The London School of Economics

and Political Science

Doctoral Thesis

Cliometric Essays on Mexican Migration to the United States

Author:
David Escamilla-Guerrero

A thesis submitted to the Department of Economic History of the London School of Economics for the degree of Doctor of Philosophy

London, October 2019

Declaration of Authorship

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 further declare:

- The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgment is made. This thesis may not be reproduced without my prior written consent.
- I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.
- The thesis consists of 51,780 words (excluding bibliography).

Statement of conjoint work:

- I confirm that Chapter 4 was jointly co-authored with Professor Moramay López-Alonso (Rice University) and I contributed with 95% of the work.

Declaration about working paper versions of the material presented in this thesis:

- A version of Chapter 2 is available as a Working Paper (2019/173) of the Economics Department, University of Oxford.
- A version of Chapter 4 is available as a Working Paper (2019/23) of the United Nations World Institute for Development Economics Research.

Abstract

There is a lack of cliometric literature addressing the characteristics of Mexican migration during the Age of Mass migration (1850–1914). To fill this void, I analyze an original data set—the Mexican Border Crossing Records (MBCRs) publication N° A3365—to disentangle the initial mechanics of Mexican migration in the early twentieth century. I first offer a historical overview on Mexican migration to the United States in Chapter 1. In Chapter 2, I introduce these novel micro data that record individual characteristics of migrants that crossed the Mexico-US border from 1906 to 1908.

In Chapter 3, I address the initial determinants of the Mexican-American migration stream. I use the migrant's location of last residence and final destination to identify migration corridors at the local level (migration streams between Mexican municipalities and US counties). In addition, I provide a quantitative assessment of the push and pull factors that may explain differences in migration intensity across corridors. These factors include the US-Mexico wage gap, market potentials, living standards and access to railways.

In Chapter 4, I use the migrant's height—a proxy for physical productivity of labor—to quantify the selectivity of Mexican migration. In addition, I exploit the Panic of 1907 as a natural experiment of history to study the speed that migrant self-selection adjust and change to both environmental and economic factors. This financial crisis provides me with exogenous variation in height to evaluate if unexpected shocks affecting the demand of immigrant workers can induce short-run changes in migrant self-selection. To explain shifts in selection patterns, I focus on labor institutions as mechanism of adjustment. Specifically, I study the

enganche, a system of labor recruiting that neutralized mobility and job-search costs.

In Chapter 5, I exploit the reported locations of birth, last residence and destination to classify migrants based on their chosen migration method: direct or stage migration. The micro data reveal that forty percent of the migrants moved within Mexico before crossing the border. I estimate correlations between stage migration and potential wage at the destination controlling for the immigrants' age, literacy, sex, marital status and birthplace. In Chapter 6, I offer some concluding remarks.

My findings expand our knowledge about the initial patterns of Mexican migration using micro data not analyzed previously. They show that in the early twentieth century, the decision to migrate was a function of diverse forces, which effects and magnitudes varied across Mexican regions. Also, Mexican migration was characterized by an intermediate or positive selection, and labor institutions involved in the migration process shaped migrant self-selection. Finally, Mexicans used stage migration to reach the US border, and it was associated with a significant wage premium at the destination.

Acknowledgments

This thesis was the result of four years of work (2015–19), but my interest in studying economic phenomena from a long-run perspective is rooted in my undergraduate studies at the Faculty of Economics – National Autonomous University of Mexico. My first approach to economic history was the lecture of Enrique Rajchenberg: *Historia Económica de México*. His structured teaching method and provocative hypotheses sparked my interest in economic research. Thanks to his support, I did a one-year academic exchange at Science Po and an academic visit at The Center for American History – University of Texas. These experiences allowed me to appreciate the relevance of economic history in economics and pushed me to look for additional intellectual challenges abroad. The supervision and guidance of Enrique set the basis for my academic career.

I had my first contact with cliometrics in 2010, during my Master's Degree at Carlos III University of Madrid. I was very fortunate to receive lectures from first-class economic historians such as Leandro Prados de la Escosura, Carlos Alvarez Nogal and Joan Rosés. Especially, Joan transmitted to me his passion for research combining econometrics, economic theory and history. Since then, I have received extraordinary academic advice and supervision from him, which has shaped my understanding of economics. After my studies in Madrid, I wanted to expand my understanding of economic theory. For that reason, I studied a Master's Degree in Economics at the University of Essex. The lectures of Melvin Coles and Rossella Argenziano equipped me with analytical skills that have helped me to develop my research. I have no question that my time in Essex

gave me the possibility to pursue my doctoral studies at the London School of Economics.

I could not have written this thesis without the supervision of Eric Schneider and Joan Rosés, my primary and secondary supervisors, respectively. During the past four years, they have challenged my ideas and corrected my mistakes. They have provided me with extraordinary insights about my research, and they have supported me in indescribable ways. The supervision of Eric Schneider has been academically invaluable. I have no doubt that his comments made me a better economist in every front. His commitment with my academic development and personal wellbeing is something that I cannot overemphasize more. Joan and Eric have been more than my supervisors, they became my mentors in academic and non-academic aspects. In addition, Chris Minns has been an extraordinary academic advisor during my studies at the London School of Economics.

The interaction with my colleagues of the Department of Economic History enlightened this thesis as well. I have obtained unique ideas from academic discussions and informal talks with my colleagues. More importantly, they always made me feel at home. Especially, I am grateful to Mattia Bertazzini for his friendship and comments on my papers. In addition, Maanik, Greta, Roger, Iván, Enrique, Juliana, Jessie, Jenny, Oliver and Sam made my four years in London unforgettable.

I have also received invaluable support from Beatriz Leycegui and Enrique Minor. They have supported my dream of studying abroad since I sent my application to the doctoral program. They have always cared about my academic, professional and personal development. For this, I am in debt with them. Similarly, the unconditional friendship of Javier, Clemente, Víctor, Lui, Rubén, José Juan, Luis Felipe, Rafa, Rodrigo, Ulises, Jessica, Diana, Karina and Jorge was fundamental to conclude my doctoral studies. They had supportive words that kept me going in difficult times.

I want to especially thank my family. They walked by my side during all these years. Without their inexhaustible love, I would not have finished my doctoral studies. My siblings, Leslie and Marcos, and my parents, Magdalena and Juan, have always provided me with everything they can. They supported my research and studies in Mexico and Europe in innumerable ways. There are no words to express my gratitude to them. I am very grateful to my aunt Mercedes, who financed part of my studies in the United Kingdom. In addition, my cousins Héctor and Carlos as well as my aunt Rosa have supported me during all this time. I deeply thank Daniela for being comprehensive and loving during these years. She always helped me to overcome every adversity during this academic journey. Moreover, she has accepted with endless patience my obsession and passion for economics.

I wish to warmly thank my examiners, Tim Hatton (Essex University) and Leah Platt Boustan (Princeton University), for agreeing to read and evaluate my work. Over the years, my research has benefited from presentations in diverse seminars, conferences and workshops. I received extraordinary comments from Ran Abramitzky, Zach Ward, Neil Cummins, Gregory Clark, Larry Neal, Geoff Clarke, Jessica Bean, David Clingingsmith, Maggie Jones, Carolyn Moehling, Jean-Laurent Rosenthal, Melissa Thomasson, Stephen Broadberry, Jane Humphries, among other scholars that attended the annual conferences of the Economic History Society (2018 and 2019), the Economic History Association (2018) and the Cliometric Society (2018 and 2019). I also received valuable insights at the Labor in History & Economics Conference (2019) and the 2018 World Economic History Congress. Additionally, I benefited from my academic visit at El Colegio de México in 2016. I am especially grateful to Dr. Fernando Pérez for his mentorship during my fellowship at The Bank of Mexico in the summer of 2018, and to Dr. Miguel Niño-Zarazúa for his mentorship during my fellowship at UNU-WIDER in the fall of 2018.

Finally, this research was developed with the financial support of: the Mexican National Council for Science and Technology (2015–18) – Scholarship N° 409165; the Mexican Ministry of Education Scholarship (2015–16); the Radwan Travel and Discovery Fund (2016) – London School of Economics; the Pre-Dissertation Exploratory Grant (2017) – Economic History Association (US); the Research Fund for Graduate Students (2017) – Economic History Society (UK); and the PhD Final Year Scholarship (2018–19) – London School of Economics.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Between 1850 and 1914 about 55 million people emigrated from Europe to destinations in North and South America and Australasia. This historical episode is known as the Age of Mass Migration and has motivated a large body of literature traditionally focused on migration to the United States, because this country absorbed about 60% of the total outflows (Hatton & Williamson, 1998, p. 3). Mexicans started to emigrate in large numbers to the United States by the turn of the twentieth century (Gratton & Merchant, 2015, p. 521). However, little has been written about the characteristics of the Mexico-United States migration during the Age of Mass Migration.

This thesis presents four essays that intend to fill this gap in the literature. These essays can be described as *cliometric* in the sense that they use concepts and approaches of applied economics to investigate a historical issue (Hatton, 2010, p. 941). To my knowledge, they constitute the first research that exploits micro data to address the initial mechanics of Mexican migration. As in most cliometric literature on international migration, the topics that are examined have clear parallels to contemporary debates, providing insights on relevant issues concerning migration today (Abramitzky & Boustan, 2017, p. 1312). Precisely,—after introducing the core data in Chapter 2—Chapter 3 addresses the initial drivers of Mexican emigration at the local level; Chapter 4 estimates the self-selection of Mexican migrants and disentangles the role of labor institutions in shaping migrant selection; and Chapter 5 evaluates if migration methods (direct

or stage migration) can explain differences in the immigrant's labor-market performance.

This introduction has three objectives. First, I introduce the Mexico-United States migration by outlining some of the literature on the topic. I pay special attention to research addressing Mexican migration from a historical perspective, since they constitute relevant references to my research. Second, I review key literature on the Age of Mass Migration to provide the reader with a benchmark for the findings of this thesis. The literature review—on the Age of Mass Migration—is organized according to the topics examined in each chapter. It is important to mention that this review omits relevant topics in the economics of migration literature—such as the effects of immigration in source and destination countries, and the political economy of immigration policy—because they are not directly related to the issues investigated in this thesis. Finally, I summarize the contributions of the thesis.

1.1 Mexican migration to the United States

As a consequence of the Mexican-American War (1846–48), Mexico lost the territories of California, Nevada, Texas, Utah, and most of Arizona, Colorado and New Mexico. Henderson (2011, p. 9) estimates that this conflict produced an initial stock of about 80 to 100 thousand Mexican immigrants in the United States. Since then, the characteristics of the Mexican-American migration flow have changed over time. Based on the immigrants' profile, migration methods, incentives to migrate and institutional environments, previous scholarship has identified diverse migration patterns, which can be understood as migration models followed during specific periods of time. The following periodization captures the migration patterns that have existed since the end of the nineteenth century and helps to identify key literature on each period.[1]

[1] Cardoso (1980), Durand (2016), González & Fernandez (2002), Henderson (2011) and Verduzco (1995) offer alternative time periods to study migration patterns.

1.1.1 Beginnings (1884–1910)

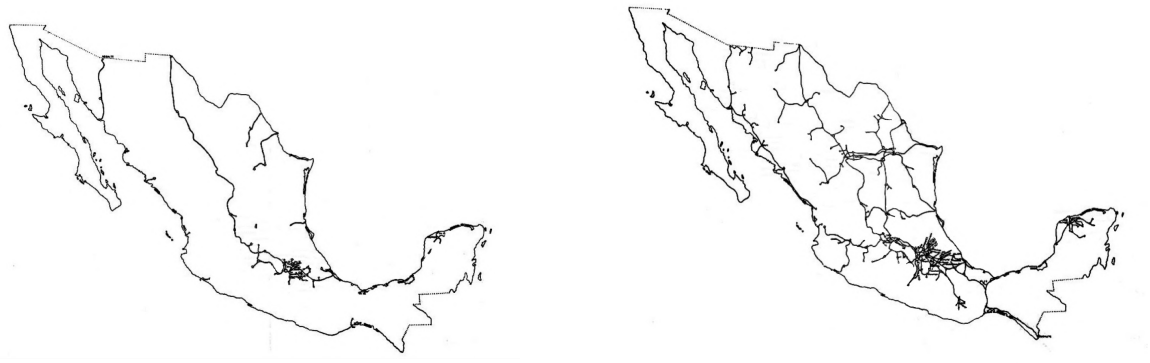
Estimations of out and return migration are scarce and imprecise for this period because neither Mexico or the United States kept systematically statistics of Mexican migration until 1906. However, previous literature agrees that Mexican emigration to the United States became notable only from 1884, when the Central Mexican Railway connected Mexico City with the American railway network at El Paso, Texas. At that time, the Sonora Railway also connected the sea port of Guaymas—in the Gulf of California—with Nogales, Arizona. By 1900, two additional rail lines from central Mexico were connected at Eagle Pass, Laredo and Brownsville, making trains the fastest transportation towards the United States. Figure 1.1 depicts the expansion of the Mexican railway network from 1884 to 1910. Yet, it is not clear if railways *per se* fostered mass emigration (González & Fernandez, 2002, p. 43), or if they were affordable for the Mexican working class (Coatsworth, 1979, p 940).

Besides the expansion of the railway network, two additional factors characterize this period: the absence of immigration restrictions for Mexicans and the *enganche*.[2] First, during this period Mexicans faced an open-door policy in both countries, which facilitated even more emigration before 1910 (Cardoso, 1980; Durand, 2016; Gamio, 1930). Mexicans were not considered immigrants who sought to settle permanently, but temporary workers that moved back and forth supplying labor without restrictions. Figure 1.2 shows Mexican-born and Mexican-origin population in the United States. From 1900 to 1910, Mexican-born population increased two-fold. This happened due to Mexican mass migration that created a Mexican-American Southwest in the early twentieth century (Gratton & Merchant, 2015).

Second, the *enganche* was a system of labor recruiting that was institutionalized in Mexico since the second half of the nineteenth century. It consisted in recruiting

[2]In fact, Durand (2016) calls this period *The Enganche Era*.

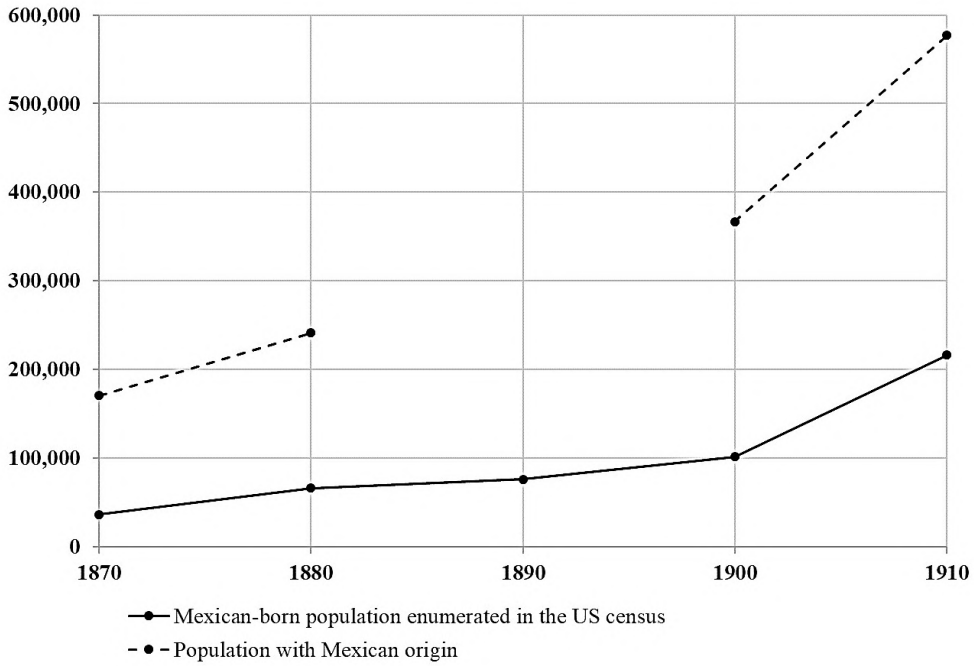
Figure 1.1: Expansion of the Mexican railway network (1884–1910)



(a) 1884 *(b) 1910*

Source: Cosío Villegas (1974).
Note: Most of the current south-north Mexican railways were constructed during the dictatorship of Porfirio Díaz, known as the Porfiriato (1877–1911). The mileage increased from 477 km in 1877 to 19,000 km in 1910 (Cosío Villegas & Bernal, 1973; Henderson, 2011).

Figure 1.2: Mexican population in the United States, 1870–1910 (Number of persons)



Source: 1. Census data from Social Explorer Dataset: 1890, 1880, 1890, 1900 and 1910 Census. Digitally transcribed by the Inter-university Consortium for Political and Social Research. Edited and verified by Michael Haines. Compiled, edited and verified by Social Explorer. Note: The 1910 Census reported the birthplace of white foreign born population.
2. Figures of population with Mexican origin from Gratton & Merchant (2015, p. 524-5). Note: The authors do not provide data for 1890.

workers in highly populated regions and transport them to remote places or with labor shortage. As part of this recruiting system, workers were paid in advance in exchange for future work at the destination, which created an indebtedness relationship with the labor contractor (Brass, 1990, p. 74). Durand (2016, p. 53–4) documents that recruiters arrived systematically to regions in central and west Mexico to recruit and transport workers to the southeast of the country.

At the beginning of the twentieth century, American companies and labor contractors adopted the *enganche* to transport and allocate labor across the southwestern United States and other regions. This was possible due to the three rail lines built between 1884 and 1900, which became the major means of transporting recruiters south into Mexico and transporting large numbers of workers north to the United States (Woodruff & Zenteno, 2007, p. 512). The increasing demand of Mexican workers sparked the creation of recruitment agencies at the border, being El Paso (Texas) the only real labor depot at the time. Other cities such as Los Angeles (California), San Antonio (Texas) and Tucson (Arizona) became distributing points of Mexican labor as well (Clark, 1908, p. 475). Although the *enganche* eliminated transportation and job-search costs for migrants, it was also characterized by the breach of contracts (changes in agreed work locations, labor tasks and wages) once the workers arrived to the United States (Durand, 2016, p. 61).

The literature addressing the characteristics of Mexican emigration before 1910 is scarce—compared to the body of literature covering subsequent periods—and can be defined mainly as qualitative or historical. The principal reference for the beginnings of the migration flow is Clark (1908). Based on observations and interviews mainly at El Paso, Eagle Pass and Laredo in 1906–07, his research concludes that most immigrants were unskilled laborers that came from the central plateau of Mexico. He argues that Mexican immigrants were seasonal laborers that moved first to employment centers such as El Paso or Los Angeles from

where they were distributed. He also describes the Mexican immigrant as fairly efficient and documents the increasing displacement of Japanese, Greeks and Italians by Mexican laborers. His analysis of wages for Mexican immigrants across industries and occupations suggests that the US–Mexico wage gap may have been the main incentive to emigrate. His research also approaches indirectly—through individual experiences or stories—the performance of Mexican immigrants in the US labor market. Similarly, González (2010) uses American and Mexican newspapers to describe the factors that influenced Mexican emigration before 1910, but focuses on documenting the labor discrimination and mistreatment that Mexican immigrants experienced in the United States.

Cardoso (1980, p. 12 & 18–37) collects daily wages available to Mexican immigrants (1900-10) from miscellaneous sources and argues that the US-Mexico wage gap was the main driver of the migration flow. In addition, he outlines other factors that influenced Mexican emigration such as the economic expansion of the American Southwest and regional droughts in Mexico. He also documents the increasing presence of Mexican immigrants across economic sectors and regions of the United States during the first decade of the twentieth century.[3] In this sense, Gratton & Merchant (2015, p. 521)—probably the only research using a structured quantitative analysis—use census data to show that Mexican mass migration to the United States started precisely from 1900, creating a Mexican American Southwest.[4] Therefore, the growing stock of Mexican immigrants in the US may have influenced emigration as well.

The reports of the US Immigration Commission (1911a) or Dillingham Commission constitute probably the most important source of data for the period. They present secular trends of gross Mexican emigration from 1820 to 1910. For some fiscal years, they also report the distribution of Mexican immigrants across

[3] Cardoso (1980, p. 35) highlights that the stock of Mexican citizens in the United States increased two-fold between 1900 and 1910.

[4] Similarly, González & Fernandez (2002, p. 43) suggests that Mexican labor began to enter the United States in sizable numbers after 1905 (see Figure 1.3).

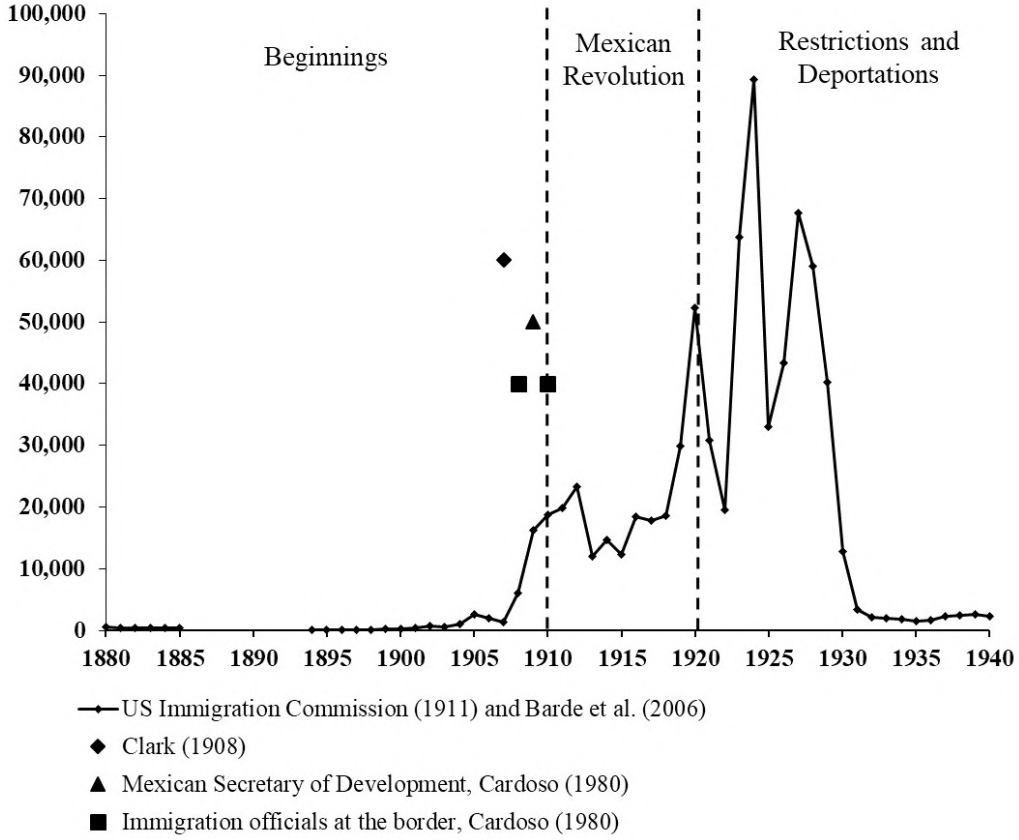
occupations; wages for male and female Mexican immigrants across US regions and economic activities; and figures of return migration. To my knowledge, these rich data has not been used to develop quantitative research. The exception is Feliciano (2001), who uses it to study the long-run assimilation of Mexican immigrants from 1910 to 1990. In general, literature studying Mexican emigration before 1910 uses scattered data of gross emigration—observed at specific locations and periods of time—to estimate the size of the migration flow. Figure 1.3 plots some of these estimates and contrasts them with the figures of the Dillingham Commission. The estimates of Clark (1908) and Cardoso (1980) are considerably larger than the official figures. However, it is difficult to assess which of them are more precise since the size and desert nature of the border complicated the registration of Mexican immigrants (Cardoso, 1980, p. 34).

In Chapter 2, I introduce an unexploited source of data for this period, which provides information on the size and characteristics of Mexican migration: the Mexican Border Crossing Records (MBCRs). Specifically, I analyze the publication N° A3365 and present evidence suggesting that some of the migration patterns described by previous literature may be imprecise. I argue that this novel data is representative for the period and it may capture 81% of the gross emigration flow. The micro data also show that return migration may have been about 6% of the gross flow. In Chapter 3, I use these data to test if Mexican emigration—during this period—was determined by the Mexico-US wage gap and to estimate the causal effect of the access to railways on emigration. The findings offer an alternative narrative to previous literature.

1.1.2 Revolution and War (1910–20)

It is important to mention that most literature considers the period from 1910 to 1920 as part of the beginnings of the Mexico-United States migration. However, the second decade of the twentieth century was characterized by two events that

Figure 1.3: Gross Mexican emigration, 1880–1940
(Number of immigrants)



Source: Author's calculations based on Barde et al. (2006); Clark (1908); Cardoso (1980); and US Immigration Commission (1911a).
Note: Estimations of Clark (1908) are based on back-of-the-envelope estimations of third-class passengers at El Paso and Eagle Pass from August 1906 to August 1907. He suggests figures between 50 and 60 thousand immigrants per year.

might have impacted Mexican emigration: the Mexican Revolution (1910–20) and a guest worker program—established by the United States to offset the effects of World War I on labor supply (Woodruff & Zenteno, 2007, p. 512).[5] The former represents a shock from the supply side that may have increased the emigration of individuals escaping from the Mexican civil war (refugees) (Gratton & Merchant, 2015, p. 528-9), and the latter is a policy-driven shock from the demand side. Although it is likely that these events may have modified migration patterns

[5]See Knight (1986) for a review of the Mexican Revolution.

existing before 1910,[6] most literature assumes that migration patterns remained unchanged during this period.[7]

In addition to violence, the Mexican Revolution impacted the population living standards through epidemic disease, starvation and inflation (Henderson, 2011, p. 24)—that is, incentives to emigrate or push factors. However, Figure 1.3 shows that the Mexican Revolution may have slowed gross emigration in the years of more intense fight (González & Fernandez, 2002, p. 43). This may be explained by the regular suspension of railway services, affecting the northward transportation of passengers (Hardy, 1934, p. 252–60). The volume of Mexican emigration regained strength in 1918 (see Figure 1.3), when the violence in Mexico was winding down and the shortage of labor in the United States increased due to World War I.[8] Therefore, the Mexican Revolution added to the emigration flow, but was not primarily responsible for rising emigration (Gutmann et al., 2000, p. 147; Cardoso, 1980, p. 53). Cardoso (1980, p. 53) estimates that this juncture induced 2 million border crossings during the entire decade.

