A. and McMaster, I. (1990): Regional migration, wages and unemployment: Empirical evidence and implications for policy, *Oxford Economic Papers*, *42*(4), 812–831.

Radu, D. (2008): Social Interactions in economic models of migration: A review and appraisal, *Journal of Ethnic & Migration Studies*, *34*(4), 531–548.

Ranis, G. and Fei, J. (1961): A theory of economic development, *The American Economic Review*, *51*(4), 533-565.

Rappaport, J. (2007): Moving to nice weather, *Regional Science and Urban Economics*, *37*(3), 375–398.

Ritsilä, J. and Ovaskainen, M. (2001): Migration and regional centralization of human capital, *Applied Economics*, *33*(3), 317–325.

Rodríguez-Pose, A. and von Berlepsch, V. (2014): When migrants rule: The Legacy of mass migration on economic development in the United States, *Annals of the Association of American Geographers*, *104*(3), 628–651.

Rodríguez-Pose, A. and von Berlepsch, V. (2015): European migration, national origin and long-term economic development in the United States, *Economic Geography*, *91*(4), 393–424.

Rodríguez-Pose, A., Ketterer, T. and Castells-Quintana, D. (2015): Do we follow the money? The drivers of migration across regions in the EU, *Region*, *2*(2), 27–45.

Rodríguez-Pose, A. and Ketterer, T. (2012): Do local amenities affect the appeal of regions in Europe for migrants?, *Journal of Regional Science*, *52*(4), 535–561.

Roy, A. (1951): Some thoughts on the distribution of earnings, *Oxford Economic Papers*, *3*, 135-146.

Ruggles, S., Genadek, K., Goeken, R., Grover, J. and Sobek, M. (2015): *Integrated Public Use Microdata Series: Version 6.0 [dataset]*. Minneapolis: University of Minnesota.

Sáenz, R. and Morales, M.C. (2015): *Latinos in the United States: Diversity and change*. Cambridge: Polity Press.

Shah, N. and Menon, I. (1999): Chain migration through the social network: Experience of labour migrants in Kuwait, *International Migration*, *37(*2), 361–382.

Simon, C. (2004): Industrial reallocation across US cities, 1977–1997, *Journal of Urban Economics*, *56*(1), 119–143.

Simpson, N. and Sparber, C. (2013): The short- and long-run determinants of less-educated immigrant flows into U.S. States, *Southern Economic Journal*, *80*(2), 414–438.

Sjaastad, L. (1962): The Costs and returns of human migration, *Journal of Political Economy*, *70*(5, Part 2), 80–93.

Spickard, P. (2007): *Almost all aliens: immigration, race, and colonialism in American history and identity*. New York: Routledge.

Staiger, D. and Stock, J. (1997): Instrumental Variables regression with weak instruments, *Econometrica*, *65*(3), 557–586.

Stock, J. and Yogo, M. (2005): Testing for weak instruments in linear IV regression, in Donald, A. and Stock, J. (eds) *Identification and Inference for Econometric Models*.

New York: Cambridge University Press.

Storper, M. and A. Venables (2004): Buzz: Face-to-face contact and the urban economy, *Journal of Economic Geography*, *4*(4), 351-370.

Tassinopoulos, A. and Kristensen, H. (1998): Mobility and migration of labour in the Euorpean Union and their specific implications for young people, *Publications of the European Communities*. Luxemburg/Thessaloniki, 1–154.

Taylor, P., Hugo Lopez, M., Hamar Martinez, J. and Velasco, G. (2012): *When labels don't fit: Hispanics and their views of identity*. Pew Research Center Hispanic Trend Project, Washington, D.C.

Todaro, M. (1969): A model of labor migration and urban unemployment in less developed countries, *The American Economic Review*, *59*(1), 138–48.

Treyz, G., Rickman, D., Hunt, G. and Greenwood, M. (1993): The dynamics of U.S. internal migration, *The Review of Economics and Statistics*, *75*(2), 209-214.

U.S. Census Bureau (2016): *Population Estimates Program (PEP)*. Population and Housing Unit Estimates. Washington, D.C.

U.S. Department of Homeland Security (2011): *Yearbook of Immigration Statistics: 2010*. Washington, D.C.

Winters, P., De Janvry, A. and Sadoulet, E. (2001): Family and community networks in Mexico-U.S. Migration, *The Journal of Human Resources*, *36*(1), 159–184.