To my knowledge, there is no quantitative research addressing the impact of the Mexican Revolution on emigration to the United States. I believe that this period represents an unique opportunity to study how wars can shape migration patterns in the short and long run. The MBCRs along with the US population censuses may be the best data sources to study this period.

1.1.3 Restrictions and Deportations (1921–41)

The Immigration Act of 1917—which required all immigrants to pass a literacy test and to pay an eight dollar head tax—was the first restriction imposed to Mexican emigration (Kosack & Ward, 2014, p. 1015). However, these restrictions

[6] Hatton (2014, p. 46) argues that immigration policies can influence the scale and composition of migration.
[7] The exception is Cardoso (1980), who considers the 1910–20 period a category in itself to study migration patterns.
[8] About one million US citizens conscripted in the military in 1918 (Henderson, 2011, p. 25).

were waived for Mexicans during wartime, and American employers persuaded the government to extend the exemption beyond the war's end. The open-door policy for Mexican emigration ended effectively in 1921, when the US Secretary of Labor ended the waivers (Cardoso, 1980, p. 98). The next year, a ten dollar visa fee was added to the eight dollar head tax already charged to each immigrant, making emigration through official entrance ports very expensive for the average Mexican worker (Henderson, 2011, p. 35). Therefore, this immigration restrictions set—for the first time in history—the incentives for illegal emigration.

The change in immigration policy was due the economic recession of 1920. The end of the economic expansion in the United States—induced by World War I—decreased the demand for labor across sectors and produced high unemployment rates. The economic recession had an important impact on the mining and railway industries, which employed a high percentage of the Mexican immigrants. Cardoso (1980, p. 97) suggests that about 21% of the legal Mexican immigrants lost their job at the beginning of the 1920s. In this sense, Mexican immigrants became an oversupply of labor that produced an anti-immigrant sentiment. For example, the American Federation of Labor (AFL) expressed that Mexican immigrants were a major menace to American workers because their willingness to work for low pay pushed down wages for all workers (Henderson, 2011, p. 38–9). Some politicians urged the US government to conduct deportations to rid the country of unemployed indigent Mexicans. At the same time, Mexican immigrants approached local consular officials asking for assistance. This situation derived in a repatriation program implemented by the Mexican government costing about 1 million dollars (Cardoso, 1980, p. 98–112).

Figure 1.3 shows that emigration regained momentum after 1922, but with an increasing participation of unauthorized emigration. To reduce illegal emigration, in 1924 the Mexican government established migration offices in key rail stations with the objective that no tickets would be sold to emigrants without a valid

work contract. Also, migration agents rode on all northbound trains to remove and send back individuals thought to be emigrants without proper documents (Cardoso, 1980, p. 111–12 & 130). From the American side, the Immigration Act of 1924 created the Border Patrol to control unauthorized immigration from Mexico. Durand (2016) documents that only in that year, the Border Patrol made 4,038 deportations. Despite the implemented restrictions, Mexicans continued emigrating in large numbers until the onset of the Great Depression. During the 1930s, the anti-immigrant sentiment increased and the US government carried out raids and massive deportations.[9] Verduzco (1995, p. 576) argues that in the period 1929–32, the US government deported 345 thousand Mexican immigrants. Moreover, gross emigration from 1931 to 1940 accounted for only 4% of the figure estimated for the period 1921–30.

Apart from historical literature, there are three investigations that provide a comprehensive quantitative analysis about the characteristics of Mexican emigration during this period. First, Gamio (1930) exploits individual postal money orders (remittances) sent by Mexican immigrants—from 1926 to 1927—to assess the composition of the migration flow: the immigrants' origin and destination. To study the performance of Mexican immigrants in the US labor market, Gamio (1930) collects data at the state level on average daily wages in Mexico and the United States across labor groups and occupations. In addition, he uses data obtained from personal contact with immigrants to portray the cultural background, religion and social mobility of Mexican immigrants. His results suggest that Mexican emigration had a regional and transient character—that is, most immigrants came from the central plateau of Mexico and migrated only for some seasons. In addition, most Mexican immigrants begin working as unskilled workers and received wages much lower than the American laborers of the same class (Gamio, 1930, p. 46–7).

[9] In addition, in March 1929 the US Congress passed legislation that made illegal entry a misdemeanor punishable by a one thousand dollar fine or up to a year in prison (Henderson, 2011, p. 43).

Second, the research of Paul S. Taylor consists of studies at locations in California, Colorado, Pennsylvania, Illinois and Texas. For example, Taylor's (1933) research at the Imperial Valley (southern California) is based on field observations carried out in 1927. The study presents tabulations of Mexican immigrants by state of birth and marital status to address the composition Mexican labor and population in the region. His results reveal that only 33.5% of the immigrants came from the central plateau, suggesting the existence of different migration patterns at the time. In addition, he presents the distribution of Mexican children in elementary schools at the district level using density dot maps. To study the region's labor market dynamics, he collected data on daily wages and characteristics of the immigrant labor by sector or crop. Taylor (1933) also analyzes land ownership—based on surnames—in each district to identify the position of Mexican immigrants in the socioeconomic ladder. A similar structure and approach is followed in the other studies. The last volume of Taylor's (1933) research presents statistics that complement the case studies. They include monthly figures of gross emigration, which he uses to analyze the volume and fluctuations of seasonal emigration. An outstanding contribution are the detailed maps showing the destination of Mexican repatriates during the Great Depression years.

In many dimensions, the investigations of Gamio (1930) and Taylor (1933) are very similar to modern cliometric research. They use novel data to analyze the composition variation of Mexican emigration over time and/or across space. Their research shows clear awareness of the factors influencing the emigration decision and the immigrant assimilation into the labor market. They also address cultural and institutional issues using qualitative methods. Their major contribution was the collection and analysis of unique data.[10]

Third, Kosack & Ward (2014) use height—reported in the MBCRs—to proxy migrant quality and estimate the self-selection of Mexican migrants into and out of the United States in the 1920s. Their empirical strategy estimates differences in

[10]Unfortunately, the raw data collected by Gamio (1930) or Taylor (1933) is lost or not available.

height between migrants and Mexican soldiers controlling for diverse factors that may influence an individual's height. The findings suggest that Mexican migrants were positively self-selected. In addition, they link their migrant sample to 1930 US and Mexican census to obtain samples of permanent and return migrants. They argue that return migrants were neutrally selected relative to permanent migrants. To my knowledge, this is the only research that exploits the MBCRs to study Mexican emigration in a historical perspective.

Overall, this period shows that Mexican emigration was very responsive to economic crises. In Chapter 4, I exploit the financial crisis of 1907 to disentangle the effect of unexpected shocks on migrant self-selection. The findings suggest that the selection patterns of Mexican migration changed as a result of this crisis. Therefore, I speculate that some dimensions of the migration stream could have changed as well from 1921 to 1941.

1.1.4 Bracero Program (1942–64)

In 1942, the Mexican and American governments signed a contract-labor program unprecedented in the history of both countries: the *Bracero Program*. Initially, it was an emergency measure to satisfy labor shortages in the US created by World War II. The program—an American initiative—was characterized by the large-scale sustained recruitment of temporary agricultural Mexican workers under a series of international agreements (Durand, 2007; García y Griego, 1983; Samora, 1982). As part of the agreement, Mexican migrants were to be paid the same wage as their American counterparts, so that migrants would not reduce wages in any activity or location (Chacón, 2009, p. 522).[11]

This has been the only episode in history in which Mexican emigration was—to some extent—regulated, and immigrants were guaranteed with decent living

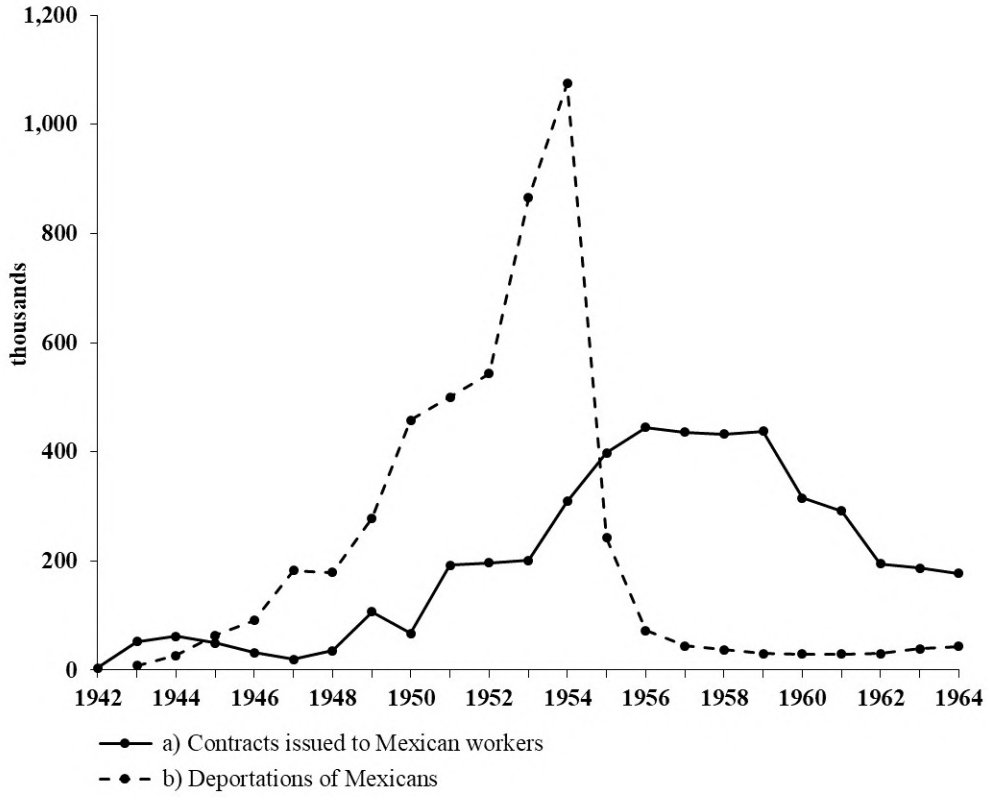
[11] The program included additional conditions such as: Mexican workers would not be enrolled in any American military service; Mexican workers would not be discriminated in any sense; and Mexican workers would obtain round transportation, housing and social security.

and working conditions. The program also helped the Mexican government to regulate the geographic composition of the bracero stream through state-level quotas. Nevertheless, unauthorized emigration continued, and the inability of both governments to control it generated political tensions. From 1950, the US government implemented diverse measures to control unauthorized migration, being Operation Wetback in 1954—a military-style expulsion campaign (Henderson, 2011, p. 84)—the most important. Figure 1.4 shows that effect of this operation. After 1954, the number of deportations (unauthorized immigration) dropped precipitously, reaching levels below the number of issued contracts (legal immigration)—a situation not experienced since the last years of World War II. In the following years, little by little the protections of the contract-labor program were dropped and employers started to use undocumented workers again (García y Griego, 1983, p. 49–67). The constant frictions between governments induced the ending of the program in 1964 (Verduzco, 1995, p. 577).

This period has motivated quantitative literature, however it remains relatively scarce. One data source for the bracero era is the Mexican Migration Project (MPP). It is a retrospective survey that reports under which legal status the migration spell occurred—that is, it identifies individuals that emigrated under the *Bracero Program* and those who emigrated illegally (unauthorized immigrants). The latest update of the MMP contains data from 170 Mexican communities and covers retrospectively the period 1915–2018.[12] Using the MPP, Massey & Liang (1989) find that migrating during the bracero period increases the likelihood to make repeated trips. Bracero migrants were also more likely to introduce their children to migration, who eventually settle permanently in the United States. Their findings shows that during this period there was an

[12]Since 1982, three to five communities across Mexican regions are randomly surveyed each year. The selection criteria is based on the existence of some migration in the community. The sample size is generally 200 households and communities with different characteristics have been chosen to provide a range of diverse sizes, regions, ethnic compositions, and economic activities.

Figure 1.4: Gross Mexican emigration, 1942–1964
(Number of immigrants)



Source: a) US Department of Labor (1965). b) INS Annual reports obtained from Samora (1982, p. 46).
Note: Prior to 1960, the figures refer to deportations. Afterwards, they refer to located deportable Mexicans. The number of contracts issued capture legal immigration, whereas the number of deportations proxy indirectly the size of illegal immigration.

intergenerational transmission of migration. Kosack (2016) shows that individuals that migrated under the Bracero Program were more likely to start a new business relative to those that made illegal trips. Hence, the program had a significant impact on economic growth and development by spurring new investment at the community-level.

In addition, Kosack (2019) uses census data and the location of the bracero recruitment centers to estimate the impact of the program on human capital investments. His results suggests that bracero migration increased primary school enrollments and the spending on education by state governments. Clemens et al. (2018) disentangle the impact of ending the program. They exploit archival data

on the geographic locations of bracero workers in the United States to estimate the impact of the bracero exclusion. Their findings provide evidence that the exclusion of braceros had little effect on wages or employment of American farm workers.

In sum, this immigration policy modified in several ways the existing migration patterns, and its end revealed the unanticipated consequences of the program: unauthorized circular migration (Durand, 2016). In Chapter 4, I show that the recruitment of Mexican laborers existed since the early twentieth century, but it was carried out by private agents. I also provide evidence suggesting that this recruitment system—the *enganche*—served as an adjustment mechanism of migrant self-selection. Therefore, private and public institutions have been shaping the Mexico-US migration for more than 100 years.

1.1.5 Unauthorized Immigration and Reforms (1965–onwards)

The decades following the Bracero Program program were characterized by the absence of an immigration policy from both governments. This strengthened unauthorized immigration and modified importantly migration patterns: women started to emigrate and the geographic composition of the flow diversified. Moreover, new immigrants came from both rural and urban Mexico (Durand, 2016; Henderson, 2011). By the onset of the 1980s, the American public debate was once again decrying a crisis over illegal immigration. However, new actors of the American society—ethnic groups, human rights defenders, academics and politicians—have embraced free market fundamentalism and agreed that legal immigration and diversity was a good thing. This movement influenced public opinion towards an amnesty program for illegal immigrants: the Immigration Reform and Control Act (IRCA) of 1986. This legislation allowed undocumented immigrants to apply for adjustment of status (Henderson, 2011, p. 114–5), which empowered and consolidated the Mexican and Hispanic-Latino communities.

Over time, the amnesty led to an immigrant naturalization process, and the circular character of Mexican migration changed to a permanent settlement of legal and undocumented immigrants. The permanent settlement of undocumented immigrants has created generations of individuals who were brought to the United States at an early age without documentation. These immigrants—known as dreamers—have assimilated the US culture and have been educated by US school systems. However, they have not been offered a permanent solution to their undefined immigration status.[13] In addition, from 2010 Mexican net migration has fallen to zero (Passel et al., 2012), and recent findings suggest that—for the first time—Mexican unauthorized immigrants in the United States no longer account for the majority of those living in the country illegally (Passel & Cohn, 2019). Recently, arrivals at the US-Mexico border of migrants and asylum seekers—families and unaccompanied children—from Central America have increased notoriously. This could change in the short-run the role of Mexico from a migrant-sending country to a destination country, which in turn may modify the Mexico-US migration patterns (O'Connor et al., 2019).

From an academic perspective, this period concentrates most literature on Mexican migration. The reason is the creation of data sources to study this phenomenon. Some of the most relevant data sources are the following. The Survey of Migration at Mexico's Northern Border (EMIF Norte) has been implemented since 1993 at the main entrance ports of the Mexico-US border and in some important Mexican airports. It captures the socioeconomic characteristics of migrants and their migratory experience in the United States. The Mexican Family Life Survey (MxFLS) contains information for a 10-year period, collected in three rounds: 2002, 2005-2006 and 2009-2012. It is a longitudinal multi-thematic survey representative of the Mexican population at the national, urban, rural

[13]The Deferred Action for Childhood Arrivals (DACA) Program was announced in 2012. The program allows certain young individuals who came to the United States as children, without documentation or who have overstayed their valid documentation, an opportunity to request work authorization and to defer their removal from the country for a period of two years.

and regional level. This survey captures migration impacts on variables as self-employment, farm work, family production, labor income, private transfers, household consumption, among others. Since the mid 1980's the characteristics of the household receiving remittances can be identified with the National Survey of Household Income and Expenditure (ENIGH). However, this survey is only representative at a national level and for some states in specific years. The 2000 Population Census also contains a migration supplement with a municipality representativeness, but to my knowledge, it was not implemented again. In the United States, the Population Censuses, the American Community Survey and Current Population Survey are the main sources to identify the characteristics of Mexican households and foreign-born individuals. Although it provides data on specific communities, the Mexican Migration Project (MPP) may be the ideal data source to study the evolution of migration patters during this period.

The previous data sources have spurred inexhaustible literature addressing drivers and effects of migration that are not normally possible to study with data from other periods. For example, Nawrotzki & DeWaard (2016) argue that climate shocks can induce emigration and the risk of migration starts to decline three years after the shock. Similarly, Nawrotzki et al. (2015) find that most of the undocumented migration comes from rural Mexican communities, where most of the climate shocks occur. Thus, social programs aiming to smooth climate changes might be an efficient migratory policy instrument.

The impact of migration on health has also captured the attention of scholars. Ortmeyer & Quinn (2015) argue that the impact of migration on health depends on the duration of the trip, and since difficult crossings need a recovery period, short repeated migration spells have a harsher effect. In addition, Wilson et al. (2014) show that the migrant's risk for developing chronic health conditions depends on the access to health services and working conditions. Ullmann (2012) develops a cohort analysis and finds that wives of migrants are more likely to

have heart diseases and overweight than their non-migrant counterparts. Brown et al. (2015) show that, in general, Mexican immigrants in the United States face declines in health.

Other literature has addressed the links between migration and crime. Interestingly, Mexican immigration is associated with a decline in property crimes in the United States (Chalfin, 2015). Also, crime faced by migrants might alter their remitting behavior. For instance, robbery rates reduce the size and incidence of remittances, negatively impacting the well-being of those households left behind (Coon, 2015).

Clearly, the forces behind sending remittances has been a popular topic due to the importance of these economic transfers.[14] Basu & Bang (2013) find that the incentive behind sending remittances is the immigrants' desire to insure against uncertainty rather than altruistic motives. In contrast, Roberts & Morris (2003) reject that altruism, insurance, and investment are the only motives for remitting. They propose that remittances are an instrument to access the labor market at a community level, being an "entrance-right" if immigrants want to have job opportunities in the United States. Similarly, Massey & Basem (1992) find that the propensity to remit is determined by indicators of social capital and community membership. Other factors influencing the remitting behavior are the migration costs and the duration of the migration spell (Ortmeyer & Quinn, 2012). Also, the financial development of the receiving country and the immigrant's legal status can influence the volume and use of remittances (Coon, 2014; Quinn, 2005; Schluter & Wahba, 2009).

In addition, the selectivity of Mexican migration has been studied. Chiquiar & Hanson (2005) estimate the self-selection of Mexican migrants using census data. They argue that Mexican migrants were positively selected in terms of schooling: while much less educated than US natives, Mexican immigrants

[14] In 2018, Mexico was the third recipient of remittances in the world (World Bank, 2019, p 14).

were more educated than the average resident in Mexico. In addition, under a common price for observable skills, Mexican immigrants would be concentrated in the middle of Mexico's wage distribution. McKenzie & Rapoport (2010) argue that selection patterns could be influenced by the quality of migrant networks. They show that Mexican immigrants coming from communities with strong migrant networks were negatively self-selected whereas those with access to weak networks were positively self-selected. Orrenius & Zavodny (2005) also argue that immigration restrictions can modify migrant selection patterns. They provide evidence suggesting that stricter border enforcements are associated with higher immigrant skill levels.

From the previous review, we can conclude that the Mexico-US migration has evolved over the last century. However, our knowledge about its characteristics regards mainly to recent periods of time. This is not the case for other historical migration streams to the United States. English, Italian, Irish, Scandinavian and Spanish migration—during the nineteenth and early twentieth century—have been studied using large data sets, linking-data techniques, and explicit theoretical models or econometric methods. This thesis present—for the first time—equivalent research for the Mexican migration flow during the Age of Mass Migration (1850–1914). The next section reviews some literature addressing the characteristics of migration during this period.

1.2 Migration during the Age of Mass Migration (1850–1914)

The Age of Mass Migration (1850–1914) marked a profound change in the distribution of global population (labor) and economic activity (Hatton & Williamson, 1998, p. 3). The mass emigration from Europe to the New World was influenced by a unique set of conditions: improvements in transportation technologies—the shift from sail to steam—that reduced trans-Atlantic travel costs; the absence of restrictive immigration policies; and notorious asymmetries between source and

destination countries arising from industrialization and structural transformation processes (Abramitzky & Boustan, 2017, p. 1314; Hatton & Williamson, 1994, p. 544). This period of free mass migration has motivated literature addressing key topics in migration economics that are relevant today. Therefore, the findings of this literature can inform current debates on immigration issues. The following paragraphs review influential research in three topics: drivers of emigration, migrant selection and the immigrant assimilation into the labor market.

1.2.1 Drivers of emigration

Traditionally, the objective of migration economics has been to understand what motivates migrants or which forces spark migration streams. Literature on this topic studies the push (incentives operating at home) and pull factors (incentives operating at the destination) that may explain fluctuations in migration intensity over time and across migration streams. This research has evolved from models considering business cycles as main driver to models abstracting the emigration decision—that is, derived from utility maximizing microfoundations. The main reference of this type of research is the extensive work of Tim Hatton and Jeffrey Williamson. They use a time series analysis to examine the determinants of emigration from Ireland (Hatton & Williamson, 1993), the United Kingdom (Hatton, 1995a) and Scandinavia (Hatton, 1995b) from 1870 to 1913. Also, they follow a similar approach to assess the drivers of emigration across eleven European countries over the period 1850 to 1913 (Hatton & Williamson, 1994). This series of papers—first in their class—form part of Hatton and Williamson's influential book addressing the causes and economic impact of this unparalleled transfer of population (Hatton & Williamson, 1998). All this research acknowledges that employment rates in the destination and home country, stock of previous migrants and demographic variables matter in the migration decision. However, their main finding is that emigration responded systematically to real wage gaps between home and abroad.

Larsen (1982, p. 115) proposes a model for Danish emigration (1870-1913), finding as main pull factors the employment in the United States and the wage-differential between these countries. In addition, she highlights the importance of population growth and age structure as push factors. For the case of Sweden, Quigley (1972, p. 121-4) also finds that emigration from 1867 to 1908 were responsive to changes in agricultural and industrial wages, but population pressures and harvest quality in Sweden induced emigration as well.

Another characteristic of this macroeconomic-type literature is that it focuses mainly on emigration from northern Europe. However, emigration from Italy has caught the attention of scholars because it became the most important migrant-sending country by the end of the nineteenth century. Although Hatton & Williamson (1998, p. 102-5) recognize that southern European migration patterns differ from the Scandinavian or Irish, their time series analysis concludes that Italian migration rates (1876-1913) were determined by much of the same forces as those underlying the emigrations from northern Europe.[15] However, recent literature suggests that during the period 1881–1921 globalization-induced agricultural-price shocks increased the propensity to emigrate (Gray et al., 2019). Sánchez-Alonso (2000a, p. 327) also argues that emigration from Spain was different since it was income constrained. Her time series analysis (1882-1914) suggests that Spanish migration rates were mostly explained by the depreciation of the Spanish currency (peseta), which increased the cost of emigrating. In addition, Boustan (2007, p. 267) studies Jewish emigration from Russia to the United States. She implements a model à la Hatton (1995a) and finds that business cycles and periodic religious violence drove the Jewish migration stream.

Aside from analyzing secular trends and variations in time series migration rates, literature on the Age of Mass Migration has proposed models to understand

[15]Italian emigration was different in the sense that Italians emigrated to a wider variety of destinations, and it was characterized by a predominance of men and high rates of return migration.

variations in emigration across regions of migrant-sending countries. In general, these approaches are more eclectic since they cannot always draw on the time-variant component of emigration. This literature recognizes that migration patterns may vary not only over time but across space at the provincial level. An example of this type of research is Hatton & Williamson (1998), which exploits the fact that Italian emigration rates in 1882 and 1912 varied widely between 69 provinces. Their cross-sectional analysis reveals that the wage gaps had a smaller effect than that implied in cross-country or time series analysis. Additionally, demographic pressures, labor force participation in agriculture, economic development and proximity to labor markets mattered for explaining emigration at the provincial level. On the contrary, emigration tradition had a small effect and literacy had no impact on emigration. Another example is Sánchez-Alonso (2000b), who implements pooled cross-section regressions of gross emigration from 49 provinces in Spain (1888-90 and 1911-13). Unlike the Italian case, improvements in literacy and income per capita at home were positively increased regional emigration rates, i.e. poverty and ignorance were the principal constraints to emigration (Sánchez-Alonso, 2000b, p. 751). For the case of Irish migration, these constraints were neutralized by migrant networks (Connor, 2019, p. 141). In contrast, Abramitzky et al. (2013, p. 2) assess the effect of parental wealth and expected inheritance on the decision to migrate. They find that present and future wealth discouraged the emigration of Norwegian men.

1.2.2 Migrant selection

From a microeconomic perspective, policy makers and scholars—in the past and present—have been interested in knowing who migrates and if migration is a selective process. In other words, how migrants compare to those choosing to stay? The observed positive, neutral or negative self-selection of migrants depend very much on the migration stream, the measure of selection used and the context analyzed (rural/urban or national/local). For example, Wegge (2002,

p. 365) quantifies the selection of German migration—specifically from Hessel-Cassel region—during 1852–57 and finds that emigrant population was positively self-selected in terms of skill, but negatively self-selected in terms of wealth (measured as land ownership). Cohn (1995, p. 394) use passenger lists to show that male emigration from England to the United States was overrepresented by laborers (landless immigrants) whereas farmers composed about 20% of the migration flow from 1836 to 1853.

Using census sources and occupational-based earnings data, Abramitzky et al. (2012, p. 1834) find evidence of negative selection of Norwegian migrants from urban areas and inconclusive results for rural migrants. Using age-heaping as a proxy for the immigrants' arithmetical capability, Mokyr & Ó Gráda (1982, p. 375) argue that during the pre-famine and famine periods (1803–46) Irish migrants were negatively self-selected. This selection pattern may have persisted during the nineteenth century since Connor (2019, p. 141) finds that sons of farmers and illiterate men were more likely to leave Ireland than their literate and skilled counterparts. Spitzer & Zimran (2018) also find that Italian migration was negatively selected at the national level from 1907 to 1925. They exploit the immigrant's height as measure of selection, because on average stature is indicative of the individual's income, cognitive ability, health and occupation skill (Bleakley et al., 2014; Borrescio-Higa et al., 2019; Deaton, 2007; Steckel, 1995). Overall, the perception is that migrants to the United States from countries of the European periphery were negatively selected (Abramitzky & Boustan, 2017; Sánchez-Alonso, 2019). In contrast, following the same approach Kosack & Ward (2014) find that—on the basis of height—Mexican migrants were positively selected in 1920.

Spitzer & Zimran (2018) also show that national-level estimates of migrant selection could mask substantial variation at the local level. In fact, Italian immigrants were positively selected at the local level and selection varied sys-

tematically within Italy, with more positive local selection from poorer provinces. Since the direction of migrant selection is closely related to the motivations of prospective migrants (Abramitzky & Boustan, 2017, p. 1322), differences in the factors inducing emigration across provinces/states could imply differences in selection patterns. Considering that the drivers of Spanish emigration were heterogeneous across provinces (Sánchez-Alonso, 2000b), we could expect differences in selection patterns within Spain during the Age of Mass Migration. Moreover, the selection patterns can change based on the country of destination. Sánchez-Alonso (2019, p. 19–20) suggests that Italian, Portuguese and Spanish immigrants moving to Latin American countries were drawn from the northern regions, which historically tended to be more literate. Therefore, on the basis of literacy, migration from the European periphery to Latin American was positively selected.

There are different factors that may influence shifts in selection patterns. Covarrubias et al. (2015) argue that migrant selection (the immigrants skill level) could be influenced by easing or tightening of the liquidity constraints to migration. In addition, an increase in GDP has a negative effect over the average skill level since it increases the participation of low-skilled workers. Immigration policies also can change the selectivity of migration. Ward (2017) argue that from 1917 to 1924 migration quotas modified the return migration decision of immigrants. Specifically, the quotas lowered the unplanned returns of those holding low-skilled jobs in the United States—that is, return migrants became less negative selected for the countries most restricted. Spitzer & Zimran (2018) evaluates the impact of the Immigration Act of 1917, which imposed a literacy test for every immigrant from Europe. They find that this requirements was associated with a an increase in positive selection of Italian migration. Similarly, Massey (2016) assesses the impact of the first immigration quota to Canadian migration. She finds that this immigration policy resulted in migrants of higher skill.

1.2.3 Immigrant assimilation and labor-market performance

This body of literature addresses the immigrant's performance in the labor market of destination. The investigations could refer to short-run outcomes, but commonly focus on long-run analysis since the process of earnings convergence between immigrants and natives is slow (Abramitzky & Boustan, 2017, p. 1326). While for some immigrants it could take generations to catch up with natives, others could experience a widening in the earnings gap. The speed of the convergence would be influenced by the wage growth or both immigrants and natives. For example, Hatton (1997) exploits surveys of male workers in Michigan and California in 1890 and 1892, respectively. He argues that the assimilation into the job market is importantly influenced by the age of the immigrant upon arrival. Although immigrants who arrived as adults face an initial earnings disadvantage, their earnings grew faster relative to their native-born counterparts. Using 1900 and 1910 US census data, Minns (2000) finds a faster growth of immigrant earnings controlling for occupational sector and arrival cohort. In contrast, Feliciano (2001) estimates wage differentials between Mexican immigrants and native whites from 1910 to 1990. She finds a constant deterioration of relative earnings during the whole period.

The previous research consist of single or repeated cross sections that may overstate the convergence of immigrants to natives. Abramitzky et al. (2014) overcome this problem by constructing a panel data set of native-born workers and immigrants from 16 sending countries using the US census of 1900, 1910 and 1920. They obtain two results to highlight. First, on average, immigrants had the same performance as natives and they moved up the occupational ladder at the same rate. Second, they find substantial variation in the immigrant labor-market performance between sending countries.[16] Pérez (2017) also exploits census data (1869 and 1895) to follow natives and immigrants in Argentina. His

[16] In addition, Abramitzky et al. (2019d) find that past and present immigrants assimilate into the US culture at similar rates.

findings suggest that first and second-generation immigrants experienced faster occupational upgrading. Furthermore, immigrants' sons experienced mobility out of unskilled occupations. Similarly, Collins & Zimran (2018) study the assimilation of famine-era Irish migrants. Their findings suggest that famine-era Irish sons experienced convergence in occupational status. Interestingly, more Catholic surnames were associated with less upward mobility.

There are diverse factors that may influence immigrant assimilation. Eriksson (2018) shows that Norwegian-born immigrants living in ethnic enclaves experienced lower occupational earnings in 1910 and 1920. Eriksson & Ward (2018) measure segregation based on whether the next-door neighbor was native born. They show that only 40% of the European households had a native-born neighbor at arrival during the period 1905–09. This result suggest that immigrants spatially assimilate at a slow rate, which is consistent with the finding of Abramitzky et al. (2014): some immigrants started out the assimilation process from lower-paid occupations relative to natives. Moreover, Ward (2019a) finds that 51% of these initial ethnic gaps in occupational income remained after three generations. This was especially true for Mexican immigrants. Kosack & Ward (2019) use data on grandchildren and great-grandchildren of Mexican immigrants. They find that the occupational income gap initially experienced by the grandfather in 1880 remained unchanged across the three generations.

Two additional factors influencing immigrant assimilation are self-presentation and English fluency. Carneiro et al. (2016) evaluate the impact of choice name on immigrant assimilation. They argue that choice of an American first name was associated with higher job occupation scores. Goldstein & Stecklov (2016) find a similar result for the period 1880–1930: immigrant children with more American-sounding names experienced a higher occupational performance. Furthermore, Abramitzky et al. (2016) acknowledge that the association between name foreignness and economic and social outcomes can have intergenerational

effects. In contrast, Ward (2019b) argues that English fluency was less important for the assimilation of migrants into the job market at the turn of twentieth century. His estimates suggest that English fluency had an associated increase in occupational-based earning of 0.5–6.5 percent.

1.3 Structure and contributions of the thesis

As can be noticed from the previous review, the study of Mexican migration to the United States has been neglected by the cliometric literature on the Age of Mass Migration. The following four chapters intend to fill this gap in the literature. The common denominator of the previous cliometric scholarship is the use of large data sets. Chapter 2 introduces an unexploited source of data: the Mexican Border Crossing Records (MBCRs). These records report individual alien arrivals at diverse entrance ports along the Mexico-US border. With these micro data, I assess the geographic composition of the migration flow during the period 1906–08. In addition, Chapter 2 estimates migration rates at the local level and addresses the migration patterns existing before 1910. The subsequent chapters of the thesis use these micro data to study other dimensions of Mexican migration in early twentieth century.

Although scholarship on Mexican migration before 1910 provides many details about the forces behind the emigration decision, it lacks a comprehensive quantitative assessment to support their arguments. Chapter 3 estimates—for the first time—the effects of the push and pull factors that may have driven Mexican emigration at the beginning of the twentieth century.

Aside from immigration restrictions and liquidity constraints, we know little about the forces that may have changed migrant selection during the Age of Mass Migration. Chapter 4 intends to fill this void. First, I quantifying the selection of Mexican migration and use a natural experiment of history—the US financial crisis of 1907—to evaluate the sensitivity of selection patterns to random shocks.

Second, I explore potential mechanisms through which selection patterns can adjust to unexpected changes in the demand of immigrant labor.

Certainly, the factors determining the economic and non-economic assimilation of immigrants has gained strength in recent years. Nevertheless, there are still unexplored factors. Chapter 5 assesses if migration methods can influence the immigrant's labor-market performance at the destination. Specifically, I estimate the wage premium for stage migration. Finally, Chapter 6 summarizes the main findings of the thesis and outlines future research directions.

Chapter 2

Revisiting Mexican migration in the Age of Mass Migration

New evidence from individual border crossings

Abstract

This chapter introduces and analyses the Mexican Border Crossing Records (MBCRs), an unexplored data source that records aliens crossing the Mexico-United States land border at diverse entrance ports from 1903 to 1955. The MBCRs identify immigrants and report rich demographic, geographic and socioeconomic information at the individual level. These micro data have the potential to support cliometric research, which is scarce for the Mexico-United States migration, especially for the beginnings of the flow (1884–1910). My analysis of the MBCRs suggests that previous literature might have inaccurately described the initial patterns of the flow. My findings diverge from historical scholarship because the micro data capture the geographic composition of the flow at the local level, allowing me to characterize the initial migration patterns with precision. Overall, the micro data reported in the MBCRs offer the opportunity to address topics that concern the economics of migration in the past and present.

2.1 Introduction

There is extensive literature addressing the characteristics of the Mexico-United States migration. Angelucci (2015, 2012); Chort & De La Rupelle (2016); Donato (1993); Hanson & Spilimbergo (1999); Massey (1987); Massey & Espinosa (1997); and Takenaka & Pren (2010) analyze the forces driving fluctuations in legal and illegal migration flows from Mexico. They evaluate factors relaxing financial constraints to migration (cash transfers and household resources), structural conditions (US-Mexico wage gap, border enforcement and violence), random shocks (droughts), and factors derived from the historical persistence of the migration flow (immigrant networks and reunification processes). Ambrosini & Peri (2012); Caponi (2011); Chiquiar & Hanson (2005); Ibarraran & Lubotsky (2007); Kaestner & Malamud (2014); McKenzie & Rapoport (2010); Fernandez-Huertas Moranga (2011); and Orrenius & Zavodny (2005) examine the selection of Mexican immigrants using diverse earnings, educational and skill measures. In addition, Caponi (2011); Garcia & Schmalzbauer (2017); Lozano & Sorensen (2015); Munshi (2003); Perlmann (2005); and Vargas (2016) assess the performance of Mexican immigrants and their descendants in the US labor market over time.[1] Most of this research covers the period from 1980 onwards, although Mexican migration to the United States has existed since the end of the nineteenth century (Durand, 2016; Cardoso, 1980; Gamio, 1930).

In contrast, there is little cliometric literature on the Mexico-United States migration. Kosack & Ward (2014) estimate the selection pattern of Mexican immigrants and return immigrants in the 1920s. Feliciano (2001) examines the performance of Mexican immigrants in the US labor market from 1910 to 1990. Lee et al. (2017) analyze the impact of Mexican repatriations on labor market outcomes of US natives during the period 1930–40. Also, Clemens et al. (2018)

[1] See Borjas (2007) for additional literature on the selection and assimilation of Mexican migration to the United States.

evaluate the exclusion of Mexican farm workers—the Bracero Program (1942–64) abrogation—from the United States; and Kosack (2019) estimates the impact of this program on human capital investment in Mexico.[2]

Furthermore, our knowledge about Mexican migration from 1884 to 1910 relies on the historical research of Cardoso (1980); Chacón (2009); Clark (1908); Durand (2016); Fogel (1978); González (2010); and Verduzco (1995). This literature describes the initial migration patterns using ethnographic methods, newspapers, reports, personal experiences, and historical documents. Therefore, the arguments and theoretical propositions used in research on historical Mexico-United States migration are not tested or supported with representative quantitative evidence of the period.

The lack of cliometric literature for the beginnings of the migration flow (1884–1910) is due to the fact that available micro data for the period has not been exploited. This chapter has two objectives. First, it introduces an unexplored data source that records individual border crossings: the Mexican Border Crossing Records (MBCRs). Second, the it analyzes the MBCRs data available for the beginnings of the Mexico-United States migration and contrasts the results with the previous literature. Specifically, I exploit the publication N° A3365 that consists of manifests listing aliens arriving at nine entrance ports in Arizona and Texas from 1903 to 1910.[3] To my knowledge, the MBCRs have been used only by Kosack & Ward (2014). However, following the classification of Durand (2016, p. 7), the period covered in their research does not belong to the beginnings of the flow, but to the *Restrictions and Deportations Era (1921–41)*. Therefore, their findings do not capture the initial patterns of the flow, and their estimates may be influenced by the Mexican Revolution (1910–20).

[2]Although Gamio (1930) does not develop a strictly cliometric research, he presents a study—based on quantitative evidence—of money sent back to Mexico by immigrants from 1919 to 1926.

[3]Publication Title: Lists of Aliens Arriving at Brownsville, Del Rio, Eagle Pass, El Paso, Laredo, Presidio, Rio Grande City, and Roma (Texas) from May 1903 to June 1909; and at Aros Ranch, Douglas, Lochiel, Naco, and Nogales (Arizona) from July 1906–December 1910.

In the remainder of the chapter, I describe the characteristics of the MBCRs and publication N° A3365 in Section 2. I also provide evidence suggesting that the MBCRs are representative for the period under analysis. In Section 3, I present for the first time the initial spatial distribution of the migration flow at the local level. My analysis of the micro data offers an alternative narrative to historical literature regarding the immigrants' locations of last residence at the time. My findings diverge importantly from previous scholarship because the MBCRs capture migration flows across diverse entrance locations over long periods of time. This allows me to characterize migration patterns with precision. I offer concluding thoughts in Section 4.

2.2 The Mexican Border Crossing Records

The reporting of alien arrivals at the Mexico-US border started in few locations ca. 1903.[4] It was implemented systematically across entrance ports (border towns) in 1906 and fully established later under the Immigration Act of 1907 (US Congress, 1907, p. 908). From 1906 arriving aliens were classified into immigrants (those who intended to settle in the United States) and non-immigrants (those in transit, tourists and aliens returning to resume domiciles in the United States). The different forms used to register arriving aliens are known as Mexican Border Crossing Records (MBCRs), and they are cataloged by the National Archives and Records Administration (NARA) in publications covering the period ca. 1903 – ca. 1955.[5]

In this chapter, I present evidence from the MBCRs publication N° A3365. It contains 5 rolls of microfilms arranged chronologically by month-year covering the period from ca. 1903 to December 1910. The microfilms reproduce two-sheet manifests (Form 500-B) listing on average 30 aliens (see Figure 2.1 and Figure 2.2).

[4]The Immigration Act of 1903 instructed the inspection of aliens along the borders of Canada and Mexico (US Congress, 1903, p. 1221).

[5]See the National Archives and Records Administration (NARA) website for a full description of the publications and forms.

These documents were filled at the entrance port by registry clerks and supervised by immigration officials. Medical officers also examined the physical and mental health of all arriving aliens (US Congress, 1907, p. 903).[6] The manifests have 29 numbered columns that report information about the alien's profile and migratory experience. They report demographic (age, sex, marital status, occupation, ability to read and write, citizenship, and race) and anthropometric (height, complexion and color of eyes, and hair) data. They also record geographic information for each individual: birthplace, last permanent residence and final destination. In addition, they report whether the immigrant had a ticket to the final destination; if he/she had ever been in the United States (dates and places); and a contact (name and address) at the final destination. The back of the manifests contains detailed instructions to fill each column and definitions for the clerk to determine the alien's race, nationality, status (immigrant or non-immigrant), etc. See Figure 2.7 in Section 2.5.

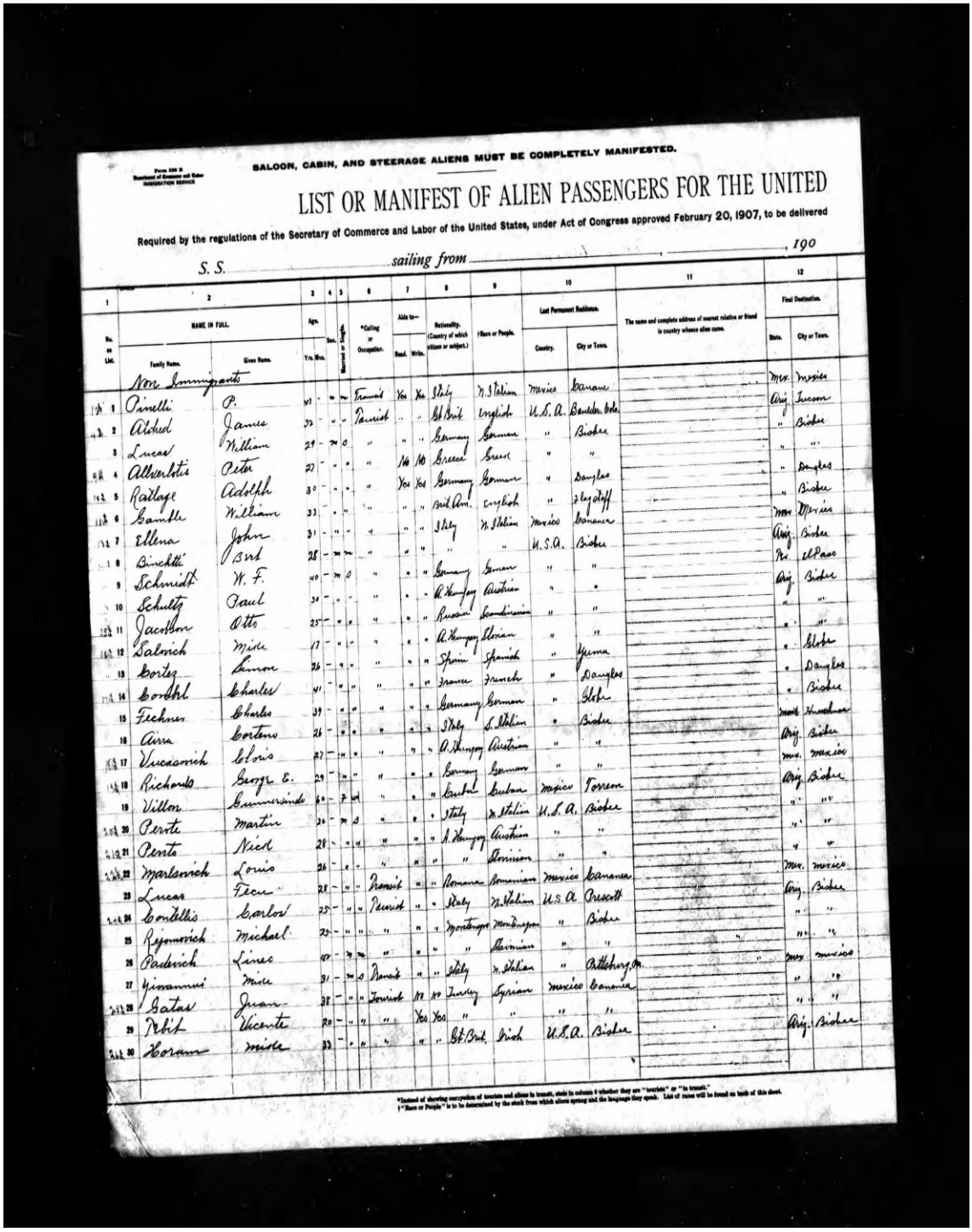
To study the initial migration patterns, it would be ideal to transcribe all data about Mexican aliens contained in the publication. However, the manifests were filled with handwriting, preventing the implementation of an automated transcription process. For this reason, I implement a sampling plan considering the large amount of data that had to be transcribed manually.

2.2.1 Sampling plan

I start by reviewing all manifests by year and entrance port to quantify the number of aliens listed as Mexican nationals: the population of interest (N). The year-entrance-port combinations or strata (s) intend to capture heterogeneity in migration patterns over time and across space. As a result of this revision, I identify that the first 115 manifests in roll 1 cover the period 1903–05 but regard to aliens others than Mexicans and do not report the entrance port consistently. Thus, I exclude them because they are not relevant for the research and cannot

[6]The medical officers should have at least two years of professional experience.

Figure 2.1: INS Form 500-B. Two-sheet manifest – Part A



SALOON, CABIN, AND STEERAGE ALIENS MUST BE COMPLETELY MANIFESTED.

LIST OR MANIFEST OF ALIEN PASSENGERS FOR THE UNITED

Required by the regulations of the Secretary of Commerce and Labor of the United States, under Act of Congress approved February 20, 1907, to be delivered

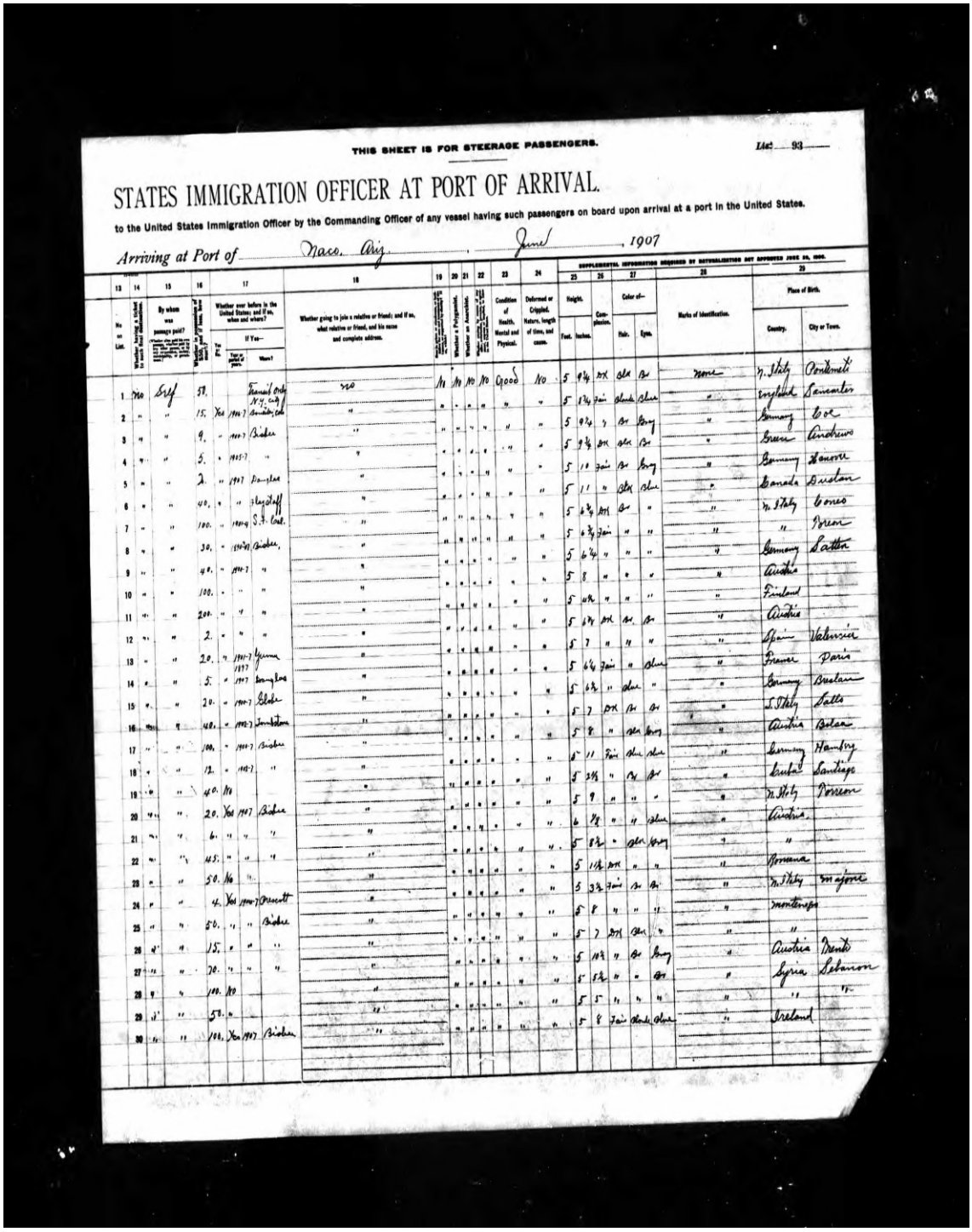
S. S. ..... sailing from ....., 190

Non Immigrants

No. on List	Family Name	Given Name	Age	Sex	Married or Single	Calling or Occupation	Able to Read	Able to Write	Nationality (Country of which citizen or subject)	Race or People	Last Permanent Residence: Country	Last Permanent Residence: City or Town	The name and complete address of nearest relative or friend in country whence alien came	Final Destination: State	Final Destination: City or Town
1	Pinelli	P.	41	M	M	Farmer	Yes	Yes	Italy	N. Italian	Mexico	Banamea		Mex.	Mexico
2	Aldred	James	32	"	"	Painter	"	"	Gt Brit	English	U.S.A.	Boulder, Colo.		Ariz.	Tucson
3	Lucas	William	29	"	S	"	"	"	Germany	German	"	Bisbee		"	"
4	Allenstein	Peter	27	"	"	"	No	No	Greece	Greek	"	"		"	Douglas
5	Railage	Adolph	30	"	"	"	Yes	Yes	Germany	German	"	Douglas		"	Bisbee
6	Gamble	William	33	"	"	"	"	"	Brit Dom.	English	"	Clayton		Mex	Mexico
7	Ellena	John	31	"	"	"	"	"	Italy	N. Italian	Mexico	Banamea		Ariz.	Bisbee
8	Binchetti	Bert	28	"	M	"	"	"	"	"	U.S.A.	Bisbee		Tex	ElPaso
9	Schmidt	W. F.	40	"	S	"	"	"	Germany	German	"	"		Ariz.	Bisbee
10	Schultz	Paul	30	"	"	"	"	"	A. Hungary	Austrian	"	"		"	"
11	Jacobson	Otto	25	"	"	"	"	"	Russia	Scandinavian	"	"		"	"
12	Salovich	Mike	17	"	"	"	"	"	A. Hungary	Slavian	"	"		"	Globe
13	Cortez	Simon	26	"	"	"	"	"	Spain	Spanish	"	Yuma		"	Douglas
14	Corthl	Charles	41	"	"	"	"	"	France	French	"	Douglas		"	Bisbee
15	Feckner	Charles	39	"	"	"	"	"	Germany	German	"	Globe		Mont	Hamilton
16	Aina	Cortens	26	"	"	"	"	"	Italy	S. Italian	"	Bisbee		Ariz.	Bisbee
17	Vucasovich	Louis	27	"	"	"	"	"	A. Hungary	Austrian	"	"		Mex.	Mexico
18	Richards	George E.	29	"	"	"	"	"	Germany	German	"	"		Ariz.	Bisbee
19	Villon	Gumersinde	40	F	M	"	"	"	Cuba	Cuban	Mexico	Torreon		"	"
20	Cerute	Martin	24	M	S	"	"	"	Italy	N. Italian	U.S.A.	Bisbee		"	"
21	Cento	Nick	28	"	"	"	"	"	A. Hungary	Austrian	"	"		"	"
22	Marlenvich	Louis	26	"	"	"	"	"	"	Slavonian	"	"		Mex.	Mexico
23	Lucas	Fecu	21	"	"	Farmer	"	"	Roumania	Roumanian	Mexico	Banana		Ariz.	Bisbee
24	Contellis	Carlos	25	"	"	Painter	"	"	Italy	N. Italian	U.S.A.	Prescott		"	"
25	Rejomovich	Michael	22	"	"	"	"	"	Montenegro	Montenegrin	"	Bisbee		"	"
26	Padovich	Lines	47	"	M	"	"	"	"	Slavonian	"	"		"	"
27	Giovanni	Mike	31	"	M	Farmer	"	"	Italy	N. Italian	"	Pittsburg, Pa.		Mex	Mexico
28	Gatas	Juan	38	"	"	Tourist	No	No	Turkey	Syrian	Mexico	Canonea		"	"
29	Ibit	Vicente	20	"	"	"	Yes	Yes	"	"	"	"		Ariz.	Bisbee
30	Koram	Mike	33	"	"	"	"	"	Gt Brit	Irish	U.S.A.	Bisbee			

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Immigration and Naturalization Service (INS) Form 500–B. *List or Manifest of Alien Passengers for the US Immigration Officer at Port of Arrival*. This form was traditionally used by vessel masters to record information about ship passengers in advance of arrival at US ports (NARA, 2000).