Yap, L. (1977): The attraction of cities: A review of the migration literature, *Journal of development economics*, *4*(3), 239–264.

Zimmermann, K. (2005): European labour mobility: Challenges and potentials, *De Economist*, *153*(4), 425–450.

8.8 APPENDIX 8A

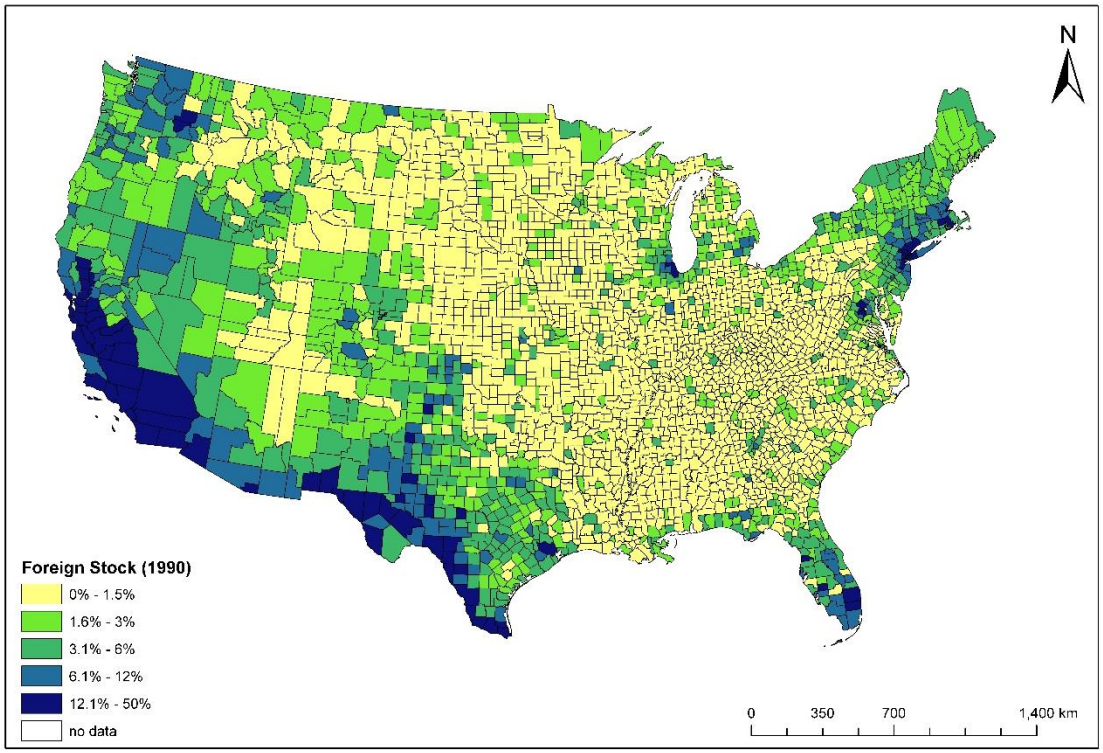


FIGURE 8-5 FOREIGN-BORN POPULATION AS SHARE OF TOTAL POPULATION, 1990

Data source: Minnesota Population Center (2016); own elaboration

8.9 APPENDIX 8B

TABLE 8-5 VARIABLES DESCRIPTIONS AND SOURCES

Variable	Description	Source
Main variables of interest		
Latinos	Share of foreign-born Latino migrants relative to total population in county *i* in year *t**. Latino is defined if originating from Mexico, the Caribbean, and Central and South American countries.	NHGIS
Migration 1880 / 1910	Share of foreign-born population relative to total population in county *i* in year t_0*	IPUMS USA
Instruments		
Population density	Population per square mile in county *i* in year 1880	ICPSR
Public good endowment	Share of employment in health, education, and public administration relative to total employment in any given county *i* in 1880 (or 1910)	IPUMS USA
Control Variables		
Income	Measure for county income: • Historical years: Constructed mean income score on county level. Individual income levels assigned to occupational data on the basis of median total income per occupation in hundreds of 1950 dollars, as natural log in t_0. Basis for construction is 1880/1910 total size of county population • 1960-2000: Natural log of average income per capita levels in 2010 US dollars on county level for county *i* (adjusted for inflation using the BLS converter)	IPUMS USA BEA and Social Explorer
Urban share	Share of urban population relative to total population in county *i* in year t_0 or *t-k*, respectively	ICPSR and Social Explorer

Variable	Description	Source
Education	Percentage of population of county *i* with college degree in *t-k*	Social Explorer
Literacy	Literacy rate in county *i* in t_0	IPUMS USA
Unemployment	Unemployment rate in county *i* in t_0 and *t-k* Proxy in 1880 (unemployment rate not available): Months unemployed last year	IPUMS USA and ICPSR
Agriculture	Percentage of labour force employed in agriculture in county *i* in t_0 and *t-k*	IPUMS USA and ICPSR
Female Participation	Female participation rate in the labour force in county *i* in t_0 and *t-k*	IPUMS USA and ICPSR
Female	Percentage of female population in county *i* in t_0 and *t-k*	Social Explorer
Black Population	Percentage of black population in county *i* in t_0 and *t-k*	Social Explorer
DistMX (ln)	Distance in km between county *i*'s centroid and the nearest point on the continental border between the US and Mexico (as natural log)	Own construction
State	State dummies	Own construction

***** t_0 = 1880 or 1910 while t = 1970, 1980, 1990, 2000, 2010. k= 10

8.10 APPENDIX 8C

TABLE 8-6 THE IMPACT OF HISTORICAL MIGRATION ON LATINO SETTLEMENT PATTERNS, OLS 1910

Dep. Var: % of Latinos in county	(1) 1970	(2) 1980	(3) 1990	(4) 2000	(5) 2010
Migration 1910	0.0159***	0.0224***	0.0279***	0.0384***	0.0370***
	(0.00460)	(0.00778)	(0.00969)	(0.0122)	(0.00822)
Income pc (ln)~	0.00121	0.00597***	0.00340	0.0471***	0.0553***
	(0.00112)	(0.00229)	(0.00302)	(0.00690)	(0.00462)
Urban share~	0.00320***	0.00554***	0.0130***	0.0315***	0.0483***
	(0.00122)	(0.00181)	(0.00240)	(0.00382)	(0.00281)
Education~	-0.000553	-0.00174	-0.00409	-0.0679***	-0.137***
	(0.00296)	(0.00542)	(0.00687)	(0.0103)	(0.00942)
Unemployment~	0.0290**	0.0465***	0.0214	0.125***	0.225***
	(0.0122)	(0.0138)	(0.0136)	(0.0299)	(0.0290)
Empl. in agri~	0.0134***	0.0210***	0.0310***	0.0760***	0.144***
	(0.00221)	(0.00362)	(0.00577)	(0.0125)	(0.0119)
Fem. labforce~	0.00500	0.0111*	0.0238**	0.00323	0.0730***
	(0.00394)	(0.00625)	(0.00958)	(0.00807)	(0.0140)
Female pop~	-0.0171	0.00528	-0.0491**	-0.222***	-0.248***
	(0.0106)	(0.0122)	(0.0240)	(0.0401)	(0.0312)
Black pop~	-0.00282	0.000354	0.00579	0.00645	-0.0159**
	(0.00176)	(0.00292)	(0.00385)	(0.00622)	(0.00676)
Distance to MX (ln)	-0.0241***	-0.0380***	-0.0431***	-0.0467***	-0.0378***
	(0.00229)	(0.00365)	(0.00411)	(0.00483)	(0.00204)
Income (ln) 1910	-7.00e-05	-0.00250	1.51e-05	0.00301	0.00282
	(0.000899)	(0.00174)	(0.00224)	(0.00341)	(0.00328)
Urban share 1910	0.00192	0.00374*	0.00327	0.00289	-0.000822
	(0.00138)	(0.00223)	(0.00305)	(0.00488)	(0.00395)
Education 1910	-0.00453	-0.000543	0.00847	0.0157*	0.0385***
	(0.00334)	(0.00560)	(0.00716)	(0.00889)	(0.00795)
Unemployment1910	-0.00304	-0.00178	-0.00476	-0.0158	-0.0273
	(0.00623)	(0.00949)	(0.0123)	(0.0168)	(0.0186)
Empl. in agri 1910	-0.00312**	-0.00195	0.00313	0.00963**	0.00905*
	(0.00135)	(0.00246)	(0.00312)	(0.00465)	(0.00465)
Fem. labforce 1910	-0.00422	-0.00538	-0.0125*	-0.0310***	-0.0214*
	(0.00301)	(0.00527)	(0.00710)	(0.0104)	(0.0115)
Female pop 1910	0.0168	0.000381	0.0145	0.0566	0.0143
	(0.0115)	(0.0213)	(0.0270)	(0.0367)	(0.0325)
Black pop 1910	0.00583***	0.00476*	0.00321	-0.0123	0.000612
	(0.00174)	(0.00257)	(0.00407)	(0.00770)	(0.0123)
States	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	3,118	3,120	3,120	3,120	3,119
R-squared	0.639	0.614	0.573	0.514	0.523