Figure 2.2: INS Form 500-B. Two-sheet manifest – Part B



THIS SHEET IS FOR STEERAGE PASSENGERS. List 93

STATES IMMIGRATION OFFICER AT PORT OF ARRIVAL.

to the United States Immigration Officer by the Commanding Officer of any vessel having such passengers on board upon arrival at a port in the United States.

Arriving at Port of Naco, Ariz., June 1907

Source: Mexican Border Crossing Records. Microfilm publication number N° A3365.
Note: Immigration and Naturalization Service (INS) Form 500–B. *List or Manifest of Alien Passengers for the US Immigration Officer at Port of Arrival*. This form was traditionally used by vessel masters to record information about ship passengers in advance of arrival at US ports (NARA, 2000).

be classified accurately. In addition, roll 5 contains data from 1909 to 1910, which I also exclude to avoid capturing any effects from the Mexican Revolution (1910–20). The objective of the paper is to study labor migration, and the presence of an armed conflict complicates the distinction between labor immigrants and refugees.[7] Table 2.1 summarizes the data contained in the publication's remainder. A total of 18,751 Mexican aliens crossed the border at nine entrance ports in Arizona and Texas from July 1906 to December 1908. The publication does not contain manifests for the first six months of 1906 or entrance ports in California.

Table 2.1 shows substantial variation in Mexican crossings between strata. As mentioned previously, the systemic registration of aliens at the Mexico-US border began in 1906 but fully enforced until later. This may explain the low number of crossings reported at Laredo and Brownsville in 1906 relative to following years. Also, in 1907 the American economy experienced one of the most severe financial crises before the Great Depression (Frydman et al., 2015, p .928; Moen & Tallman, 1992, p.611; Odell & Weidenmier, 2004, p. 1003). Banks and financial institutions of many cities limited or suspended their cash payments (Andrew, 1908, p. 497), and around two thousand firms and over one hundred state banks failed (Markham, 2002, p. 32). This event may have affected the number of border crossings in 1907. Furthermore, El Paso, Eagle Pass and Laredo were terminus stations of railways connecting central Mexico with the border (see Figure 2.8 in Annex A; Woodruff & Zenteno, 2007), which may explain the relatively high number of crossings at these entrance ports.

To select the data to be transcribed, I follow a criteria that considers the crisis of 1907 and the heterogeneous distribution of data between strata. First, since 1907 was an unusual year, potentially characterized by return migration and changes in the composition of migrants, I transcribe all data for this year regardless the entrance port. I also transcribe all data in strata that on average report 100 or less Mexican crossings per month. This allows me to capture with precision

[7] According to Dell (2012) insurgency events related to the Mexican Revolution started in 1909.

Table 2.1: Mexican and non-Mexican crossings (July 1906 – December 1908)

	Jul – Dec 1906			Jan – Dec 1907			Jan – Dec 1908		
	Total	*Mexicans*	*Share* [a]	*Total*	*Mexicans*	*Share* [a]	*Total*	*Mexicans*	*Share* [a]
Arizona									
Nogales	283	182	64	779	447	57	174	39	22
Naco	522	432	83	3,091	2,647	86	159	105	66
Douglas	202	172	85	627	405	65	197	153	78
Texas									
El Paso	3,722	2,815	76	4,678	974	21	3,293	2,361	72
Del Rio	8	8	100	81	74	91	201	200	99
Eagle Pass	180	180	100	1,679	138	8	1,073	697	65
Laredo	363	43	12	2,076	536	26	6,205	5,258	85
Roma				12	12	100	1	1	100
Brownsville	83	68	82	410	360	88	469	444	95
Total	**5,363**	**3,900**	**73**	**13,433**	**5,593**	**42**	**11,772**	**9,258**	**79**

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: [a] Percent. The table summarizes the data contained in rolls 1 to 4. Data contained in roll 5 record crossings in 1909 and 1910, which I did not considered to avoid capturing effects of the Mexican Revolution (1910–20). I identify Mexican aliens based on the reported nationality and country of birth. Non-Mexican crossings regard mainly to European and Asian aliens. After reviewing the microfilms, I did not find data for entrance ports in California.

patterns that may be underrepresented in the overall migration flow and that may have followed local dynamics. In these strata, differences between the number of transcribed and total Mexican crossings are due to the poor quality (unreadable or damaged) of some microfilms.

Second, in strata reporting on average more than 100 Mexican crossings per month (El-Paso-1906, El-Paso-1908 and Laredo-1908), I implement an equal probability systemic sampling. These strata capture 72% and 82% of all Mexican crossings in 1906 and 1908, respectively. For El-Paso-1906 stratum, I aim to transcribe 50% of the crossings, implying a fixed sampling interval of two observations—that is, I transcribe every 2^{nd} crossing if it is readable. For El-Paso-1908 and Laredo-1908 strata, I aim to transcribe 30% of the crossings. In these cases, the fixed sampling interval was three observations. The starting point for transcribing was determined by the random-number generator function of Stata.[8] Table 2.2 presents the transcribed sample: 10,895 Mexicans who crossed the border during the period July 1906 – December 1908.

[8]This function generates random integers from an specified interval.

Table 2.2: Transcribed Mexican crossings (July 1906 – December 1908)

	Jul – Dec 1906			Jan – Dec 1907			Jan – Dec 1908		
	Total	*Transcribed*	*Share* [a]	*Total*	*Transcribed*	*Share* [a]	*Total*	*Transcribed*	*Share* [a]
Arizona									
Nogales	182	154	85	447	447	100	39	39	100
Naco	432	372	86	2,647	2,163	82	105	105	100
Douglas	172	172	100	405	405	100	153	152	99
Texas									
El Paso	2,815	1,304	46	974	963	99	2,361	723	31
Del Rio	8	8	100	74	74	100	200	200	100
Eagle Pass	180	150	83	138	138	100	697	421	60
Laredo	43	43	100	536	506	94	5,258	1,513	29
Roma				12	12	100	1	1	100
Brownsville	68	68	100	360	360	100	444	402	91
Total	**3,900**	**2,271**	**58**	**5,593**	**5,068**	**91**	**9,258**	**3,556**	**38**

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: [a] Percent. The table summarizes the Mexican crossings contained in the rolls 1 to 4, and the sample transcribed by year and entrance port.

Finally, I estimate the weight of all units in each strata as:

$$w_s = \frac{N_s}{n_s}. \tag{2.1}$$

The application of weights (w_s) makes the transcribed sample match the population of interest in each strata (N_s). In the following sections, I use these weighting factors to estimate and analyze diverse aspects of the Mexican migration flow registered in the publication N° A3365.

2.2.2 Refinement of the data

The transcribed data in Table 2.2 constitute a gross flow of Mexican aliens that were not necessarily immigrants. Therefore, I apply a series of refinements to estimate accurately the flow of Mexicans migrants. First, I drop from the sample individuals whose final destination was in Mexico (return immigrants); and individuals whose last residence and final destination was in the United States (tourists or non-immigrants). Return migration represented 6.6% of the flow and the share of non-immigrants was 9.6%. Second, I drop immigrants with unreported or insufficient geographic data (last residence and final destination), which is necessary to estimate the migration flows. Finally, I classify the reported

locations of last residence and final destination as Mexican municipalities and American counties, respectively; and I drop the observations with unclassified locations. The final sample consists of 8,420 immigrants with full classified geographic information, representing 77.3% of the transcribed Mexican crossings (see Table 2.3). I obtain a flow of 15,215 immigrants by applying weighting factors to the refined sample. Table 2.4 presents its distribution by year and entrance port.

Table 2.3: Sample refinement

	Obs.	*Share (%)*
Transcribed crossings	10,895	100
Return immigrants	718	6.6
Non-immigrants	1,045	9.6
Immigrants	9,083	83.4
Last residence in Mexico		
Unreported	405	3.7
Not classified	10	0.1
A. Classified as Mexican municipalities	**8,668**	**79.6**
Final destination in the United States		
Unreported	203	1.9
Not classified	82	0.8
B. Classified as American counties	**8,798**	**80.8**
C. Final sample (A ∩ B)	**8,420**	**77.3**

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: *Return immigrants* refer to Mexican individuals whose final destination was in Mexico. *Non-immigrants* refer to Mexican individuals whose final destination and last permanent residence was in the United States. *Immigrants* refers to Mexican individuals whose last permanent residence was in Mexico and final destination was in the United States. C = Mexican immigrants whose last permanent residence and final destination was reported and classified in Mexican municipalities and US counties, respectively.

Table 2.4: Refined sample. Weighted flow (1906–08)

	Jul - Dec 1906		**Jan - Dec 1907**		**Jan - Dec 1908**		**Jul 1906 - Dec 1908**	
	Crossings	*Share* [a]	*Crossings*	*Share* [a]	*Crossings*	*Share* [a]	*Crossings*	*Share* [a]
Arizona								
Nogales	124	3.6	309	8.1	36	0.5	469	3.1
Naco	254	7.3	1,573	41.2	96	1.2	1,923	12.6
Douglas	101	2.9	194	5.1	125	1.6	420	2.8
Texas								0.0
El Paso	2,774	79.7	905	23.7	1,920	24.3	5,600	36.8
Del Rio	3	0.1	51	1.3	155	2.0	209	1.4
Eagle Pass	144	4.1	88	2.3	482	6.1	714	4.7
Laredo	28	0.8	382	10.0	4,698	59.3	5,108	33.6
Roma			12	0.3			12	0.1
Brownsville	54	1.6	302	7.9	404	5.1	760	5.0
Total	**3,483**	**100**	**3,816**	**100**	**7,916**	**100**	**15,215**	**100**

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: [a] Percent.

2.2.3 Representativeness of the sample

To assess the representativeness of the sample, it is necessary to consider that neither Mexico nor the United States kept systematically statistics of Mexican labor migration before 1910, making the MBCRs the only data capturing flows of immigrants. The open border policy of both governments and the uncontrolled 3,200 km long border made it difficult to record accurately the number of Mexican immigrants entering into or leaving the United States (Cardoso, 1980, p. 28 & 34). Thus, the few statistics available correspond to estimates from particular areas and specific periods of time.

Previous scholarship has accepted that, on average, 50 thousand Mexican immigrants crossed the US border every year during the first decade of the twentieth century.[9] This number—first proposed by Clark (1908, p. 520)—is a calculation from an official of the Mexican Central Railway. This figure consists of third class passengers who crossed the border at El Paso and Eagle Pass from August 1906 to August 1907. Taking this figure as true, the average crossings per month were 4,166. In the same period and entrance ports, my final weighted sample records 309 crossings per month, about 7% of Clark's monthly estimates. However, Clark (1908, p. 474) also argues that from January to September 1907, 26 thousand Mexican laborers entered to the United States through El Paso (2,888 laborers per month). My sample records 509 immigrants in July 1907, approximately 18% of the monthly flow estimated by Clark. Similarly, Cardoso (1980, p. 35) documents that from July 1908 to February 1909, 16,471 workers were recruited in El Paso. Assuming all laborers were Mexican, on average 2,058 immigrants were recruited per month. My sample records at this entrance port 215 crossings per month from July to December 1908, accounting for 10% of Cardoso's figure.

[9] This number is commonly extrapolated to estimate a flow of 500 thousand immigrants during the 1900–10 period (Cardoso, 1980, p. 34).

None of this research provides disaggregated statistics capturing the composition of the migration flow. Hence, I use other sources to assess if the composition of my sample is representative. One of them is El Economista Mexicano (1907), a Mexican newspaper reporting that 1,215 Mexicans migrated via El Paso in September 1907. Although my sample does not provide information for this month, the average monthly crossings during July and August 1907 accounts for 33% of this figure. More importantly, the newspaper presents statistics broken-down by the immigrants' state of origin.[10] Table 2.5 compares the statistics of El Economista Mexicano (1907) against my sample. Despite their difference in size, both samples present similar compositions: Bajio immigrants constitute more than 86%, which in fact matches the migration pattern described by previous historical scholarship.

The second source are the Abstracts of Reports of the Immigration Commission (US Immigration Commission, 1911a). The Immigration Act of 1907 (US Congress, 1907) established the creation of a commission to make a full investigation into the subject of immigration (US Immigration Commission, 1911a, p. 9). The Commission compiled existing data, and it secured original information from field investigations that were implemented across the United States from December 1908 to July 1909 (US Immigration Commission, 1911a, p. 15–20). I use the statistics on Mexican immigration for the fiscal years 1899 to 1910. Panel A of Table 2.6 shows that according to the Commission's calculations about 70% of the immigrants were laborers and 17% skilled workers. Farm laborers and professionals represented less than 5%. Also, 57% of the immigrants could neither read or write, and 66% were males (Panel B and C, respectively). Following the criteria and categories of the Immigration Commission, I calculate the composition of my sample based on the immigrants' occupation, sex and literacy. Table 2.6 shows that both compositions are very similar, suggesting that the manifests do not

[10]The newspaper does not clarify if the statistics refer to the place of last residence or place of birth.

capture disproportionately a specific immigrant profile.

Table 2.5: Composition of the migration flow at El Paso, Texas (1907)

	El Economista Mexicano *September*		**Border Crossing Records**[a] *July*		*August*	
	Immigrants	Share [b]	Immigrants	Share [b]	Immigrants	Share [b]
Panel A. States						
Guanajuato*	593	48.8	229	45.0	138	45.4
Michoacan*	279	23.0	72	14.1	64	21.1
Jalisco*	179	14.7	39	7.7	16	5.3
Zacatecas*	137	11.3	52	10.2	39	12.8
Durango*	14	1.2	17	3.3	12	3.9
Chihuahua	6	0.5	40	7.9	19	6.3
Mexico City	4	0.3	1	0.2	1	0.3
Aguascalientes*	3	0.2	32	6.3	3	1.0
Panel B. Regions						
Bajio	1,205	99.2	441	86.6	272	89.5
Border	10	0.8	41	9.3	20	6.6
Total	**1,215**	**100**	**509**	**100**	**304**	**100**

Source: El Economista Mexicano (1907) and Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: [a] Weighted flow. [b] Percent. *Bajio states.

Table 2.6: Composition of Mexican immigration to the United States (1899–1910)

	Immigration Commission (1899–1910)		**Border Crossing Records**[a] **(1906–08)**	
	Immigrants	*Share (%)*	*Immigrants*	*Share (%)*
Panel A. Occupations				
Laborers	15,763	69.3	7,144	72.1
Farm laborers	541	2.4	397	4.0
Skilled workers	3,918	17.2	1,036	10.5
Professionals	440	1.9	37	0.4
Other	2,095	9.2	1,292	13.0
Total[b]	**22,757**	**100**	**9,906**	**100**
Panel B. Literacy				
Illiterate	18,717	57.2	8,272	64.6
Total[c]	**32,721**	**100**	**12,810**	**100**
Panel C. Sex				
Males	27,676	66.0	10,992	72.2
Total	**41,914**	**100**	**15,215**	**100**

Source: US Immigration Commission (1911a, p. 97-101) and Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: [a] Weighted flow. [b] Immigrants without occupation were not considered. [c] Immigrants 14 years of age or over.

Considering that most figures presented in historical literature are back-of-the-envelope calculations, it is difficult to assess the real share of the migration flow registered in the publication N° A3365. However, the Immigration Commission

provides annual estimations based on diverse sources including statistical surveys. The Commission estimates a gross flow of 6,067 Mexican immigrants in 1908 (US Immigration Commission, 1911a, p. 95). My sample records 4,931 immigrants in the same year, 81% of the Commission's figure.

In sum, I believe that the MBCRs capture a representative sample of the Mexican immigration at the time, and the publication N° A3365 might record an important share of the gross flow. Furthermore, my sample presents a composition similar to the only statistics reporting the immigrants' location of origin at the state-level during the period under analysis (1906–08). This allows me to argue that the sample is representative for the migration flow entering through El Paso, which according to Clark (1908, p. 475) was the only real labor depot in the border. In addition, my sample presents a composition similar to studies addressing the characteristics of Mexican immigrants observed in the United States from 1899 to 1910. Together these comparisons provide evidence that my sample is representative for the Mexico-United States migration during the 1900s.

2.2.4 Limitations of the data

An important limitation of the sample is that it records crossings only at official entrance ports: documented immigration. However, estimations of undocumented Mexican immigration are scarce and imprecise for the period, because Mexicans had an undefined immigration status in the United States. Before 1910, Mexicans were not considered immigrants who sought to settle permanently, but temporary immigrants who moved back and forth supplying labor without major restrictions (Fogel, 1978, p. 10; Samora, 1982, p. 35).[11] Hence, the first Mexican immigrants did not have a clear incentive to avoid official entrance ports as it is nowadays, suggesting that MBCRs could be a reliable data source for the period.

[11]The Immigration Acts of 1903 and 1907 exempted incoming Mexicans from the head tax of $2.00 and $4.00, respectively (Cardoso, 1980, p. 34).

The desert in Arizona and New Mexico also complicates immigration through places others than the entrance ports in these states (see Figure 2.8 in Section 2.5).

A second limitation is that the geographic information was self-reported, leading to potential inaccuracies in the identification of birth, last residence and destination locations. For example, the manifests report the immigrant's "final destination", but it is likely that the records show intended destinations rather than the actual or final destinations of the immigrants. This could lead to a disproportionate representation of counties that were considered distributing points of Mexican labor (Clark, 1908, p. 475).

Potential problems of selection and under-enumeration could be a third limitation. Figure 2.8 in Section 2.5 shows that all entrance ports had direct access to railways (except Del Rio, Texas). Therefore, it could be that immigrants with access to railways or with resources to afford a train ticket are disproportionately recorded in the manifests. The data could also present different levels of under-enumeration between entrance ports. For example, entrance ports processing large amounts of immigrants could be more susceptible to under-enumeration than less dynamic ports.

Despite these issues, the MBCRs represent a unique source of data. To my knowledge, they are the only immigration data at the individual level, with which we can identify the characteristics of the Mexico-United States migration in its beginnings (1884–1910).

2.3 Initial patterns of Mexican migration

In this section, I address the Mexico-US migration patterns in the early twentieth century considering the immigrants' locations of last residence. My analysis exploits immigrant crossings registered at the main entrance ports during a time-span of 30 consecutive months. I also present, for the first time, the initial spatial distribution of the migration flow at the local level.

Figure 2.3: Migration regions and entrance ports (1906–08)

Source: Based on Durand (2016, p. 28) and Mexican Border Crossing Records. Microfilm publication N° A3365.

2.3.1 Municipalities of last residence

To study the characteristics of Mexican migration, previous literature has defined migration regions based on historical and geographic criteria. These regions (Bajio, Border, Center and Southeast) capture different migration patterns across Mexico that persist to this day (Durand, 2016, p. 27). I use these categories to contrast my results against previous scholarship. Figure 2.3 depicts the migration regions and the location of the entrance ports in Arizona (Nogales, Naco and Douglas) and Texas (El Paso, Del Rio, Eagle Pass, Laredo, Roma and Brownsville).

The Bajio region comprises the states lying just north of the Valley of Mexico and chiefly on the western slope of the central plateau (Clark, 1908, p. 468). These states were among the most populated in the beginning of the twentieth century, and they were characterized by their large agricultural and mining

centers.[12] The Border region covers the northern Mexican territory that was poorly populated until the 1950s. However, throughout the border states were consolidated economic centers connected to the United States and central Mexico by the railways of the time. The Center region covers the Valley of Mexico, which economic and political dynamism gravitated towards Mexico City, the capital of the country. The South region comprises the farthest states from the US border, which were relatively isolated from the rest of the country, except for the state of Veracruz where the most important seaport of Mexico was located.

Previous literature has agreed that in the beginnings of the flow most Mexican immigrants came from the Bajio, also known as the traditional or historical immigrant-sending region (Cardoso, 1980, p. 26; Clark, 1908, pp. 467–8; Durand, 2016; Gratton & Merchant, 2015, p. 528; p. 27–9 & 59–60; Henderson, 2011, p. 14; Ríos-Bustamante, 1981, p. 21; among others). However, the micro data suggest a different pattern. Table 2.7 shows that most immigrants actually came from the Border region. Immigrants from the Bajio represent only one third of the sample, and migration flows from the Center and South of the country were almost nonexistent.

Table 2.7: Region of last residence. Total weighted flow (1906–08)

	Crossings	*Share (%)*
Border	9,783	64.3
Bajio	5,178	34.0
Center	244	1.6
South	11	0.1
Total	**15,215**	**100**

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Figure 2.3 depicts the migration regions in Mexico.

[12]The Bajio states are: Durango, Zacatecas, San Luis Potosi, Nayarit, Aguascalientes, Guanajuato, Jalisco, Colima and Michoacan. Before 1917, the state of Nayarit was called Tepic. See Figure 2.3 for guidance.

Table 2.8: Twenty most important migrant-sending municipalities (1906–08)

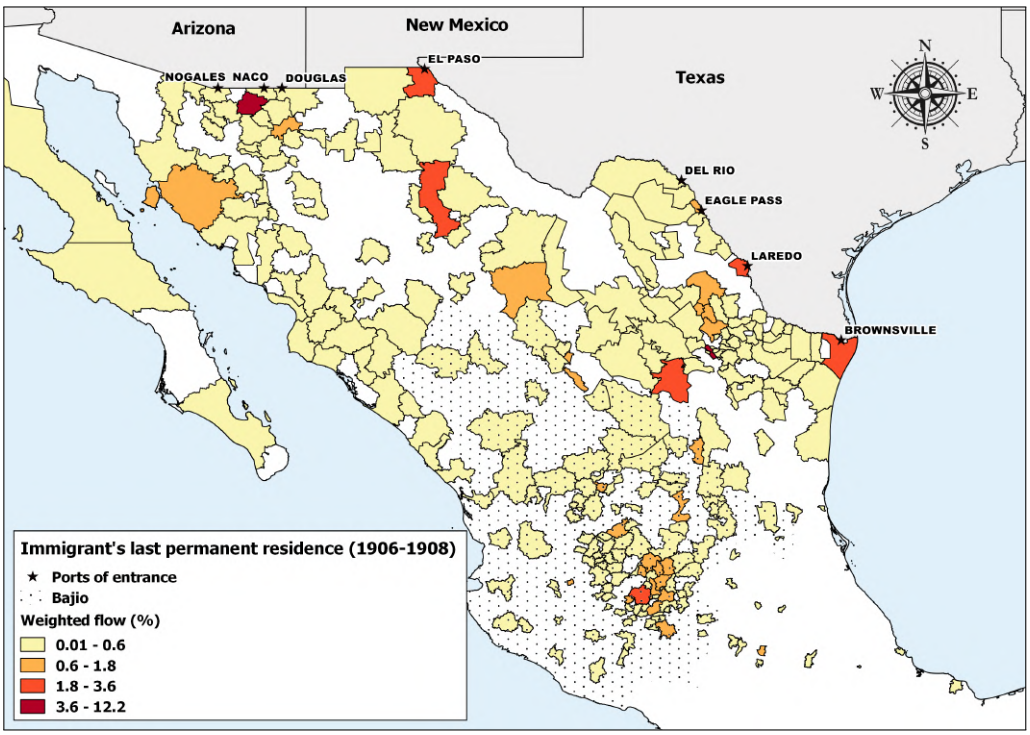
Municipality	*State*	*Weighted Flow*	*Share (%)*	*Migration rate*
Monterrey	Nuevo Leon	1,862	12.2	21.6
Cananea	Sonora	1,649	10.8	111.1
Chihuahua City	Chihuahua	550	3.6	10.2
Matamoros	Tamaulipas	521	3.4	32.5
Nuevo Laredo	Tamaulipas	489	3.2	54.9
Penjamo	Guanajuato*	439	2.9	7.9
Juárez City	Chihuahua	398	2.6	33.8
Saltillo	Coahuila	349	2.3	6.5
San Luis Potosi	San Luis Potosi*	275	1.8	3.3
Leon	Guanajuato*	259	1.7	2.9
Piedras Negras	Coahuila	259	1.7	21.5
Guadalajara	Jalisco*	254	1.7	2.1
Morelia	Michoacan*	234	1.5	2.9
Zacatecas	Zacatecas*	231	1.5	8.0
Villaldama	Nuevo Leon	223	1.5	33.5
Silao	Guanajuato*	211	1.4	5.9
Hermosillo	Sonora	206	1.4	9.1
Bustamante	Nuevo Leon	199	1.3	56.9
Irapuato	Guanajuato*	195	1.3	3.7
Mexico City	Mexico City	193	1.3	0.3

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: * Bajio states. See Figure 2.3 for the states location. I estimate migration rates (per 1,000 people) based on population levels from the 1910 Population Census. Mexico City's town halls were considered as a whole.

Furthermore, immigrants might have come disproportionately from specific states or municipalities within regions. To identify migration patterns at the local level, I estimate the total outflow of immigrants from each municipality that was reported as last permanent residence. Table 2.8 shows the top twenty municipalities that make up 60% of the total outflow. Four of these locations belong to the state of Guanajuato in the Bajio, and they account for 7.3% of the total outflow. From a local perspective, they make up 54.4% of the outflow from Guanajuato, implying that migration was highly clustered in few municipalities within the state. Considering that in 1910 the state had 45 municipalities, we can argue that migration was not a generalized experience, but a local phenomenon. Similarly, the state of Michoacan has an important participation in the total outflow (5.6%), but three municipalities (Morelia, La Piedad and Pururandiro) make up most migration (57.7%) from this state. The same pattern holds considering the state of Zacatecas. Jointly, Zacatecas City and the municipalities of Jerez and Nochistlan

concentrate three fourths of the state's outflow. In other words, the migration from the Bajio followed local dynamics before 1910.

Figure 2.4: Immigrant's last permanent residence (1906–08)

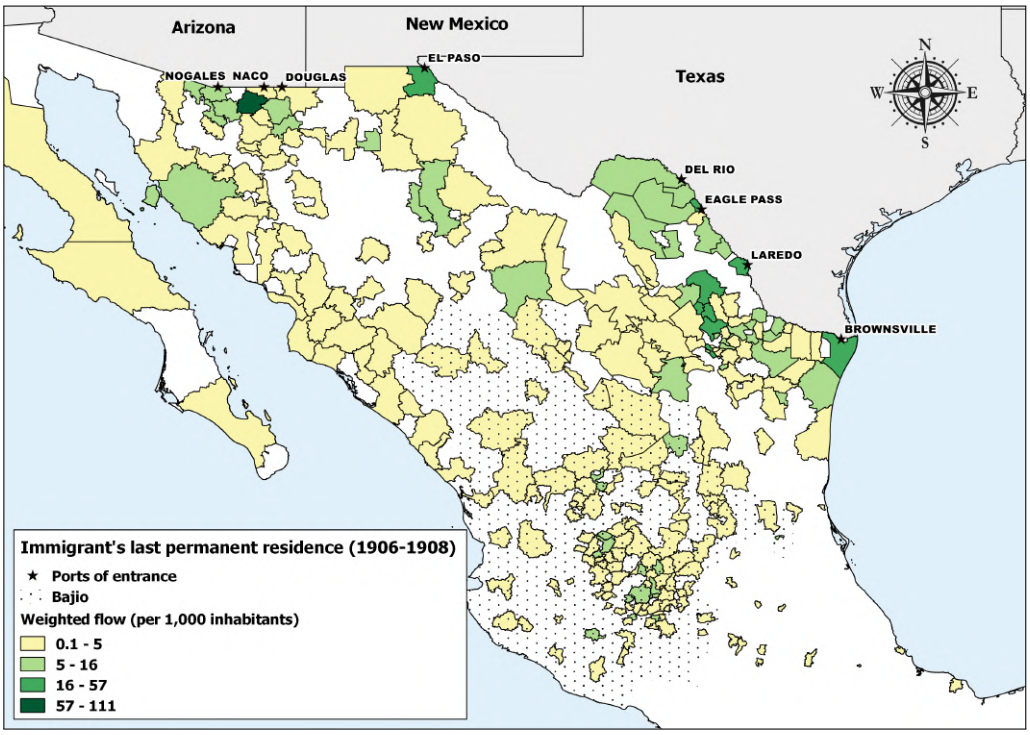


Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Spatial distribution of the Mexican-American migration flow from July 1906 to December 1908. The polygons display the immigrant's last permanent residence (municipalities) and their shares in the overall weighted flow (quartiles calculated with Jenks natural breaks classification method). The shaded area covers the states of the Bajio region.

Figure 2.4 presents the initial spatial distribution of the Mexico-US migration. Most immigrants from the Bajio actually came from a small group of adjoining municipalities in the states of Guanajuato, Jalisco and Michoacan. These locations were characterized for their intensive economic activity. By 1890, there were 31 haciendas in Guanajuato, which provided commodities to the region and 46 local mining centers (De Cardona, 1892). Although the importance and productivity of these centers varied, all of them extracted silver and gold. This attracted workers from all over the country, keeping labor supply high and consequently low salaries in the region. Migration from other Bajio municipalities was scarce and had low shares in the total outflow. Table 2.8 and Figure 2.5 confirm that

migration rates in the region were relatively low: on average, two immigrants per 1,000 people.[13]

Figure 2.5: Migration rates – last permanent residence (1906–08)



Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Spatial distribution of the Mexican-American migration flow from July 1906 to December 1908. The polygons display the immigrant's last permanent residence (municipalities) and their migration rate per 1,000 persons (quartiles calculated with Jenks natural breaks classification method). The shaded area covers the states of the Bajio region.

In the Border region, Nuevo Leon, Sonora, and Chihuahua were the main immigrant-sending states, which were poorly populated until the second half of the twentieth century. Thus, its geographic location might have driven their migratory importance. Similar to the Bajio, migration in the Border region was concentrated in few municipalities, but these locations were distributed across the region. Monterrey and Cananea present the highest shares in the total outflow (12.2 and 10.8 percent, respectively. The former was a dynamic smelter city and the latter emerged in the mid-nineteenth century as an important mining center (Cardoso, 1980, p. 17). The average migration rate in the Border region was six

[13]The states of Guanajuato, Jalisco and Michoacan were among the most populated in the country (see Figure 2.9 in the Annex). Hence, the low share of Bajio immigrants in the sample also reflects low migration rates.

immigrants per 1,000 people, but in the top ten municipalities, it was about 41 immigrants per 1,000 people. This corroborates that migration was intense in several municipalities of the Border region (see Figure 2.5). These results line up with recent findings suggesting that from 1900 until 1920, Mexican migration to the United States was characterized by a high level of circular cross-border mobility of young men (Gratton & Merchant, 2015, p. 532).

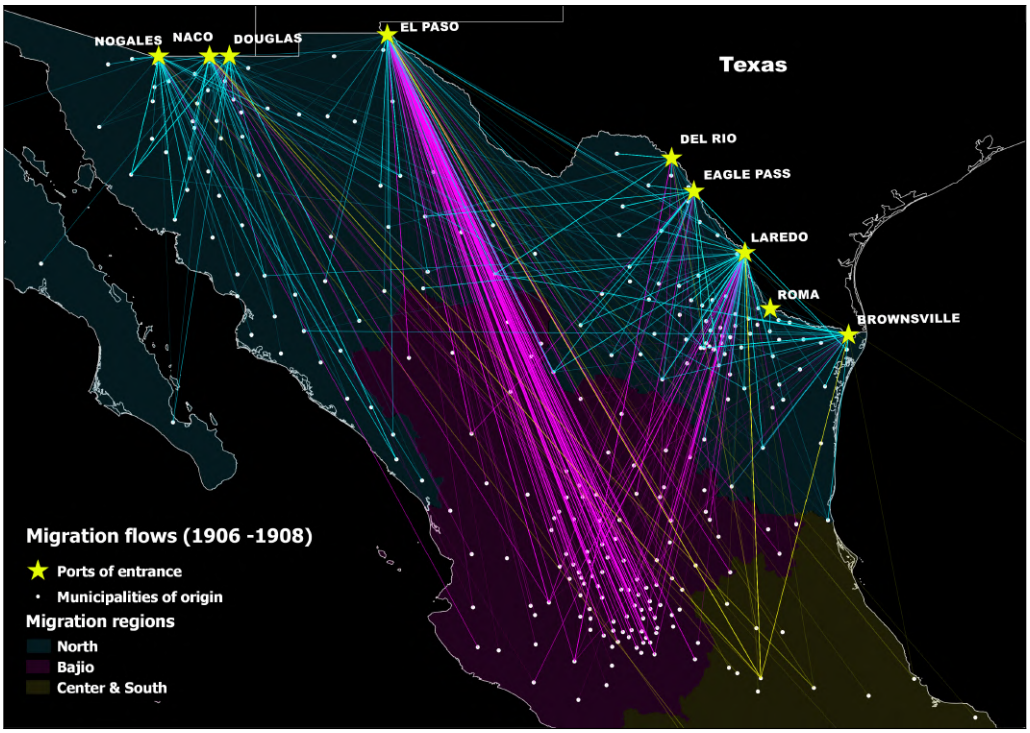
2.3.2 Explaining the divergence of patterns

Why does the previous migration patterns diverge importantly from the previous historical literature? The answer to this question is because the influential work of Clark (1908), which is the most cited reference for the period, might be biased to a large extent. When one analyzes his paper, it is clear that entrance ports other than El Paso are not analyzed in detail or even mentioned. Although he addresses the labor conditions and available wages for several places along the border, his seminal work describes the composition of the migration flow via el El Paso and Eagle Pass only. Figure 2.6 depicts the intensity of the migration flows at the time. It shows that most immigrants registered at El Paso came from Bajio states. For this reason, Clark (1908, p. 468) concludes that in 1908 most of the migration flow occurred between the Bajio and El Paso.

However, this is not precise. My sample reveals that migration via Arizona is not insignificant as Clark (1908) suggests. On the contrary, the flow of Mexican immigrants registered at Naco was greater than in Eagle Pass in 1906 and 1907 (see Table 2.4). Also, migration via Laredo was more intense than the registered at El Paso or Eagle Pass in 1908. In this sense, my results diverge from Clark's because my sample captures immigration across a broader array of entrance locations and over a longer period of time.

On the other hand, the micro data support findings from literature studying immigration at locations other than El Paso. For example, Gamio (2002, p. 182)

Figure 2.6: Intensity of emigration streams by entrance port (1906–08)



Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Each line represents an individual. Overlapping lines capture the intensity of a migration corridor by adding pixel values of one line with the other. Hence, brighter lines represent more intensive migration flows.

documents that Mexicans working in the south of Texas came mostly from Nuevo León and Tamaulipas. In my sample, 68% of the immigrants registered at Laredo came from those states. The same pattern is observed when analyzing the flow registered at Brownsville: 92% of the immigrants came from Nuevo León and Tamaulipas. Immigrants from the Bajio represented less than 17% and 2% of the crossings registered at Laredo and Brownsville, respectively.

Another example is González (2010, p. 12 & 18), who documents that in 1888 there was a constant flow of families migrating from Sonora to Arizona; and that there was a notorious flow of Mexicans migrating from Sonora and Sinaloa to Kansas by 1907. In my sample, 90% of the Mexicans crossing the border via Nogales, Naco and Douglas came from Sonora and Sinaloa. Registers of Bajio immigrants at these ports were almost nonexistent (see Figure 2.6). In sum, the

micro data from the MBCRs capture better the geographic composition of the flow, allowing to characterize the initial migration patterns with more precision.

2.4 Conclusions

I have presented evidence suggesting that historical scholarship may have described inaccurately the initial patterns of the Mexico-United States migration. Based on the immigrants' last residence, my findings confirm that there was a geographic selection of Mexican immigrants at the beginning of the flow. However, most immigrants came from the Border region and not from the Bajio as suggested by Clark (1908); Cardoso (1980); Durand (2016); among others. Moreover, Bajio immigrants actually came from a small group of adjoining municipalities. This suggests that the Bajio was still not consolidated as the principal immigrant-sending region and probably its migration culture was in the process of gaining strength.

In addition, my local-level analysis reveals two additional characteristics of the migration flow: immigrants came from specific municipalities, and migration rates were heterogeneous within and across states. The immigrant-sending municipalities were economically dynamic and populated locations. By themselves, these municipalities attracted laborers from all over Mexico, but labor market pressures jointly with the higher wages offered in the American Southwest might have motivated immigrants to continue moving north (Clark, 1908, p. 470; Durand, 2016, p. 61). In other words, migration at the time did not follow regional but local dynamics. These results do not necessarily contradict migration patterns described by previous literature, but they expand our knowledge about Mexican migration using quantitative evidence not analyzed previously.

The individual-level data reported in the MBCRs offer the opportunity to address diverse topics in migration economics. New statistical methods developed by Abramitzky et al. (2019a) and Abramitzky et al. (2019b) can be implemented to

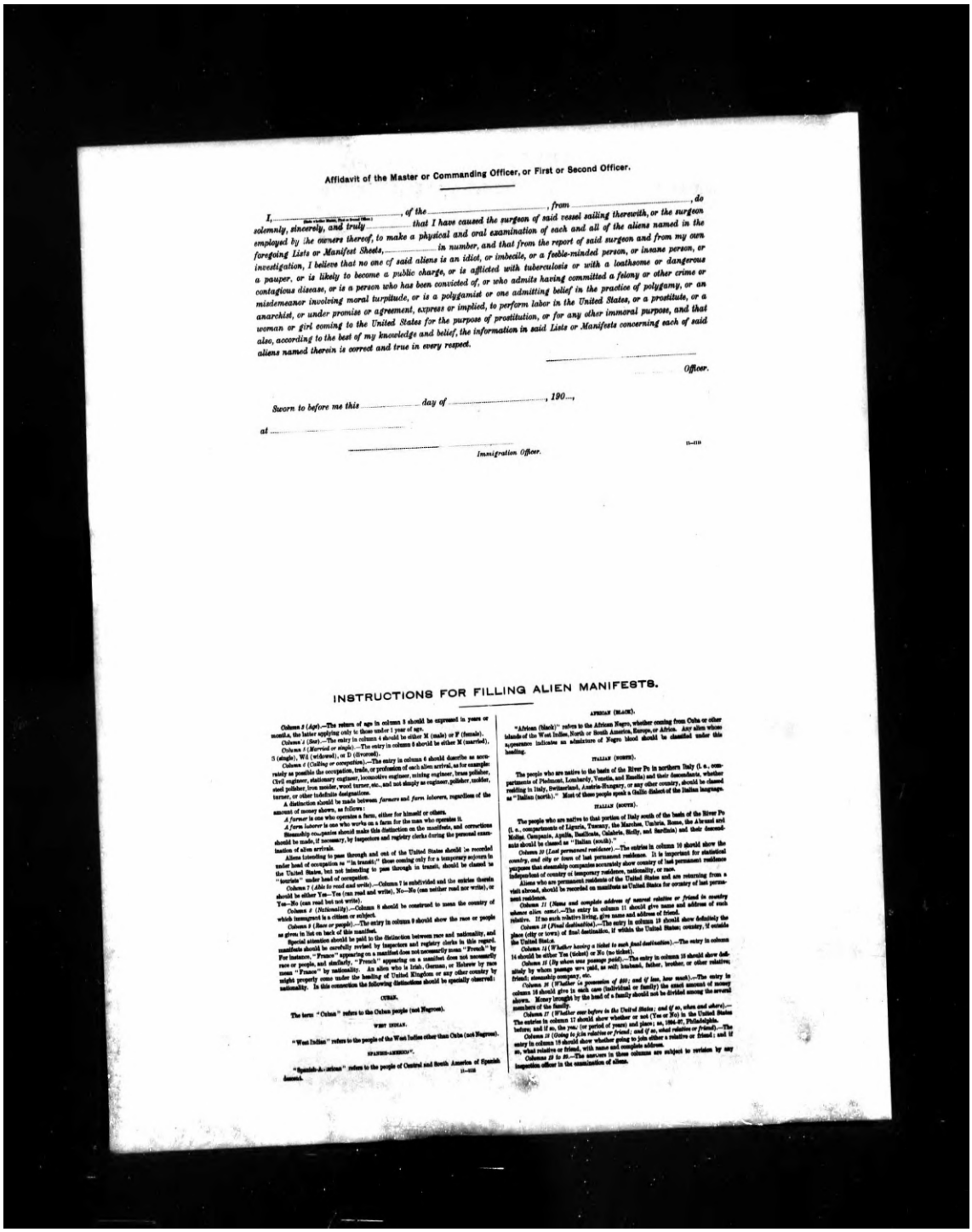
link immigrants recorded in the MBCRs with other historical sources. This could allow the development of research similar to Abramitzky et al. (2014); Inwood et al. (2019) and Ward (2019b), who study the assimilation and performance of immigrants during the early twentieth century. Since the MBCRs record return migration, it is also possible to examine the selection pattern into migration and into return migration like Abramitzky et al. (2019c); Kosack & Ward (2014).

The geographic data reported in the MBCRs allow to estimate initial migration rates at the local level, which can be used in approaches similar to Sequeira et al. (2019) for evaluating the long-run effects of Mexican migration on economic and development outcomes in both Mexico and the United States. Also, migration models à la Hatton & Williamson (1993); and (Hatton & Williamson, 1994; Hatton, 1995b,a) can be tested to study the determinants of Mexican migration in the Age of Mass Migration.

In sum, the MBCRs represent a unique source of micro data to develop cliometric research addressing the initial mechanics of the most intense and persistent migration of the twentieth century.

2.5 Appendix

Figure 2.7: INS Form 500-B. Instructions for filling alien manifests



Affidavit of the Master or Commanding Officer, or First or Second Officer.

I, ..., of the ....., from ....., do solemnly, sincerely, and truly ..... that I have caused the surgeon of said vessel sailing therewith, or the surgeon employed by the owners thereof, to make a physical and oral examination of each and all of the aliens named in the foregoing Lists or Manifest Sheets, ... in number, and that from the report of said surgeon and from my own investigation, I believe that no one of said aliens is an idiot, or imbecile, or a feeble-minded person, or insane person, or a pauper, or is likely to become a public charge, or is afflicted with tuberculosis or with a loathsome or dangerous contagious disease, or is a person who has been convicted of, or who admits having committed a felony or other crime or misdemeanor involving moral turpitude, or is a polygamist or one admitting belief in the practice of polygamy, or an anarchist, or under promise or agreement, express or implied, to perform labor in the United States, or a prostitute, or a woman or girl coming to the United States for the purpose of prostitution, or for any other immoral purpose, and that also, according to the best of my knowledge and belief, the information in said Lists or Manifests concerning each of said aliens named therein is correct and true in every respect.

.....
Officer.

Sworn to before me this ..... day of ....., 190...,

at .....

.....
Immigration Officer.

INSTRUCTIONS FOR FILLING ALIEN MANIFESTS.

Column 3 (Age).—The return of age in column 3 should be expressed in years or months, the latter applying only to those under 1 year of age.

Column 4 (Sex).—The entry in column 4 should be either M (male) or F (female).

Column 5 (Married or single).—The entry in column 5 should be either M (married), S (single), Wd. (widowed), or D (divorced).

Column 6 (Calling or occupation).—The entry in column 6 should describe as accurately as possible the occupation, trade, or profession of each alien arrival, as for example: Civil engineer, stationary engineer, locomotive engineer, mining engineer, brass polisher, steel polisher, iron molder, wood turner, etc., and not simply as engineer, polisher, molder, turner, or other indefinite designations.

A distinction should be made between *farmers* and *farm laborers*, regardless of the amount of money shown, as follows:

A farmer is one who operates a farm, either for himself or others.

A farm laborer is one who works on a farm for the man who operates it.

Steamship companies should make this distinction on the manifests, and corrections should be made, if necessary, by inspectors and registry clerks during the personal examination of alien arrivals.

Aliens intending to pass through and out of the United States should be recorded under head of occupation as "in transit," those coming only for a temporary sojourn in the United States, but not intending to pass through in transit, should be classed as "tourists" under head of occupation.

Column 7 (Able to read and write).—Column 7 is subdivided and the entries therein should be either Yes—Yes (can read and write), No—No (can neither read nor write), or Yes—No (can read but not write).

Column 8 (Nationality).—Column 8 should be construed to mean the country of which immigrant is a citizen or subject.

Column 9 (Race or people).—The entry in column 9 should show the race or people as given in list on back of this manifest.

Special attention should be paid to the distinction between race and nationality, and manifests should be carefully revised by inspectors and registry clerks in this regard. For instance, "France" appearing on a manifest does not necessarily mean "French" by race or people, and similarly, "French" appearing on a manifest does not necessarily mean "France" by nationality. An alien who is Irish, German, or Hebrew by race might properly come under the heading of United Kingdom or any other country by nationality. In this connection the following distinctions should be spatially observed:

CUBAN.

The term "Cuban" refers to the Cuban people (not Negroes).

WEST INDIAN.

"West Indian" refers to the people of the West Indies other than Cuba (not Negroes).

SPANISH-AMERICAN.

"Spanish-American" refers to the people of Central and South America of Spanish descent.

11—113

AFRICAN (BLACK).

"African (black)" refers to the African Negro, whether coming from Cuba or other islands of the West Indies, North or South America, Europe, or Africa. Any alien whose appearance indicates an admixture of Negro blood should be classified under this heading.

ITALIAN (NORTH).

The people who are native to the basin of the River Po in northern Italy (i. e., compartments of Piedmont, Lombardy, Venetia, and Emilia) and their descendants, whether residing in Italy, Switzerland, Austria-Hungary, or any other country, should be classed as "Italian (north)." Most of these people speak a Gallic dialect of the Italian language.

ITALIAN (SOUTH).

The people who are native to that portion of Italy south of the basin of the River Po (i. e., compartments of Liguria, Tuscany, the Marches, Umbria, Rome, the Abruzzi and Molise, Campania, Apulia, Basilicata, Calabria, Sicily, and Sardinia) and their descendants should be classed as "Italian (south)."

Column 10 (*Last permanent residence*).—The entries in column 10 should show the country, *and city or town* of last permanent residence. It is important for statistical purposes that steamship companies accurately show country of last permanent residence independent of country of temporary residence, nationality, or race.

Aliens who are permanent residents of the United States and are returning from a visit abroad, should be recorded on manifests as United States for country of last permanent residence.

Column 11 (*Name and complete address of nearest relative or friend in country whence alien came*).—The entry in column 11 should give name and address of such relative. If no such relative living, give name and address of friend.

Column 12 (*Final destination*).—The entry in column 12 should show definitely the place (city or town) of final destination, if within the United States; country, if outside the United States.

Column 13 (*Whether having a ticket to such final destination*).—The entry in column 14 should be either Yes (ticket) or No (no ticket).

Column 15 (*By whom was passage paid*).—The entry in column 15 should show definitely by whom passage was paid, as self; husband, father, brother, or other relative; friend; steamship company, etc.

Column 16 (*Whether in possession of $50; and if less, how much*).—The entry in column 16 should give in each case (individual or family) the exact amount of money shown. Money brought by the head of a family should not be divided among the several members of the family.

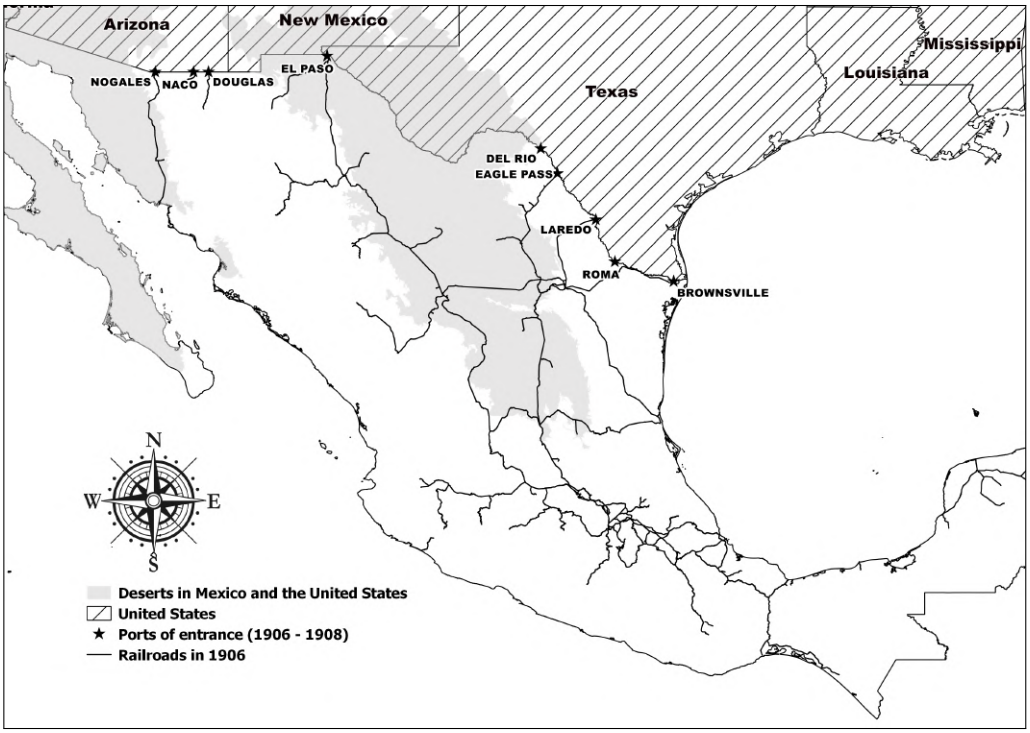
Column 17 (*Whether ever before in the United States; and if so, when and where*).—The entries in column 17 should show whether or not (Yes or No) to the United States before; and if so, the yes. (or period of years) and place; as, 1904-07, Philadelphia.

Column 18 (*Going to join relative or friend; and if so, what relative or friend*).—The entry in column 19 should show whether going to join either a relative or friend; and if so, what relative or friend, with name and complete address.

Columns 20 to 29.—The answers in these columns are subject to revision by any inspection officer in the examination of aliens.