~ Controls from 10 years prior to dependent variable

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

9 CONCLUSION

In a world that has struggled and continues to struggle with a lasting economic downturn, labour market and security issues, as well as increasing welfare burdens, rising population mobility and pressures by migrants on national borders have brought the analysis of the economic impact of migration to the very fore of the academic and political debate. From a purely scholarly perspective – and despite a few voices emphasizing the downsides of population mobility – the overwhelming majority of the migration economics literature considers inward migration growth-enhancing. Immigration fosters innovation, technological progress and a rise in productivity within the receiving countries.

One crucial aspect of population mobility has, however, been largely overlooked: the long-term implications of inward migration. Migration research has generally focused on the impact migrants have on the receiving society in the short-term. The time horizon of most studies covers only a few years, with very few venturing into the impact of immigration over ten or twenty years. Hence, the potential far-reaching consequences of inward migration over longer periods of time have been largely neglected by the scientific community and, as a consequence, have been absent from the political as well as policy debate.

Only recently has the long-term economic impact of migration on receiving regions begun to attract the attention it deserves. Less than a handful of studies so far have extended the traditional few years to one-generation type of analysis of migration to considerably longer time-frames. Rodríguez-Pose and von Berlepsch (2014), for example, have revealed that migration shocks leave a very long-lasting economic

impact: areas of the US that attracted important migration stocks more than 130 years ago were found to be significantly more prosperous today than those that were bypassed by successive migration waves. Sequeira et al. (2017) have confirmed these results. These findings point towards a need for further research analysing the reasons for this robust and enduring impact as well as evaluating the potential transmission channels. This dissertation has intended to fill a number of these gaps in our knowledge.

The thesis has resorted to the analysis of historical migration stock and foreign-born settlement patterns in the US at the turn of the 19th to the 20th century and connected them with current levels of economic development. The first part of the research has focused on different compositional features of the migrant stock, such as diversity, gender, and distance travelled. The second part has delved into one potential transmission channel explaining the impact of historical migration across time: path dependence.

Diversity was shown to be one of the decisive factors determining long-term regional economic development. Measured in its two dimensions – fractionalisation and polarisation – population diversity in US counties was found not only to affect local economic performance in the short-term, but also to leave an economic trace for decades well after the initial migration shock had taken place. Stretching the time dimension within the analysis from 20 to 130 years, it was found that counties that attracted migrants from a multitude of different national and international origins between 1880 and 1910 are significantly more prosperous today than those which were marked by more homogenous societies in the past. Migrant diversity, measured by population fractionalisation, is thus highly beneficial for long-term economic development.

The second employed measure of diversity – population polarisation – has, by contrast, long-lasting negative economic effects. Counties with highly polarised population compositions experienced a significant negative economic impact, lowering the county's future growth path. Highly polarised US counties at the height of the great migration wave of the late 19th and early 20th centuries are considerably poorer today. More importantly though, the impact of both dimensions of diversity on county-level economic trajectories remains extremely robust over time. While the effect was higher in the initial 10 to 30 years after migration, past population diversity associated with inward migration of heterogeneous population groups remained a very robust predictor of territorial disparities in the US for more than 100 years.