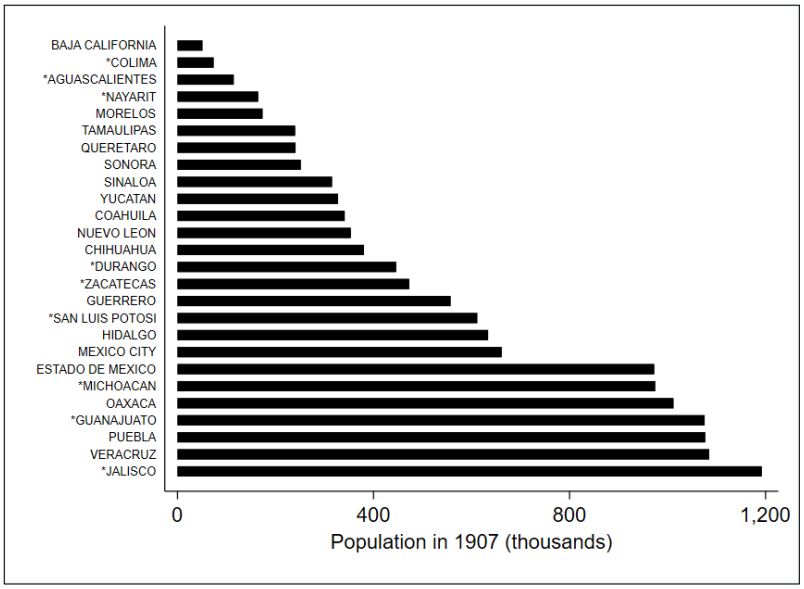
Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Immigration and Naturalization Service (INS) Form 500–B. *List or Manifest of Alien Passengers for the US Immigration Officer at Port of Arrival*. The back of the manifests contains detailed instructions to fill each of the 29 columns. Also, they contain definitions for the clerk to determine the alien's race, nationality, occupation/status, etc.

Figure 2.8: Entrance ports (1906–08), railroads in Mexico ca. 1906 and deserts



Source: Secretaría de Comunicaciones y Obras Públicas (1906), United States Environmental Protection Agency and Mexican Border Crossing Records. Microfilm publication number A3365.

Figure 2.9: Mexican population by state in 1907



Source: Secretaría de Economía (1956).
Note: *Bajio states. Considering the population levels in 1907, the Bajio states were among the most populated. The states of Guanajuato, Jalisco and Michoacan were more populated than Mexico City at the time. This could explain the low migration rates observed in Bajio municipalities and the high migration rates in the Border region locations before 1910. Before 1917, the state of Nayarit was called Tepic.

Chapter 3

Initial determinants of Mexican mass migration

Abstract

Exploiting novel micro data consisting of individual border crossings from 1906 to 1908, this chapter addresses the initial determinants of Mexican emigration at the local level. First, it estimates gross emigration between Mexican municipalities and US counties, obtaining 892 migration corridors or municipality-county pairs. Second, it evaluates diverse push and pull factors that may explain differences in size across migration corridors. The findings suggest that differences in economic conditions and Mexican immigrant networks in the United States were the main pull factors rather than the US-Mexico wage gap. Regional droughts in Mexico were the main push factor. Despite their importance for the Mexican economy, railways had a limited effect on the emigration decision at the beginning of the Mexico-United States mass migration.

3.1 Introduction

In the beginnings of the Age of Mass Migration (1850-1913), 300 thousand Europeans migrated to New World destinations every year (Hatton & Williamson, 1998, p. 7). In the last decade of that period, a new episode of mass migration started but from Mexico to the United States. Between 60 to 100 thousand Mexican laborers crossed the border in 1908 (Clark, 1908, p. 466). This figure remains at similar levels today, making the Mexico-US migration the most intense and persistent labor migration of the twentieth century.

While there is extensive literature addressing the Mexico-US migration in recent times, our knowledge about its initial characteristics has been limited to the historical research of Clark (1908) and Cardoso (1980). They argue that, at the beginning of the flow (1884-1910), most immigrants came from the densely populated central plateau of Mexico (known as the Bajio) (see Figure 3.1). In addition, previous scholarship has focused on the Mexico-US wage gap and the expansion of Mexico's railways network to explain the emigration of Mexicans during this period (Durand, 2016; Gratton & Merchant, 2015; Henderson, 2011; among others). Thus, the story of Mexican emigration was similar to the general mass migration story (Hatton & Williamson, 1993, Hatton & Williamson, 1994; Hatton, 1995b; Hatton, 1995a; and Williamson, 1998).[1] For the first time, I provide a quantitative assessment of the initial determinants of the Mexico-United States emigration using micro data not analyzed previously. The results offer an alternative narrative to previous literature.

The core data come from manifests that recorded individual border crossings. These documents are known as Mexican Border Crossing Records (MBCRs), and they were used by American authorities to collect information about individuals that crossed the Mexico-US border. The sample used in this research consists of

[1]The general narrative of the Age of Mass Migration is that long run trends of emigration responded systematically to real wage gaps between home and abroad, declining transportation costs and migrant networks; being the wage gaps in the core.

15,215 immigrants that crossed the border through nine ports of entrance (see Figure 3.1). The data was transcribed following a stratified sampling plan in which I select the subsamples of each strata (ports of entrance) using an equal probability systematic sampling. The data covers the period from July 1906 to December 1908. I did not consider data from 1909 to 1920 to avoid capturing effects from the Mexican Revolution (1910–20), an event that complicates the distinction between labor immigrants and refugees. The manifests report rich and diverse demographic, geographic and anthropometric information for each immigrant. However, this research uses the immigrant's last permanent residence and final destination. These locations were classified as Mexican municipalities and US counties to estimate the gross flow of immigrants in 892 migration corridors or municipality-county pairs. These data capture the regional composition of the flow, favoring the identification of different migration models at the local level. Chapter 2 provides a full description of the MBCRs and the sampling plan I follow.

Figure 3.1: Migration regions and entrance ports (1906–08)

Source: Based on Durand (2016, p. 28) and Mexican Border Crossing Records. Microfilm publication N° A3365.

Like other literature about Mexican migration, I used the Border, Bajio, Center and Southeast regions as reference to contrast my results (Durand, 2016, p. 28). The Bajio region comprises the states lying just north of the Valley of Mexico and on the western slope of the central plateau. These states were among the most populated in the beginning of the twentieth century and were characterized by its large agricultural and mining centers. The Border region covers the northern Mexican territory that was relatively depopulated until the 1950s. However, throughout the border states were consolidated economic centers connected to the United States and central Mexico by the railways of the time. The Center region covers the Valley of Mexico, which economic and political dynamism gravitated towards Mexico City, the capital of the country. The Southeast region comprises the farthest states from the US border and were relatively isolated from the rest of the country, except for the state of Veracruz where the most important seaport of Mexico was located (see Figure 3.1).

The findings of this research are two-fold. First, I quantify the effects of diverse forces that determined the flow at the time: distance costs, market potentials/economic conditions, wage differentials, immigrant networks in the United States, relative dryness in Mexican municipalities, and living standards in Mexico. These push and pull factors have been mentioned by previous literature as determinants of the flow. Yet, we do not know their effects' magnitude nor their individual impact when controlling for the others. To evaluate these factors, I use ordinary least squares (OLS) regressions. Contrary to Hatton & Williamson (1998), differences in the Mexico-US relative wage were not the main driver of the migration flow. However, differences in population size at home and climate shocks mattered. This result reveals that, on average, locations with large populations represented a source of frictions in the Mexican labor market, pushing laborers to migrate. The estimates also suggest that the flow was consistently driven by the social capital formation at the destination: immigrant networks. Just as Massey & Espinosa (1997) and Takenaka & Pren (2010) find for

recent periods, migrating to a county with a large Mexican community might have increased the net benefits of migration for the average immigrant. Therefore, for more than one hundred years, immigrant networks have represented a self-perpetuating social asset that provides information and assistance, which reduces the costs and risks of migrating. The regional estimates reveal that two migration models existed at the time. In the Border region, distance costs, networks, and labor-market conditions in the origin and destination influenced the decision to migrate. In contrast, migration from the Bajio was importantly determined by immigrant networks and/or labor institutions.

Second, Cardoso (1980), Durand (2016) and other scholars argue that railways were a fundamental determinant of the flow. To evaluate the importance of railways in encouraging migration, I implement an instrumental variable strategy à la Banerjee et al. (2012). The identification strategy exploits differences in distance between municipalities of origin and historical transportation corridors that proxy for the access to railways. The results show that, in the Border region, the access to railways had a small effect on the migration flow relative to the distance cost elasticity of migration: railways might have not been necessary to reach the border. However, the proximity to a transportation corridor significantly influenced the decision to migrate in the Bajio. Precisely, for Bajio migrants, mobility costs did not arise from long distances but from the access to transportation towards the US border: railways were fundamental to explain the migration flow from this region. However, since only one-third of the overall migration flow came from the Bajio, this result confirms that in the beginnings of the twentieth century railways were not accessible in all migrant-sending municipalities and/or they were not used for migrating because they did not reduce significantly migration costs.

The contributions of this chapter to the cliometric literature on international migration are two-fold. First, the micro data confirm the intuition of Abramitzky

& Boustan (2017, p. 1326): in the Age of Mass Migration, the decision to migrate was a function of diverse forces, and the effects and magnitudes of these forces varied across regions of sending countries. Therefore, research providing national-level results might serve little to identify migration patterns accurately during this period.

Second, the paper presents evidence that migration costs were not the same for everyone as it is commonly assumed (Hatton, 2010, p. 944). Thus, not only migration costs evolve along time, but they might be different in nature across regions of sending countries.

The rest of the chapter is organized as follows. The next section presents the historical context of the research. Subsequently, I describe the characteristics of the data. In the fourth section, I present the spatial distribution of the flow and its main characteristics. The fifth and sixth sections address the empirical strategy and results. The last section concludes.

3.2 The beginnings of Mexican mass migration (1884–1910)

The migration of Mexicans to the United States started from a historic perspective in 1848, when the Mexican-US War ended. As principal consequence of the conflict, Mexico conceded more than a half of its territory, and lost 75% of the population living in those lands (Verduzco, 2000).[2] Those that remained in the new American territories became immigrants without ever leaving their home, but more importantly, they became the first Mexican immigrant network.[3] It is not clear when Mexicans started to emigrate in large numbers,[4] but during the 1900s Mexican immigration increased sharply and expanded its geographic range

[2]The lost territories were California, Utah, Nevada, and most of Arizona, New Mexico, and Colorado. The Mexican-US War also formalized the loss of Texas, admitted to the Union in 1845.
[3]Henderson (2011) estimates that this initial network of Mexican immigrants was about 80 to 100 thousand.
[4]Previous research suggests that Mexican mass migration took place during the Mexican railways expansion (Feliciano, 2001, p. 388; Cardoso, 1980, p. 13).

of settlement, creating a Mexican-American Southwest (Gratton & Merchant, 2015, p. 521 & 528).[5] This was possible due to the robust economic growth of the US economy that demanded an inexhaustible source of labor.[6]

Furthermore, there were important differences in wage levels between both countries. In the agriculture sector, where most of the Mexicans were employed in their hometowns, a peasant could earn four to eight times more in the US, depending on the crop cultivated (Cardoso, 1980, p. 22). A similar relation is observed in the mining sector, and the wages offered in the US railways were at least three times higher than in the Mexican Central Railroad (Kuntz, 1995). The source of this differentials were the stagnation of Mexican salaries (see Table 3.8 and Table 3.9 in Section 3.8) and an unfavorable exchange rate.[7] An additional incentive to emigrate was rooted in the deep impoverishment of the rural population, whose living conditions worsened during the dictatorship of Porfirio Diaz (Cardoso, 1980; Henderson, 2011; Oñate, 1991)—known as the *Porfiriato* (1877–1911).[8] Thus, the sustained US labor demand, the wage differentials between the countries and the deteriorated living standards in Mexico are the structural push and pull factors behind the migration flow.

However, the Porfirian Mexico experienced a profound modernization as well. Most of the current south-north Mexican railways were constructed during this period. In fact, the mileage increased from 477 km in 1877 to 19,000 km in 1910 (Cosío Villegas & Bernal, 1973; Henderson, 2011).[9] This technological change shortened distances and increased considerably unit savings on freight opera-

[5] Mexican immigrants satisfied labor demand in farms, mines and railroads across Arizona, New Mexico and Texas (Fogel, 1978, p. 10).
[6] The US economy achieved an average GDP growth rate of 4.5 to 5% in the last decades of the nineteenth century (Balke & Gordon, 1989; Romer, 1989; Rhode, 2002).
[7] The exchange rate at the time was 2 pesos per US dollar (Clark, 1908, p. 480; Kuntz, 1982, p. 46).
[8] López-Alonso (2007) argues that statures declined for most of the second half of the nineteenth century. This evidence confirms the deterioration of living standards before and during the Porfiriato.
[9] Figure 3.8 depicts the fast expansion of the Mexican railways during this period.

tions (Coatsworth, 1979).[10] Similarly, Cardoso (1980); Durand (2016) and Gamio (1930) argue that railways were fundamental to understand Mexican emigration, because they reduced importantly travel times to the border. Nevertheless, the stagnated salary of the Mexican working class questions the social savings on passenger transportation associated to this technology (Coatsworth, 1979, p. 960). This condition may have promoted the use of the *enganche*—a search-matching labor institution—by Mexican and American recruiters to transport and allocate seasonal laborers in the United States (Clark, 1908; Durand, 2016; Gamio, 1930). Since recruiters covered the transportation costs in exchange of future labor, the *enganche* was a mechanism that eliminated mobility and job-search costs associated to emigration.

The previous factors operated under a favorable policy environment. Before 1910, Mexicans were not considered immigrants who sought to settle permanently, but temporary immigrants who moved back and forth supplying labor without major restrictions (Fogel, 1978, p. 10; Samora, 1982).[11] The absence of immigration restrictions can be understood as a pull factor that left Mexican immigrants with an undefined immigration status (Durand, 2016, p. 74). This legal lacunae makes the beginnings of the migration flow a period with no constraints on labor mobility. In addition, two US immigration policies favored even more the Mexican migration. In 1882, the Chinese exclusion law prohibited the importation and utilization of Chinese labor; and in 1907 the Japanese immigration was prohibited through the Gentleman's Agreement (Samora, 1982). These policies generated a scarcity of cheap labor, specially in the agriculture sector and railways industry (Durand, 2016, p. 73). Thus, Mexicans faced a constant labor demand arising from the buoyant growth of the US economy, and could migrate freely to satisfied it since they count with mechanisms to overcome migration costs.

[10] The economic impact of railways in Mexico was higher than in other countries. Herranz-Loncán (2014) argues that railroads accounted for 24% of the Mexico's income per capita growth before 1914.

[11] Incoming Mexicans were exempted from the head of $2.00 and $4.00 levied, respectively, by the Immigration Acts of 1903 and 1907 (Cardoso, 1980, p. 34).

In sum, Mexico and the United States reached the end of the nineteenth century with clear development asymmetries. These structural differences along with other factors such as droughts,[12] immigrant networks, the access to railways and labor institutions influenced the emigration of Mexicans before 1910, and set the conditions for a mass labor migration throughout the twentieth century.[13]

3.3 Data

To evaluate the determinants of the Mexican emigration, this research exploits an original dataset with four core components: 1) an immigrant sample that records the gross flow of migrants between Mexican municipalities and American counties; 2) population data of Mexican municipalities and US counties that capture the labor market potential at home and abroad; 3) wage data of Mexican regions and US counties that capture the potential labor income in both countries; and 4) the distance between origin and destination locations. The population and wage data represent the economic push (incentives to leave Mexico) and pull (incentives to move to the US) factors of the flow, and the distance capture migration costs.

3.3.1 Immigrant sample

To estimate gross emigration between Mexican municipalities and American counties, this research uses a weighted sample of 15,215 immigrants who crossed the border from July 1906 to December 1908.[14] These data come from the Mexican Border Crossing Records (MBCRs) publication N° A3365, which consists of manifests listing arriving aliens at nine ports of entrance. They report rich individual-level data including the immigrants' place of last permanent residence

[12] Cardoso (1980, p. 12) argues that some Mexican regions experienced droughts during the period covered in this research.

[13] According to Feliciano (2001, p. 388), Mexicans living in the US represented 0.8% of the foreign-born population in 1890. A hundred years later, this figure increased to 21.7%.

[14] Data from 1910 to 1920 was not considered because the Mexican Revolution took place during this period, complicating the distinction between labor migrants and refugees.

and destination—information that this chapter exploits.[15] The data was transcribed following a stratified sampling plan in which I select the subsamples of each stratum (ports of entrance) using an equal probability systematic sampling. The strata are intended to capture different migration patterns driven by their geographic location or migration tradition. Chapter 2 describes in detail the sampling strategy as well as the characteristics and limitations of the sample. It also provides evidence suggesting that the sample is representative for the period under analysis.

Table 3.1 presents the distribution of the sample across entrance ports. As can be noticed, crossings in 1906 and 1907 are considerably lower than in 1908. This is because the manifests start to report consistently the entrance port from July 1906. This also may explain the low share of crossings at Laredo in this year.[16] The low number of crossings in 1907 are explained by the Panic of 1907—the most important financial crisis in the United States before the Great Depression—that sparked in May and affected the demand of immigrant labor (Andrew, 1908; Frydman et al., 2015; Odell & Weidenmier, 2004).

Table 3.1: Total weighted flow (1906–08)

	Jul - Dec 1906		**Jan - Dec 1907**		**Jan - Dec 1908**		**Jul 1906 - Dec 1908**	
	Crossings	*Share* [a]	*Crossings*	*Share* [a]	*Crossings*	*Share* [a]	*Crossings*	*Share* [a]
Arizona								
Nogales	124	3.6	309	8.1	36	0.5	469	3.1
Naco	254	7.3	1,573	41.2	96	1.2	1,923	12.6
Douglas	101	2.9	194	5.1	125	1.6	420	2.8
Texas								0.0
El Paso	2,774	79.7	905	23.7	1,920	24.3	5,600	36.8
Del Rio	3	0.1	51	1.3	155	2.0	209	1.4
Eagle Pass	144	4.1	88	2.3	482	6.1	714	4.7
Laredo	28	0.8	382	10.0	4,698	59.3	5,108	33.6
Roma			12	0.3			12	0.1
Brownsville	54	1.6	302	7.9	404	5.1	760	5.0
Total	**3,483**	**100**	**3,816**	**100**	**7,916**	**100**	**15,215**	**100**

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: [a] Percent. See Figure 4.5 for the location of entrance ports.

[15] Publication Title: Lists of Aliens Arriving at Brownsville, Del Rio, Eagle Pass, El Paso, Laredo, Presidio, Rio Grande City, and Roma (Texas) from May 1903 to June 1909; and at Aros Ranch, Douglas, Lochiel, Naco, and Nogales (Arizona) from July 1906–December 1910.
[16] Microfilms 145 to 199 of roll 1 do not always report the entrance port and year.

3.3.2 Population data

To my knowledge, there is no population data at the local for the period 1906–07. Hence, I obtain municipality-level population data from the Mexican Census of 1910.[17] I use Table VI (Chapter III) of the census official results that presents the number of residents in each municipality of the country (Secretaría de Agricultura y Fomento, 1910, p. 108). Similarly, I obtain county-level population data from the US Census of 1910 (Bureau of the Census, 1910).

3.3.3 Wage data

Mexico

I use regional data on Mexican wages from Arnaut (2018), who provides locally-adjusted real wages for the period under analysis (1906–08). Although Arnaut (2018) and Rosenzweig (1965, p. 447) provide wage data by economic sector, it was not considered because it is not broken down by region, i.e. variation across space would be lost when using wages by sector. Hence, each municipality was assigned the wage level of the region to which it belongs. Table 3.2 shows wages across Mexican regions. Clearly, at the end of the Porfiriato there were notorious asymmetries within the country. While the Gulf and North Pacific regions had the highest wage level, laborers in Central Mexico could earn 60% less than the average minimum wage in the country.

United States

To my knowledge, US wage data at the county-level does not exist for the period. Hence, I collect wage data from miscellaneous sources reporting wage levels in specific cities and economic activities. I classify wages by economic sector

[17] At the time, Mexico was organized in 27 states, 3 territories and one Federal District (Mexico City). The states and territories were divided in municipalities with different categories (*Pueblo*, *Ciudad*, *Villa*, among others) depending on their population density (Secretaría de Agricultura y Fomento, 1910, p. 43). In 1910, the Baja California peninsula (nowadays divided into two states: Baja California Norte and Baja California Sur) was a single federal territory. Therefore, the peninsula of Baja California is considered a single state along this research.

(agriculture, manufactures and mining) and type of wage earner (Mexican or average laborer). Wages for Mexicans in the three sectors were obtained from Clark (1908). General wage levels for agriculture activities come from bulletins published by the US Department of Agriculture (Flint, 1900; Hudson, 1914). Wages in manufacture activities come from the 1905 Census of Manufactures (U.S. Department of Commerce and Labor, 1906a,b) and statistics of wages in manufacture industries (U.S. Department of Commerce and Labor, 1907).

Table 3.2: Minimum wage in Mexico by region (1906–08)
(Cents per day - US dollars)

	1906	**1907**	**1908**
Mexico	*22.0*	*21.5*	*20.5*
North	19.9	21.0	21.4
Gulf	33.8	31.1	28.8
North Pacific	24.3	23.4	23.6
South Pacific	16.8	16.9	15.8
Center	14.1	13.9	13.8

Source: Arnaut (2018, Annex - p. 5).
Note: Regional price deflators (1900=100). The states constituting each region are the following. **North:** Coahuila, Tamaulipas, Chihuahua, Nuevo León, San Luis Potosi*, Durango* and Zacatecas*. **Gulf:** Yucatán, Campeche, Veracruz and Tabasco. **North Pacific:** Baja California, Sonora, Tepic and Sinaloa. **South Pacific:** Colima*, Chiapas, Guerrero and Oaxaca. **Center:** Mexico City, Morelos, Aguascalientes*, Puebla, Querétaro, Tlaxcala, Hidalgo, Estado de México, Guanajuato*, Jalisco* and Michoacán*. * Bajio states. The table shows that most of the Bajio states had the lowest salaries at the time. Values for Mexico are the weighted average of the regionally adjusted wages.

County-level wages were imputed from these sources, prioritizing wage levels for Mexican workers over general wage levels because Mexican immigrants commonly faced labor discrimination (Clark, 1908, p. 479).[18] When wage levels for more than one economic sector were available, I prioritize the composition of the migration corridor based on the immigrant's reported occupations in the manifests, e.g. if most of the immigrants were farm laborers, I imputed wage levels from agriculture activities. Finally, if more than one wage level was available for each economic activity, I imputed the lowest value to obtain minimum salaries. Table 3.10 in Section 3.8 presents this data in detail.

[18] At the time, Mexican immigrants were seen as cheap and labor. Normally, they were employed at second-class jobs, which were characterized by arbitrary lower salaries, breach of contracts, payment in rations, besides racial discriminating (Clark, 1908, p. 477 & 488; Durand, 2016, p. 71).

Table 3.3: Average wage in the United States by state and economic activity (US dollars per day)

	Agriculture	Manufactures	Mining
Arizona	2.16	2	2
California	1.75	1.65	.
Colorado	.	1.75	1.5
Illinois	.	1.52	.
Iowa	.	1.5	.
Kansas	.	1.44	.
Lousiana	.	1.4	.
Missouri	.	1.46	.
New Mexico	.	1.25	2.18
New York	.	1.42	.
Ohio	.	1.37	.
Oklahoma	.	1.68	.
Texas	0.56	1.49	1.3
Wisconsin	.	1.29	.

Source: Miscellaneous sources detailed in Table 3.10.
Note: Estimates of wages at the state-level. Data from wages at US destination counties.

These sources provide wage levels in 114 counties across 14 states. Table 3.3 shows that wages in manufacture activities were relatively similar across states. Texas had the lowest wage levels in all sectors, suggesting that wages in locations near the border or easily accessible might have been lower (Clark, 1908, p. 478). Finally, for those counties without available data, I assigned the wage level from the nearest county with available data. The nearest county was identified with the geographic distance between county centroids.

3.4 Migration corridors

Using the immigrants' last permanent residence and final destination, I identify 892 migration corridors—municipality-county pairs—and estimate the size of gross emigration in each of them. Table 3.4 presents the twenty most intense corridors at the municipality-county level, which represent 44% of the weighted flow. Seven of these corridors had their origin in municipalities of Guanajuato. They concentrate 10.6% of the total outflow, and Penjamo-El Paso was the most intense corridor of the state. However, Guanajuato's share in the total outflow

was only 13.4%, implying that emigration was highly concentrated in no more than 10 municipalities. Considering that in 1910 the state had 45 municipalities, we can argue that emigration was heterogeneous within states. Similarly, the state of Michoacan had an important participation in the flow (5.6%), but emigration was mainly concentrated in the Morelia-El Paso corridor and streams from La Piedad, Pururandiro and Huaniqueo. The same pattern holds considering the state of Zacatecas.

Furthermore, most migration corridors from the Bajio had El Paso as destination (see Figure 3.2). Clark (1908, p. 475) argues that among the border cities, El Paso was the only distributing point of Mexican labor. At El Paso, immigrants were met by representatives of companies or private agents, who recruit and distribute workers to diverse locations in the United States. Representatives also went to inland Mexico to recruit and transport laborers to the United States. This search-matching mechanism of immigrant labor is known as the *enganche*, and it was used to supply all the labor needed in the railway and mining industries (Durand, 2016, p. 56 & 63).

The most intense migration corridors from the Border region had their origin in municipalities relatively close to the border. These corridors are captured by the brighter blue lines in Figure 3.2. Thus, geography might have driven emigration from this region. These results line up with recent findings suggesting that from 1900 until 1920, Mexican migration to the US was characterized by a high level of circular cross-border mobility of young men (Gratton & Merchant, 2015, p. 532). Similar to the Bajio, emigration from the Border region was concentrated in few municipalities, but it had a greater diversity regarding the immigrants' destinations.

These two migration patterns are depicted in Figure 3.2. Most emigration streams from the Bajio ended at the border, but it is unlikely that the local economy could employ all these laborers. In fact, Mexican immigrants reported

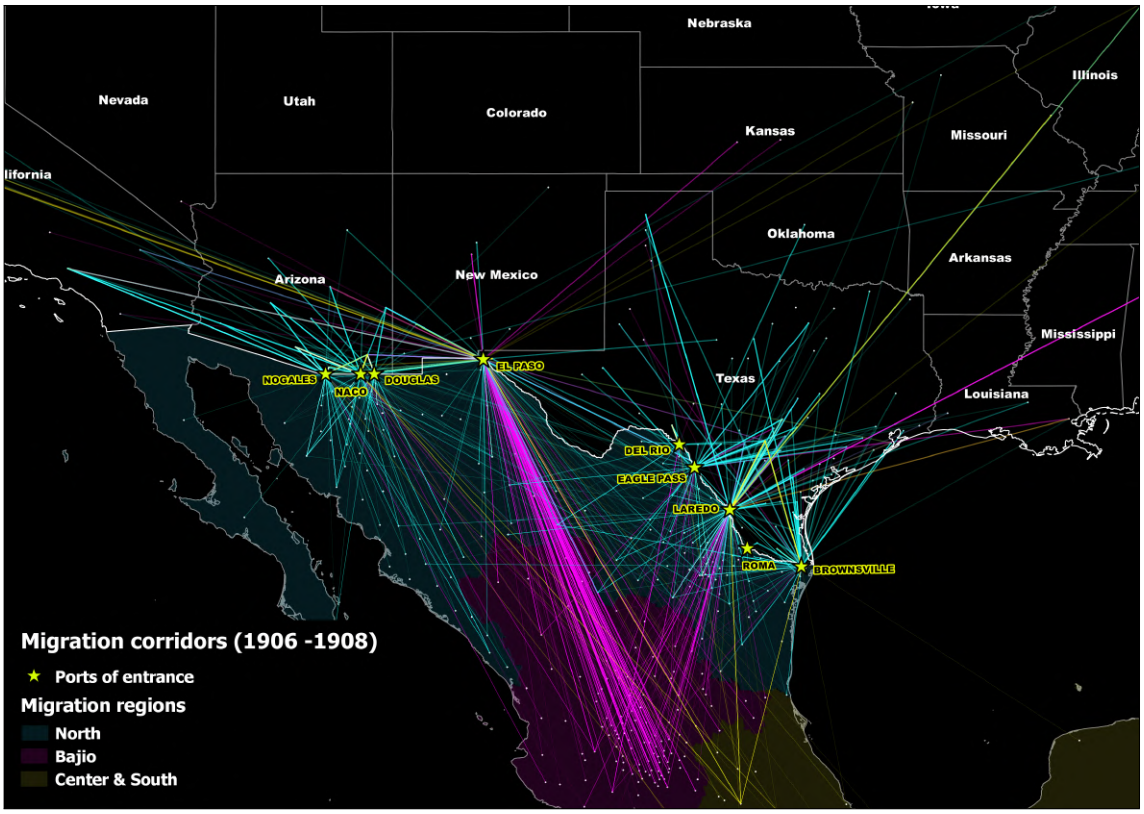
Table 3.4: Twenty most important migration flows from Mexican municipalities to US counties (1906–08)

Last Permanent Residence		**Final Destination**			
State	*Municipality*	*State*	*County*	*Weighted Flow*	*Share (%)*
Sonora	Cananea	Arizona	Cochise	1,264	8.3
Nuevo Leon	Monterrey	Texas	Webb	908	6.0
Chihuahua	Chihuahua	Texas	El Paso	451	3.0
Guanajuato*	Penjamo	Texas	El Paso	424	2.8
Nuevo Leon	Monterrey	Texas	Bexar	399	2.6
Tamaulipas	Nuevo Laredo	Texas	Webb	392	2.6
Tamaulipas	Matamoros	Texas	Cameron	357	2.3
Chihuahua	Juarez	Texas	El Paso	345	2.3
Michoacan*	Morelia	Texas	El Paso	223	1.5
Guanajuato*	Leon	Texas	El Paso	221	1.5
Sonora	Cananea	Arizona	Pima	207	1.4
Guanajuato*	Silao	Texas	El Paso	192	1.3
Zacatecas*	Zacatecas	Texas	El Paso	188	1.2
Coahuila	Piedras Negras	Texas	Maverick	181	1.2
Jalisco*	Guadalajara	Texas	El Paso	179	1.2
Guanajuato*	Irapuato	Texas	El Paso	170	1.1
Nuevo Leon	Villaldama	Texas	Webb	148	1.0
Guanajuato*	San Fco. del Rincon	Texas	El Paso	144	0.9
Guanajuato*	Guanajuato	Texas	El Paso	143	0.9
Guanajuato*	Abasolo	Texas	El Paso	142	0.9
			Total	**6,680**	**43.9**

Source: Mexican Border Crossing Records. Microfilm publication number A3365.
Note: The table presents the most intense migration corridors between Mexico and the United States before 1910. While flows from the Bajio are present, previous literature underestimated the importance of flow originated in the Border region. *Bajio states.

El Paso as final destination because they did not know where they would end up working. Some Bajio immigrants might have arrived by themselves at El Paso to secure employment at the contracting houses and be transported to other destinations. However, recruiters transporting laborers from the Bajio might have speculated with labor as well. That is, they may have hold immigrants at their headquarters in El Paso until they secure high commissions for delivering workers to companies in need of workers (Clark, 1908, p. 475). This concentration of immigrants at labor depots suggests that Bajio emigration was driven by specific migration traditions and/or the presence of labor institutions. In contrast, Figure 3.2 shows that emigration streams from the Border region ended at

Figure 3.2: Migration corridors (1906–08)



Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Each line represents an individual. Overlapping lines capture the intensity of a migration corridor by adding pixel values of one line with the other. Hence, brighter lines represent more intensive migration corridors.

diverse counties in Arizona, Southern California and Texas. Immigrants from the border states might have not used the *enganche* as part of the emigration process. Probably, the proximity to the United States facilitated the access to information relevant for the emigration decision, and/or the higher wages in the north of Mexico removed income constraints to emigration. The next section tests the existence of these migration patterns and evaluates the push and pull factors that may explain differences in size between migration corridors.

3.5 Push and pull factors of Mexican emigration

At the turn of the twentieth century, there were important differences in wage levels between Mexico and the United States, and both economies were connected by railways. Clark (1908), Cardoso (1980), Durand (2016), Henderson (2011),

Gratton & Merchant (2015) among others argue that these factors are "key" to understand the mass migration of Mexicans. Yet, we do not know their effects' magnitude nor their individual impact when controlling for other forces. In this section, I provide a quantitative assessment of the factors that may account for the initial rise of mass emigration to the United States. I evaluate demographic and economic fundamentals considered in most literature on the Age of Mass Migration (see, for example, Hatton & Williamson, 1993; Hatton, 1995a; Hatton & Williamson, 1998; Quigley, 1972; Sánchez-Alonso, 2000a). I also assess other forces that may have been important for the Mexican case.

3.5.1 Empirical strategy

I use an equation similar to Crozet (2004); Flores et al. (2013); and Ramos (2016). They provide a framework to understand the determinants of bilateral migration flows between countries. However, in the MBCRs data on return migration (inflows to Mexico) is scarce and non-representative. For this reason, the empirical strategy captures the effect of factors explaining differences in gross emigration across corridors. I choose a double log specification with the dependent variable expressed as natural logarithm of migration for two reasons. First, when emigration rate is used—as the dependent variable—in single cross-sectional econometric models, estimations are found to fail the test for functional form. Specially, if there is a wide spread in the measured emigration rates and since these are bounded below at zero (Hatton & Williamson, 1998, p. 260). In Chapter 2, I show that emigration rates varied widely within and across Mexican regions. Second, the logarithmic form of migration can be easily interpreted. The coefficients unambiguously reflect elasticities of Mexican migration with respect to each independent variable. In addition, comparability between different time periods is assured, and the coefficients of the independent variables are also easily interpreted (Quigley, 1972, p. 120). The basic estimating equation is:

$$ln\, M_{ij} = \alpha + \beta_1\, ln\, D_{ij} + \beta_2\, ln\, Pop_i^{MX} + \beta_3\, ln\, Pop_j^{US} + e_{ij}. \tag{3.1}$$

Gross emigration between the Mexican municipality i and US county j, M_{ij}, is a function of: the geographic distance between the origin and destination, D_{ij}; and the population sizes in such locations, Pop_i^{MX} and Pop_j^{US}. In Equation 3.1, distance arises as the cost for moving: *ceteris paribus* the greater the distance, the greater are the resources needed to emigrate, and consequently the smaller is the outflow of migrants. At the time even if railways were accessible, a migrant will often travel 480 km or more on foot to reach the border (Clark, 1908, p. 478). Distance is measured as kilometers and represent the shortest walking route that a migrant might have followed. As Poot et al. (2016), I computed the distance using Google Maps because the estimated walking route controls for Mexican and American ruggedness.

Since employment rates in the origin and destination are not available, I use population sizes (Pop_i^{MX} and Pop_j^{US}) as a proxy for the probability of employment. According to Crozet (2004, p. 440), agglomeration may occur because the access to markets positively influences the location choices of workers. However, depending on the context of the country, large populations may be a source of labor frictions (unemployment) as well. I estimate Equation 3.1 using ordinary least squares (OLS) and clustering the standard errors at the state-level because—as I argued in the previous section—Mexican emigration followed local dynamics rather than regional.

To consider additional factors influencing emigration, I expand Equation 3.1 as follows:

$$\begin{aligned} ln\, M_{ij} = {} & \alpha + \beta_1\, ln\, D_{ij} + \beta_2\, ln\, Pop_i^{MX} + \beta_3\, ln\, Pop_j^{US} + \beta_4\, Wage_{ij} + \beta_5\, Cont_i \\ & + \beta_6\, ln\, Network_j + \beta_7\, Drought_i + \beta_8\, HDI_i + e_{ij}. \end{aligned} \tag{3.2}$$

Where $Wage_{ij}$ is the US-Mexico relative wage, expressed in the form of: $Wage_{ij} = ln\,(wage_j^{US}) - ln\,(wage_i^{MX})$. This variable captures the migrant's comparison of future expected incomes at home and abroad (Hatton, 2010, p. 943). The relative wage is expected to have a positive and significant effect on the size of the migration corridors, considering the important differences in wage levels between both countries.

$Cont_i$ is an indicator variable for Mexican municipalities sharing border with the United States. Residing in a contiguous municipality may have influenced emigration to the United States because geographic distance has little relevance to cross the border from these municipalities (Flores et al., 2013, p. 203). $Network_j$ is the stock of Mexican immigrants registered in the 1900 US Census (foreign population born in Mexico), and it captures the access to friends, relatives, and a familiar community. Immigrant networks represent assistance to new migrants: they facilitate the adjustment or assimilation for newcomers at the destination (Hatton, 2010; Flores et al., 2013).

$Drought_i$ is the Palmer Drought Severity Index and captures the relative dryness in the immigrant's municipality of last permanent residence. In other words, it controls for the presence of droughts. It is a standardized index that spans from -6 (dry) to +6 (wet). Values below -4.0 represent extreme droughts while values above +4.0 represent extreme wet spells.[19] The data come from Stahle et al. (2016), who provide reconstructions of the self-calibrating Palmer Drought Severity Index (PDSI) on a 0.5° latitude/longitude grid centered over Mexico from AD 1400-2012. This variable captures climate shocks influencing the decision to emigrate.

HDI_i is a Quasi-Human Development Index at the state-level in 1910. I assign to each Mexican municipality i the index value of the state to which it belongs.

[19]The DSI was estimated from 1901 to 1911 at a municipality level. The values were assigned to each immigrant according to the year of crossing and last permanent residence. The form I use in the regressions is $DSI_{i,t} = DSI_{i,t-1}$ with $t \in \{1906, 1907, 1908\}$.

The index is the simple average of standardized health, education and income proxies; and it captures differences in economic development within Mexican regions through variables others than wage.[20] The data come from Campos-Vázquez & Vélez-Grajales (2012, p. 611), who study living standards during the Porfiriato. Table 3.7 in Section 3.8 presents summary statistics of the previous variables.

3.5.2 Results: national-level drivers

I start my analysis considering the three years as a pooled cross-sectional sample. Column 1 of Table 3.5 shows that the estimated coefficient β_1 is significant and has the expected sign. An increase of one percent in the distance reduces the migration flow by 0.46%. The population sizes are significant, and their effects fit in the context of the research. On the one hand, a one percent increase in the population of origin would lead the migration flow to increase by 0.26%. This result is consistent with the fact that population in both countries was growing at the time, but living standards in Mexico were deteriorating. In consequence, the growing population in Mexico represented a more constrained society with higher incentives to emigrate. On the other hand, populated cities in the United States represented for Mexican laborers a higher probability to be employed with superior salaries and better living conditions. Precisely, an increase of one percent in the US population size leads to an increase of 0.15% in the flow. These initial findings suggest that migration costs were more important that conditions at home. It is difficult to compare the coefficient on distance cost since migration costs have been absent in most cliometric studies on the Age of Mass Migration (Hatton, 2010, p. 944).

[20] The variables used in the Quasi-Human Development Index are: health (number of physicians per 10 thousand people), education (school enrollment and literacy rates) and income (urbanization rates - proportion of population living in places with more than 2,500 people). The authors follow the estimation method of the standard Human Development Index.

Table 3.5: Determinants of Mexican migration to the United States (1906–08). Dependent variable: Gross emigration (lnM_{ij})

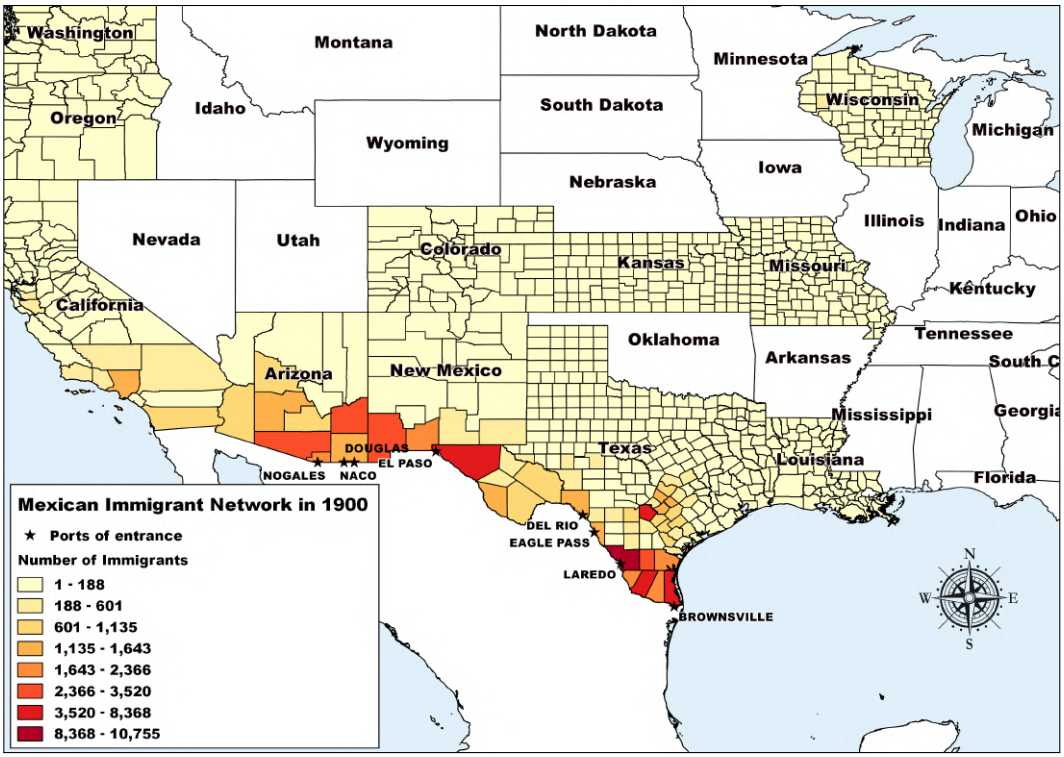
	Complete Sample					Bajio	Border
Independent Variables	1	2	3	4	5	6	7
ln Distance	-0.457***	-0.493***	-0.439***	-0.406***	-0.447***	-0.315	-0.799***
	(0.117)	(0.086)	(0.097)	(0.102)	(0.098)	(0.282)	(0.099)
ln Population^{MX}	0.261***	0.264***	0.282***	0.286***	0.300***	0.236***	0.370***
	(0.077)	(0.077)	(0.067)	(0.058)	(0.067)	(0.037)	(0.067)
ln Population^{US}	0.147***	0.136***	0.099***	0.086**	0.094**	-0.011	0.130**
	(0.046)	(0.043)	(0.035)	(0.040)	(0.037)	(0.102)	(0.047)
ln Wage$^{US/MX}$		0.156	0.081	0.164*	0.165*	0.341	-0.033
		(0.134)	(0.120)	(0.094)	(0.094)	(0.201)	(0.101)
ln Network^{1900}			0.176***	0.191***	0.182***		
			(0.034)	(0.032)	(0.033)		
Drought				-0.134**	-0.129**	-0.072	-0.149
				(0.054)	(0.053)	(0.164)	(0.079)
HDI^{MX}					-0.793	-1.539	1.527
					(0.974)	(1.134)	(2.412)
Contiguity					-0.070		-0.118
					(0.142)		(0.109)
ln Network^{1900} × Distance						0.370***	0.101***
						(0.039)	(0.023)
Constant	0.560	0.585	-0.735	-1.140**	-0.843	-4.229*	-0.001
	(0.758)	(0.785)	(0.696)	(0.529)	(0.619)	(1.849)	(0.724)
Observations	892	892	866	866	866	258	570
R-squared	0.073	0.076	0.123	0.134	0.136	0.144	0.220

Notes: * = Significant at 10% level; ** = Significant at 5% level; *** = Significant at 1% level.
Robust standard errors in parenthesis clustered at a state-level.
Distance = Geographic distance between origin and destination (kilometers).
Network = Stock of Mexican immigrants registered in the 1900 US Census.
Drought = $DSI_{i,t-1}$ with $t \in \{1906, 1907, 1908\}$. The values were assigned to each immigrant according to the year of the crossing and last permanent residence, and then collapsed by migration corridor.
HDI = Quasi-Human Development Index at the state-level (Mexico) in 1910.
Contiguity = Dummy variable for Mexican municipalities sharing border with the United States.

Column 2 of Table 3.5 shows that the relative wage do not explain differences in size across migration corridors. However, the fundamental push and pull factors remain significant. This result is consistent with the argument of Hatton (2010, p. 943): employment probabilities take a larger weight in the emigration decision because migrants are risk averse and greater uncertainty attaches to the employment probabilities than to the wage rates. The result is also in line with previous literature finding that when including proxies for wages and employment, the latter usually dominate the regression results.[21]

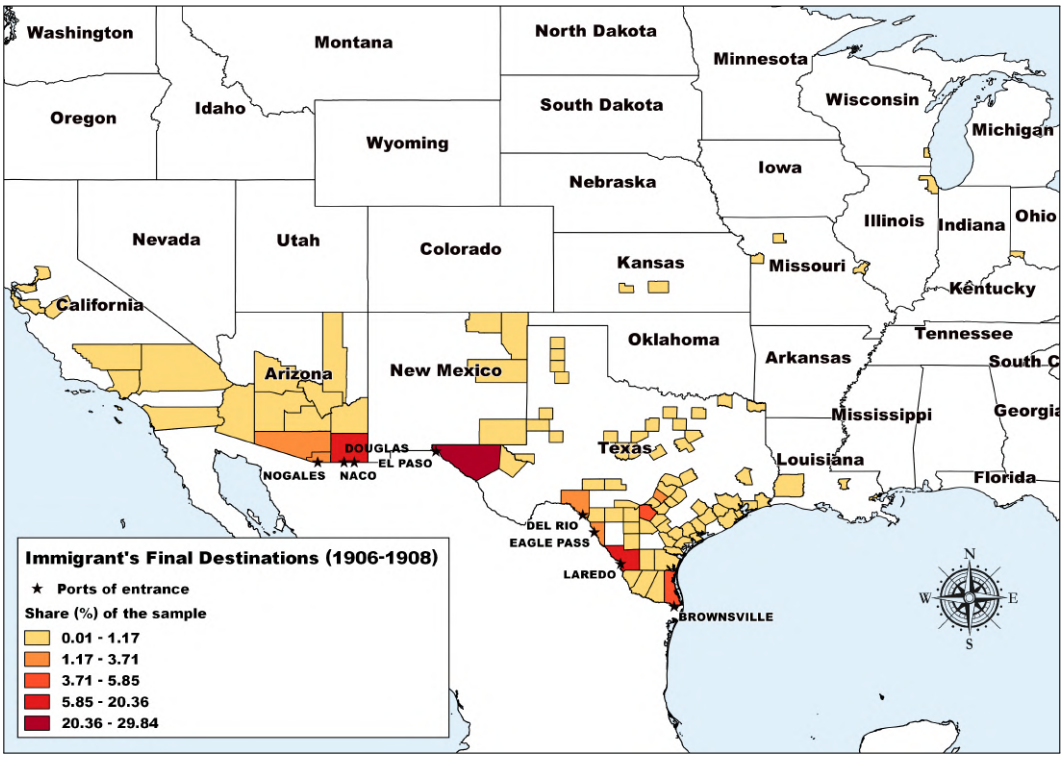
[21] See Gould (1979) for a review.

Figure 3.3: Mexican immigrant network in the United States (1900)



Source: 12th Population Census of the United States (Bureau of the Census, 1900).

Figure 3.4: Immigrant's destinations in the United States (1906–08)



Source: Mexican Border Crossing Records. Microfilm publication N° A3365.

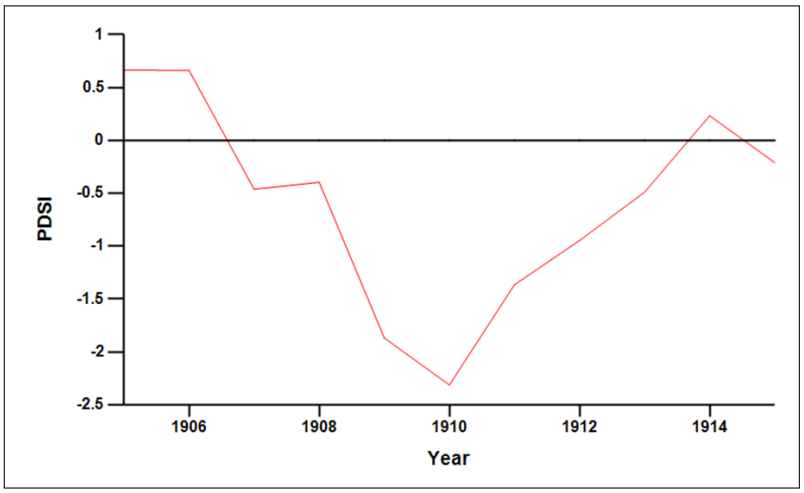
In contrast, immigrant networks influenced emigration (see Column 3). Large networks could have represented lower job and housing search costs, and lower psychological costs of being away from family and friends (Poot et al., 2016, p. 7). In this sense, a one percent increase of the Mexican community in the county of destination increases the migration flow by 0.17%. While the marginal effect seems to be small, we need to consider that Mexican emigration and consequently immigrant networks were increasing exponentially in the early twentieth century. Clark (1908, p. 520 & 521) argues that 25% to 50% of the Mexicans crossing the border settled permanently in the United States. According to Cardoso (1980, p. 36) the number of Mexican citizens living permanently in the United States increased 300% from 1900 to 1910. If we assume that the network elasticity of migration is constant, a change of this magnitude might have increased gross emigration by 51%. This powerful long-lasting effect of the immigrant stock also characterized migration streams from Europe in late nineteenth century (Hatton & Williamson, 1994). Figure 3.3 and Figure 3.4 depict the spatial distribution of the stock of Mexican immigrants in 1900 and the immigrants' destinations in the United States (1906–08), respectively. It is clear the spatial correlation of both variables: the immigrant sample concentrates in counties with the largest Mexican immigrant communities in 1900.

Differences in relative dryness across municipalities had a significant effect on the migration flow (see Column 4). An increase of one index point reduces gross emigration by 13%. Figure 3.5 shows the average DSI in Mexico during the period under analysis and confirms that the years of 1907 and 1908 were part of a drought that lasted until 1910. Figure 3.9 in Section 3.8 depicts the regions affected by this phenomena and how it intensified at the eve of the Mexican Revolution (1910–20). Although, the relative-wage variable becomes weakly significant (at the 10% level), its coefficient is numerically small relative to distance cost coefficient. This result is consistent with the general view of previous literature: short-run and cross-sectional fluctuations are largely accounted for employment proxies,

while longer term trends can be explained by slowly changing income gaps between origin and destination (Hatton & Ward, 2018, p. 3).

Column 5 of Table 3.5 reveals that living standards across Mexico—captured by the HDI—did not mattered and neither residing in a border municipality. Interestingly, the magnitudes and significance of the distance cost, population sizes and immigrant network coefficients are hardly changed, suggesting that these factors may have been the initial systemic drivers of the migration flow.

Figure 3.5: Droughts in Mexico. Palmer Drought Severity Index (1905–15)



Source: Stahle et al. (2016).
Note: The Palmer Drought Severity Index (PDSI) uses temperature and precipitation data to estimate relative dryness. It is a standardized index that spans from -6 (dry) to +6 (wet). However, values below -4.0 represent extreme droughts while values above +4.0 represent extreme wet spells.

3.5.3 Results: regional-level drivers

The previous results capture average effects across emigration regions. Hatton & Williamson (1998, p. 95–122) and Sánchez-Alonso (2000b) show that Italian and Spanish emigration was segmented, respectively. Was this the case of Mexican emigration by the end of the Age of Mass migration? Columns 6 and 7 of Table 3.5 uncover the particular push and pull factors in the Bajio and Border region.

When I estimate Equation 3.2 only for the Bajio region, the distance cost loses explanatory power. This result should be interpreted with caution. From an

statistical perspective, this result arises because—as I described previously—most migration corridors from this region ended at El Paso: there is little heterogeneity in distance across migration corridors. The concentration of immigrants at El Paso may be to both the presence of labor recruiting companies (*enagnche* agencies) and Mexican immigrant networks. The *enganche* was a system of labor recruiting practiced to transport and allocate Mexican workers across the United States (see Chapter 1 for a description). During the 1900s *enganche* agencies established at border towns, from where Mexican workers were distributed to farther locations.