The analysis of the long-term impact of population diversity presented in Chapter 5 thus exposes the relevance of diversity for the long-term prosperity of places. It also discovers how the relationship between diversity and economic performance can be double-edged, depending on the dimension of diversity considered. Migration-prone counties have significantly benefited from a diverse migrant inflow during the Age of Mass Migration. Bringing their variety of skill-sets, abilities, ideas, perspectives, and experiences along with them, the interaction of migrants in a common location spurred high levels of economic activity. This surge of new ideas triggered economic dynamism resulting not only in substantial short-term economic prosperity as predicted by Jacobs (1961, 1969), Florida (2002), or Saxenian (2006), among others. It also built the foundations for sustainable long-term regional wealth. This long run elevated growth path, however, is shown to hinge on one condition: the capacity to establish bridges across ethnic, cultural, and/or national groups. The presence of deep divides among migrant groups inevitably leads to polarisation, social tension, and even conflict. Once 'bridging' across groups ceases to exist, once a particular group

becomes too large and dominant, the positive externalities linked to diverse backgrounds, perspectives, and ideas fade away and become overshadowed by rising polarisation within the population. Paving the way for the formation of ghettos, rivalry and conflict, polarisation significantly undermines economic development – again, not only in the short term, but for decades to come.

A different compositional feature of the migrant population was analysed in chapter 6. The gender component of migration was shown to have a significant and ongoing macroeconomic impact. Contrary to expectations, high concentrations of immigrant women at the turn of the 19th to the 20th century were found to significantly lower county-level economic prospects, both over the short- as well as over the long-term. Areas marked by a lower presence of immigrant women, however, realised significantly higher levels of prosperity. These findings seem to be heavily influenced by the low levels of female participation in the official labour force, as one century ago the large majority of women left their jobs at a relatively young age, mostly following marriage, in order to become 'homemakers'. Their active engagement in the shadow economy was not accounted for in traditional measures of economic activity.

Immigrant women, however, influenced the prosperity of the territories where they settled in two ways: directly, via their contribution to creating large geographical concentrations of migrants (as shown above) and indirectly, via their children. Communities with large shares of children born to immigrant mothers during the Age of Mass Migration have coped significantly better over the past century than areas where the majority of children descended from two American-born parents. The highest economic development levels were achieved in counties where foreign-born mothers were more likely to have successfully integrated, usually by marrying American-born men. Children from 'mixed' foreign/American parentage contributed

to making their communities more dynamic, increasing economic activity and leaving a significantly beneficial imprint on their county's long-term growth prospects.

In the end, women managed to leave a positive long-term legacy for economic development. Even though the social norms of the past seriously dented their capacity to use their talent and entrepreneurship and transform their economic potential into official, recorded, economic activity, they left a positive economic trace in a more intricate and concealed way. By shaping the gender balance in the local migrant stock as well as in their role as 'carriers of culture' for their children (to name only one potential explanation), they became a fundamental transmission channel for medium- to long-term territorial prosperity. We speculate women immigrants to have passed the risk-seeking, hard-working, dynamic, and entrepreneurial character normally associated with immigrants to their children. They also left them the traditions, culture, and mentality of their place of origin in legacy. The more integrated these women became, the higher their indirect economic effect. One could assume that children born to immigrant women and locals grew into more independent and determined individuals, capable of building a better future for themselves and their offspring. Spurred on and supported by their 'mixed' families, this young, first-born American generation seems to have been at the base of an economic dynamism that seems to have become engraved in the very territories where they were born. This dynamism is still distinguishable in the economic development of the counties that attracted masses of immigrants more than a century ago and has not waned despite decades of high population mobility in the US.

A similar strong and positive long-term impact has been identified for internal migrants in chapter 7. Counties that attracted a big inflow of US-born citizens in the late 19th and early 20th centuries are considerably wealthier today (once other factors

are controlled for) than areas avoided by internal migrants. Hence, in a similar way to international immigrants, domestic migrants have proven to be a source of economic activity, raising the long-term economic growth path of territories. Demoting domestic migration therefore to a simple population redistribution mechanism at one moment in time, ignores an important factor behind the long-term economic dynamism of places.

One decisive element in this interdependence was found to be the distance travelled by the migrant between outset and destination. Counties which attracted large shares of long-distance domestic migrants have been decidedly more dynamic over time. Large shares of population born locally and/or originating from the same state have not triggered the same levels of economic activity, resulting in significantly lower long-run economic development. A self-selection mechanism of internal migrants is assumed to be the root-cause behind this connection. Long-distance domestic migrants have transmitted a similar economic legacy as international migrants. Their attitude when taking high risks crossing hundreds of kilometres across often undeveloped or unorganized, often dangerous territory, wiping their slate clean to start all over is assumed to be at the root of the substantial economic dynamism of the areas where they settled. The internal migrants' determination to seek a more promising future in combination with the institutional constructs they brought from their far-flung places of origin seem to have become etched in the places of destination, making them more resilient and prosperous in the long-term.

But how can the economic dynamism spurred by large and diverse groups of ancient migrants still affect today's economic performance? The institutional transfer mechanism from past migration to current levels of development is analysed in chapter 8. This chapter seeks to answer whether the institutional imprint of past migration has

become embedded into the very territory, leading to the formation of migration-prone and migration-averse regions. By comparing the two largest waves of migration to the US – the European immigration of the 19th century and the Latino immigration from 1970 until today – evidence is found to support the idea that migrants, regardless of origin and ethnicity and regardless of how much time has passed between their arrival, are connected across time by an invisible cord. Despite many apparent differences between the two migrant waves – different continents, different ethnicity, traditions, customs, and migration incentives – and the time gap between both movements – three to five generations – historical migration patterns have, to a large extent, determined recent migration destinations. Regions that received large stocks of European migrants more than 100 years ago have and are still attracting significantly more population from Latin America today than those areas of the US which did not appeal to migrants around the turn of the 19th to the 20th century. Historical migration has thus acted as pull factor for current migrants in a stronger and more robust way than most of the traditional characteristics of the places of destination identified within the scientific literature. A high education level, low unemployment, or high levels of income have been less of a draw to Latin American immigrants than historical migration patterns.

These findings provide additional evidence of the presence of an institutional imprint created by migrant settlement patterns more than 100 years ago. Immigrants to (and within) the US in the late 19th century seem to have radically transformed the territories where they settled. Large numbers of migrants created not only a macroeconomic shock for the labour market and local economy at the time, but shook established institutional frameworks to the bone. The surge in population diversity across many US communities coupled with the risk-seeking and entrepreneurial mentality of the migrant bringing along the cultural baggage from their home regions and countries,

altered local habits, values, traditions and customs. The (what we assume to be) migration 'buzz' in receiving areas seems to have become ingrained in the territories' own institutional frameworks. It has acted as a pull factor for new migrant waves today, generating and perpetuating a path dependence which has extended over generations of migrants. Indeed, "as migrants settle, they rework the destination itself" (Silvey and Lawson 1999: 125).

Overall, the findings presented in this dissertation have important policy implications. Even though conditions, circumstances and global challenges today are highly different from those prevailing in the late 19th and early 20th centuries, in times where the US as well as many other developed countries are trying to lower immigration numbers, build up restrictions that often allow them to hand-pick those who enter (e.g. McGovern, 2012), and refuse entry to immigrants from 'unwanted' countries, cultures or religions, the results of this research provide considerable food for thought about the economic and social implications of adopting these sort of measures.

Restrictive migration policies are often justified by politicians proposing them as a way to 'protect' the locals or the 'native' population from the ills of lower wages and unemployment. While there is limited scientific evidence that preventing population mobility will achieve any of these goals, restricting migration will also have long-term, often ignored, economic consequences. Curbing migration limits diversity levels and, thus, prevents potentially far-reaching, growth-enhancing economic mechanisms from materialising. The aging societies of the developed world, in particular, are especially vulnerable to this. In the long-term, they cannot afford to miss out on the variety of ideas, perspectives, skills and experiences that immigrants bring. Massively restraining migration would imply foregoing much-needed innovation and technological progress. And yet, it is exactly these countries that are trying to shield

themselves from these positive externalities associated to migration, on the basis of frequently ill-justified security, welfare or economic concerns related to migration. Restricting mobility and thus curtailing population diversity levels has negative economic consequences in the short-, medium- and long-term that societies increasingly relying on the innovative capacity of their populations cannot shun.

Nevertheless, as the research has shown, facilitating population mobility is not a magic solution. More migration and a resulting rise in diversity will only lead to a surge in economic potential if the incoming population becomes rapidly integrated into the communities of the receiving areas. In order for the long-term growth enhancing mechanism associated with migration to kick off, a system to prevent the formation of ghettos and facilitate interaction, avoid the formation of deep divides between different population groups and foster a dialogue between the foreign and native populations has to be set in place. Placing emphasis on the long-lasting and growth-enhancing impact of population diversity, while simultaneously nurturing the necessary integration mechanisms can prepare the ground for not only a short-term peaceful integration of migrants, but also for durable long-term regional prosperity.

Migration policy should also pay special attention to the gender balance of the migrant population. Hugely male dominated migration flows can lead not only to detrimental social consequences, but entail short- and long-term undesirable economic effects. Migration policies (unintentionally) granting preferential entry to men due to, for example, placing emphasis solely on skill-levels and personal income, can end up in missing out on important positive economic influences of female migration. However, in order for female migrants not to impact negatively in economic terms, the results indicate that women need to become rapidly integrated into the receiving societies. Migration policy therefore needs to also pave the way for the participation of

immigrant women in local labour markets and societal structures. They need to be free to take their own decisions in issues such as education, employment and marriage. Only if integration is successful, women migrants can demonstrate their full positive economic potential, consistently boosting economic activity levels over very long time frames.

The results of the dissertation also provide considerable food for thought for those involved in designing regional development strategies. Internal migrants stimulate economic activity at their places of destination. They leave a long-lasting positive impact for economic development. Hence, taking into account local push- and pull factors which determine the region's attractiveness to domestic migrants is a must for local and regional development policy design. Cities and regions need to become and remain attractive for long-distance (internal and external) migrants. The more a place can attract immigrants from far flung areas, the more it will benefit economically in the short-, medium- and long-run. Simply relying on migrants from neighbouring areas will not do the trick.

Regional policy design also needs to consider the region's migration history, especially with regards to international migration. Migrant streams established in historical times shape the settlement pattern and direction of future migration waves, independent of the origin, ethnicity or culture of the migrants. Generating a 'migration buzz' and engraving it in the institutional framework of the territory represents the best guarantee for becoming a migration-prone region. Places that manage to create this welcoming and open migration vibe at any given point in time have the best chance of attracting future generations of migrants. This also assists in remaining ahead of the rest in wealth generation for long periods of time. Regions lacking this migration 'buzz', by contrast, will see their potential limited.

Finally, the research conducted in this dissertation brings to the fore the real importance of migration for economic development in the long-term. The composition of the migrant stock, their skills, cultural baggage, institutional constructs and historic settlement patterns carry long-term economic consequences that still affect economic development in the US (and possibly elsewhere) to this very day. Migration is, by far, the most important factor shaping long-term territorial growth in the US. Whether a county was rich or poor, more or less educated, or had a specific industrial structure 100 years ago is inconsequential for current levels of development. However, whether it managed to attract migrants and integrate them has contributed to determine subsequent economic dynamism for more than a century. Hence, there is a need to trumpet that current migration policies, aimed at restricting overall migration numbers and, in some cases, at preventing certain population groups, nationalities or religious denominations from entering a country can entail very long-term and unintended negative economic implications. By restricting migration, individual regions as well as entire countries can miss out on the massive economic potential and substantial growth-enhancing dynamics that migration brings in the long-term. Election cycles, public opinion and short-term quick-fixes should not be the only pillars policy-making decisions about migration are based on. While burning issues should most definitely not be ignored, there is need to consider what lies beyond immediate challenges and look at migration from a big-picture perspective. Shutting our doors is not the solution and will end up in significantly negative economic implications for generations to come.

9.1 REFERENCES

Florida, R. (2002): *The rise of the Creative Class: And how it's transforming work, leisure, community and everyday life*. New York: Perseus Book Group.

Jacobs, J. (1961): *The death and life of the great American cities*. New York: Random House.

Jacobs, J. (1969): *The economy of cities*. New York: Random House.

McGovern, P. (2012): Inequalities in the (de-)commodification of labour: Immigration, the nation state, and labour market stratification. *Sociology Compass*, *6*(6), 485–498.

Rodríguez-Pose, A. and V. von Berlepsch (2014): when migrants rule: The legacy of mass migration on economic development in the United States. *Annals of the Association of American Geographers*, *104*(3), 628-651.

Saxenian, A. (2006): *The new Agronauts: Regional advantage in a global economy*. Cambridge, MA: Harvard University Press.

Sequeira, S., Nunn, N., and Qian, N. (2017): Migrants and the making of America: The short- and long-run effects of immigration during the Age of Mass Migration. *National Bureau of Economic Research Working Paper Series*, No. 23289.

Silvey, R. and Lawson, V. (1999): Placing the migrant. *Annals of the Association of American Geographers, 89(*1), 121-132.