To test that for Bajio immigrants income constraints to migration were relaxed by networks and/or the *enaganche*, I add an interaction term between the migrant stock and distance to Equation 3.2. The results in Column 6 show that the interaction-term coefficient is numerically large and statistically significant. This confirms that for Bajio immigrants distance becomes less important, the larger the stock of previous immigrants. In addition, the remaining explanatory variables lose explanatory power, but conditions (population size) at the origin. In their cross-sectional analysis of Italian migration, Hatton & Williamson (1998, p. 112) find a similar result: population pressures represented limited opportunities for southern Italians from 1902 to 1912.

I follow the same procedure to evaluate the determinants of emigration from the Border region. Column 7 of Table 3.5 shows that emigration from the northern states of Mexico was influenced by different factors. Border immigrants relied less on immigrant networks and more on labor market conditions at home and abroad. Indeed, the interaction-term coefficient is statistically significant but small in size relative to the Bajio estimates. The distance-cost coefficient remains numerically large and significant. This suggests that Border migrants may have financed emigration with their own resources and only used networks and institutions to travel long distances. The proximity to the border may have allowed them to observe labor market dynamics in the American Southwest and

obtain information more promptly to maximize their probability of employment. Similarly, Hatton & Williamson (1998, p. 114) show that proximity to labor markets had a potent impact on emigration rates from Italy to continental Europe during the Age of Mass Migration.

In addition, the geographic location of the Border region may have had two effects. First, the possibility to emigrate facing shorter distances favored Border immigrants to develop long-term relations with American employers to supply labor every season in the same location.[22] Hence, they depended less on Mexican immigrant networks or labor institutions (*enganche*). Second, some municipalities of the Border region experienced a constant flow of immigrants that came from locations within the region and from the Bajio as well. This internal migration might have generated market frictions, making emigration from the Border region more responsive to population pressures. Indeed, the estimated coefficient β_2 reveals that a one percent increase in the Mexican population would increase emigration by 0.37%, about 50% more than in the Bajio.

The previous results disentangle the effect of geography, markets and immigrant networks on the Mexican-American migration flow during the last years of the 1900s. These effects abstract the forces behind emigration from the Bajio and Border region, and show that Mexican emigration was segmented like in the European periphery during the last decades of the Age of Mass Migration. Undoubtedly, climate shocks (drought) may have been an important push factor; however they lose explanatory power in the regional regressions because heterogeneity temperature within regions is limited. As a robustness check, I estimated Equation 3.2 for each year of the sample. Table 3.11 in Section 3.8 shows that the main results hold. The next section addresses an additional factor that may have shaped Mexican emigration: the access to railways.

[22] On average, immigrants from the Bajio traveled 1,460 km while their Border region counterparts traveled 658 km only (see Table 3.7).

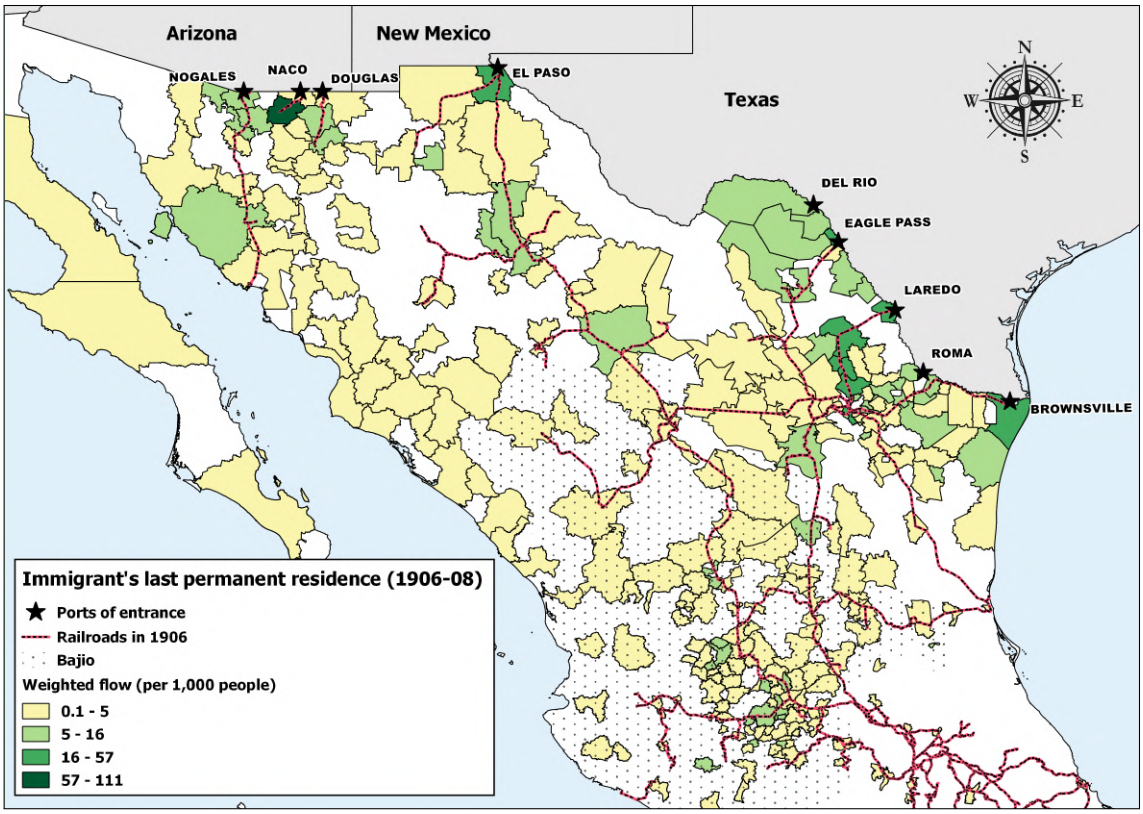
3.6 The impact of railways

The distance elasticity of migration declines over time due to communication and transport technologies (Greenwood, 1997). The access to transportation infrastructure facilitates labor mobility since migrants can return home easily whenever needed Banerjee et al. (2012, p. 10). However, the impact of railways on Mexican emigration—in the beginnings of the twentieth century—is not obvious. Although railways reduced migration costs—measured as time units—they were not accessible for all immigrants. Figure 3.6 shows that in 1906 railways crossed municipalities with the highest migration rates, but most of the Mexican West coast, the Lower California and several municipalities of the Border and Bajio regions had no direct access to this transportation technology (Coatsworth, 1979, p. 941). Therefore, it is questionable whether railways were a fundamental factor to explain the mass migration of Mexicans.

3.6.1 Empirical strategy

The effect of the access to transportation infrastructure on emigration might be endogenous: railways might have arisen in response to the demand for transportation towards northern Mexico. To correct for endogeneity, I follow Banerjee et al. (2012) and construct an instrumental variable consisting on straight lines connecting historically important cities in Mexico with the US border. The cities of Aguascalientes, Chihuahua, Colima, Durango, Guadalajara, Guanajuato, Hermosillo, Mexico City, Monterrey, Morelia, Puebla, Queretaro, San Luis Potosí, Oaxaca, Veracruz and Zacatecas were selected due to their political and economic relevance from 1790 to 1846. These historical cities were identified with the First Colonial Population Census of 1790, also known as the *Revillagigedo* Census (Castro Aranda, 2010) and with the Historical Statistics of Mexico (Instituto Nacional de Estadística Geografía e Informática, 1986).

Figure 3.6: Migration rates (1906–08) and Mexican railways (1906)



Source: Mexican Border Crossing Records. Microfilm publication number A3365.
Note: Spatial distribution of the Mexican-American migration flow from July 1906 to December 1908. The polygons display the immigrant's last permanent residence (municipalities) and their migration rate per 1000 persons (quartiles calculated with Jenks natural breaks classification method). The black line represents the railways system in Mexico c.a. 1906, which connected the principal migrant-sending municipalities with the US border. The shaded area covers the states of the *Bajio* region.

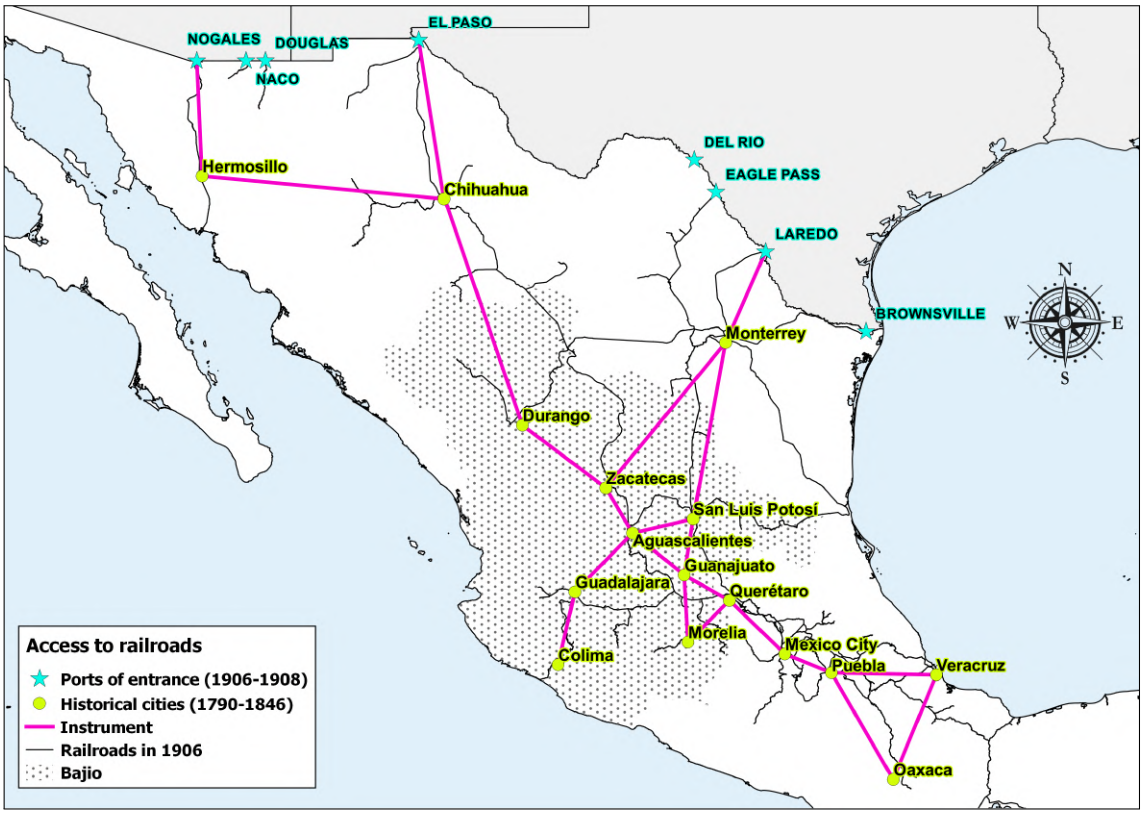
I draw the lines using the following decision rule. I draw a straight line from each historically important city to the nearest entrance port at the US border and/or to the nearest historically important city. If there were two cities or ports where the difference in distances is less than 60 km, I draw a line to both.[23] Along each straight line, I projected a train station every 30 km. Finally, I computed the distance from the centroid of each municipality, reported as the last permanent residence, to the nearest projected train station. The computation of the centroids and distances was made by QGIS. Figure 3.7 shows that the instrument (straight lines) coincide relatively well with the railways network except in the regions of

[23] I chose the 60 km criteria because it is the minimum distance needed to connect Puebla and Queretaro with Mexico City.

the Lower California and the Yucatan Peninsula, which were isolated by the Gulf of California and the Gulf of Mexico, respectively. For this reason, I excluded the municipalities belonging to these regions.

The idea behind using straight lines is that they abstract for transportation corridors that have been present since colonial times. Therefore, being on or near a straight line between two historical cities makes it much more likely that a transportation route will be present, making more feasible emigration from such locations compared to similar areas far away to the straight lines. To check if the lines proxy for transportation infrastructure, I estimate the following equation.

Figure 3.7: Historical cities (1790–1846) and straight lines as instrument (60 km criteria)



Source: The historical cities were identified with the First Colonial Population Census of 1790, also known as the *Revillagigedo* Census (Castro Aranda, 2010) and with the Historical Statistics of Mexico (Instituto Nacional de Estadística Geografía e Informática, 1986), tables 1.4.1 through 1.4.27. The ports of entrance along the US border are the ones identified in the Mexican Border Crossing Records. Microfilm publication number A3365. Note: The black line represents the railways system in Mexico c.a. 1906. The shaded area covers the states of the *Bajio* region. As expected, the instrument (straight lines) coincide well with the railways network except in the regions of the Lower California and the Yucatán Peninsula, which were relatively isolated by the Gulf of California and the Gulf of Mexico, respectively.

$$ln\, S_{1900,i} = \rho\, ln\, S_{hist,i} + e_i. \tag{3.3}$$

The natural logarithm of the distance from the centroid of each municipality i to the nearest train station as they existed in 1900, $S_{1900,i}$, is a function of: the natural logarithm of the distance to the nearest projected station, $S_{hist,i}$. The distance estimates for 1900 were kindly shared by Woodruff & Zenteno (2007).[24] Equation 3.3 should not be interpreted as the first stage of a two-stage least squares (2SLS) strategy, because the lines might proxy as well for other transportation infrastructure between the historical cities, e.g. roads. To the author's knowledge, detailed data about roads before 1910 is nonexistent.[25]

Adding $S_{hist,i}$ as explanatory variable to Equation 3.2 results in the equation:

$$ln\, M_{ij} = \alpha + \Phi\, ln\, S_{hist,i} + \mathbf{W}'_{ij}\beta + \mathbf{X}'_i\delta + \mathbf{Y}'_j\eta + e_{ij}. \tag{3.4}$$

Where $\mathbf{W}_{ij}$, $\mathbf{X}_i$ and $\mathbf{Y}_j$ are the full set covariates defined previously. The coefficient of interest is Φ, which reflects the effect of having access to a transportation corridor. This identification strategy provides me with an exogenous source of variation in access to transportation infrastructure: differences in distance from the municipality i to the nearest projected train station, $S_{hist,i}$; which goes back at least forty years before the Mexican and American railways were connected (1884) and 60 years before the period of analysis (1906–08). Hence, any economic and urbanization pattern before 1846 would have had ample chance to relocate by the beginning of the twentieth century, as argued Banerjee et al. (2012, p. 4).

[24] Figure 3.11 depicts the map used by Woodruff & Zenteno (2007) in their estimates.

[25] In Latin America, pre-railway overland transport was very precarious. Most roads were not accessible to carts, and a large share of freight transport depended exclusively on pack animals (Bignon et al., 2015, p. 1279).

3.6.2 Results

I estimate the effect of Φ using ordinary least squares (OLS) and excluding targeted poles (terminus municipalities and the straight lines' nodes) that might have influenced the railway network design (Jedwab & Moradi, 2016; Bertazzini, 2018). Table 3.6 shows that the distance to a historical transportation corridor might have been an important factor in the decision to emigrate. Considering the complete sample, the access to transportation elasticity of migration is about 25% of the distance elasticity of migration. This result suggests that, the access to railways was costly at the turn of the twentieth century. On average, migrants had to travel 45 km to the nearest station (see Table 3.7 in Section 3.8).

The regional results reveal that in the Bajio (column 2), the distance to a transportation corridor has a large and significant effect on the migration flow: increasing the distance to a transportation corridor by one percent would reduce the migration flow by 0.22 percent. This suggests that for Bajio migrants, moving costs did not arise from long distances but from the access to transportation towards the US border. In other words, railways might have been fundamental to explain emigration from this region.

In the Border region (column 3), increasing the distance to a transportation corridor by one percent would reduce the migration flow by 0.06 percent (one third than in the Bajio). But, since the distance elasticity of migration remains significant, emigrate by train from this region might have implied additional costs. Moreover, the access to transportation elasticity of migration is only 13 percent of the distance elasticity of migration, confirming that either: railways were not accessible in all migrant-sending municipalities and/or they might not have been necessary to emigrate because they did not reduce migration costs significantly. As Coatsworth (1979, p. 940) argues: "Mexicans did ride trains, but not because they were much cheaper than walking".

Table 3.6: Impact of the access to railways on Mexican migration (1906–08). Dependent variable: Gross emigration (lnM_{ij})

Variables	60 km criteria			100 km criteria		
	Complete sample	**Bajio**	**Border**	**Complete Sample**	**Bajio**	**Border**
	1	2	3	4	5	6
ln Distance to straight line (km)	-0.099*	-0.228*	-0.067**	-0.099*	-0.194	-0.060*
	(0.054)	(0.119)	(0.021)	(0.049)	(0.106)	(0.025)
ln Distance to final destination (km)	-0.390**	0.037	-0.487**	-0.395**	0.033	-0.489**
	(0.156)	(0.330)	(0.174)	(0.155)	(0.259)	(0.175)
Migration corridor-specific covariates	YES	YES	YES	YES	YES	YES
Municipality-specific covariates	YES	YES	YES	YES	YES	YES
County-specific covariates	YES	YES	YES	YES	YES	YES
MX State FE	YES	YES	YES	YES	YES	YES
Observations	642	197	433	642	197	433
R-squared	0.176	0.188	0.176	0.176	0.186	0.176

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Notes: * = Significant at 10% level; ** = Significant at 5% level; *** = Significant at 1% level. Robust standard errors in parenthesis clustered at a state-level. Targeted poles excluded (terminus municipalities and the straight lines' nodes). Following Equation 3.4:
lnM_{ij} = Weighted emigration
$S_{hist,i}$ = Distance to straight line
$\mathbf{W}_{ij}$ = Migration corridor-specific covariates: US-Mexico wage gap.
$\mathbf{X}_i$ = Municipality-specific covariates: Population size, development index and temperature.
$\mathbf{Y}_j$ = County-specific covariates: Population size and Mexican immigrant stock in 1900.

As a robustness check, I used an expanded set of straight lines based on a 100 km maximum difference in distances criteria when connecting two historical cities (see Figure 3.10 in Section 3.8). The econometric results are similar in magnitude and significance for the complete sample and the Border region, showing that the number of lines do not drive my results (columns 4 to 6). Nevertheless, the estimates are not significant for the Bajio region.

These results do not necessarily contradict previous historical research arguing that railroads were a fundamental factor to explain the mass emigration of Mexicans to the United States before 1910, but disentangle the real influence they might have had on the migration flow based on micro data. Future research could deepen in this matter by assessing if the Mexican peasantry could afford a train journey at the time (Kosack & Ward, 2014, p. 1022). Finally, even though Mexican railroads connected the principal cities of the Border and Bajio region with the United States, other stories overcoming the structure of the Mexican

railways system might be considered to study the push and pull factors of the Mexican-American migration flow.

3.7 Conclusions

Exploiting a novel dataset consisting of individual border crossings from 1906 to 1908, in this chapter I refine our knowledge on the initial determinants of the Mexican emigration to the United States. To assess the push and pull factors of gross emigration, I identify forces that could have influenced the emigration decision at the local level. The results suggest that differences in economic conditions and differences in immigrant networks at the United States were the main pull factors rather than differences in wage rates between countries. Economic conditions at home and regional droughts experienced at the eve of the Mexican Revolution (1910) pushed Mexican peasants to emigrate. The region-level analysis reveals the existence of two emigration models. In the Border region, distance costs and market-oriented incentives (market potentials) influenced the decision to emigrate. In contrast, the flow from the Bajio was determined by labor-market pressures in Mexico and immigrant networks.

Despite its importance for the Mexican economy, railways had a limited effect on Mexican emigration at the turn of the twentieth century. The estimates suggest that the access to railways induced emigration, but only in the Bajio. Therefore, railways did not reduce migration costs significantly or/and they were not accessible for the average migrant. This finding supports Coatsworth (1979) view: "Mexicans did ride trains, but not because they were much cheaper than walking". In other words, railroads were more a catalyst than a determinant, in the sense that Mexican emigration would have occurred anyways due to the presence of other forces.

The policy implications of these results are worth to highlight. First, differences in wage gaps were not the main driver, but differences in the stock of Mexican

immigrants at the destination. Indeed, as Massey & Espinosa (1997) found for the 1987-1992 period, the flow was driven by the social capital formation in the destination. In other words, Mexican immigrant networks have been a self-perpetuating social asset that provides information and assistance, which reduces the costs and risks associated to emigration. Therefore, the persistence of immigrant networks as the main driver of the flow, questions if the convergence of US and Mexican real wages would be an effective mechanism to reduce or control the Mexican emigration to the United States.

Second, in the Bajio (the poorest Mexican region at the time), emigration might have been influenced by immigrant networks and/or labor institutions. These factors made the emigration decision less income constrained. Nowadays, the Bajio states present as high living standards as northern Mexico (Campos-Vazquez et al., 2017). This regional convergence in economic development might have occurred through persistent flows of remittances, and the diffusion of values and behaviors acquired in the United States (Pérez-Armendáriz & Crow, 2010). Therefore, persistent emigration could be an effective mechanism to reduce development asymmetries within sending countries.

Third, like in 1908, droughts induced emigration from 1995 to 2002 (Chort & De La Rupelle, 2016). This confirms that in rural Mexico, emigration remains a mechanism that neutralizes the effect of climate shocks on the household's income.

Finally, the Mexico-United States migration has been influenced by forces that are commonly not analyzed by policy makers. An integral migratory policy should consider the different incentives behind the emigration decision as well as their evolution along time and across Mexican regions. Only then, both countries will maximize the benefits of labor migration and minimize the problems derived from it.

3.8 Appendix

A. Summary statistics

Table 3.7: Determinants of Mexican migration to the United States (1906–08). Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
Complete Sample (892 corridors)				
Flow of Immigrants	17	66	1	1264
Distance (km)	966	686	27	4,587
Population in Mexico	39,861	103,803	1,099	720,753
Population in the US	72,615	244,119	561	2,762,522
Mexican wages	0.19	0.04	0.13	0.26
American wages	1.45	0.46	0.20	2.90
Migrant Stock in 1900	3,601	3,533	0	10,755
Drought Severity Index				
1906	0.59	1.00	-1.43	4.88
1907	-0.78	0.98	-2.27	2.49
1908	-0.93	1.25	-3.22	2.85
Distance to Train Stations in 1900 (km)	44.7	54.5	0.5	667.6
Distance to straight lines - IV (km)	90.7	89.8	0.85	709.1
Border Region (586 corridors)				
Flow of Immigrants	17	75	1	1264
Distance (km)	658	496	27	4,587
Population in Mexico	19,614	21,957	1,099	86,294
Population in the US	62,327	225,109	561	2,762,522
Mexican wages	0.21	0.02	0.20	0.24
American wages	1.43	0.50	0.20	2.25
Migrant Stock in 1900	3,027	3,262	0	10,755
Drought Severity Index				
1906	0.52	0.96	-1.29	4.07
1907	-0.97	1.05	-2.27	2.49
1908	-1.46	1.05	-3.22	1.10
Distance to Train Stations in 1900 (km)	46.1	53.6	0.5	462.3
Distance to straight lines - IV (km)	112.1	94.4	3.0	532.1
Bajio Region (266 corridors)				
Flow of Immigrants	19	44	1	424
Distance (km)	1,460	528	533	4,338
Population in Mexico	34,957	30,097	2,232	123,506
Population in the US	82,910	254,076	1,255	2,762,522
Mexican wages	0.15	0.04	0.13	0.24
American wages	1.48	0.40	0.20	2.90
Migrant Stock in 1900	4,901	3,758	0	10,755
Drought Severity Index				
1906	0.56	0.84	-1.11	4.88
1907	-0.54	0.55	-2.05	2.26
1908	-0.07	0.85	-2.97	1.17
Distance to railway stations in 1900 (km)	36.8	39.8	0.5	161.2
Distance to straight lines - IV (km)	45.1	44.0	0.85	148.3

Figure 3.8: Expansion of the Mexican railway network (1884–1910)

(a) 1884

(b) 1910

Source: Cosío Villegas (1974). Most of the current south-north Mexican railways were constructed during the Porfiriato (1877–1911). In fact, the mileage increased from 477 km in 1877 to 19,000 km in 1910 (Cosío Villegas & Bernal, 1973; Henderson, 2011).

Table 3.8: Minimum salary in Mexico by economic activity (1877–1911). Cents per day (US dollars)

	All Sectors		Agriculture		Manufactures		Mining	
Year	*Current prices*	*1900 prices*	*Current prices*	*1900 prices*	*Current prices*	*1900 prices*	*Current prices*	*1900 prices*
1877	11	16	11	16	11	16	11	16
1885	11	14	11	13	14	17	13	15
1892	15	14	14	13	16	13	16	15
1898	17	19	15	18	19	25	20	23
1902	18	16	17	16	20	18	23	21
1911	24	15	22	13	29	18	59	36

Source: Rosenzweig (1965, p. 447).
Note: The stagnation of real wages, specially in the agriculture sector characterized the Porfirian period (1877–1911). The mining salaries are the exception since they presented a considerable growth from 1898. However, for most of the country this stagnation was translated to important differences in wage levels between Mexico and the US.

Table 3.9: Growth of regionally-adjusted wages by sector and region. Average annual growth rates (1900–08)

	North	Gulf	North Pacific	South Pacific	Center
Agriculture	0.20	2.21	0.77	2.36	-1.47
Industry	2.01	0.10	1.23	-0.24	-1.52
Mining	6.08	-1.32	5.57	4.79	3.53

Source: Arnaut (2018, p. 53).
Note: Adjusted with regional deflators (1900 = 100). The annual growth rates by region confirm that real wages in the agriculture sector were stagnated and deteriorated in the North and Center, respectively. Most Mexican migrants came from these regions. However, mining salaries presented a considerable growth in last decade of the Porfiriato (1877-1911). The regions are the following. **North:** Coahuila, Tamaulipas, Chihuahua, Nuevo León, San Luis Potosi*, Durango* and Zacatecas*. **Gulf:** Yucatán, Campeche, Veracruz and Tabasco. **North Pacific:** Baja California, Sonora, Tepic and Sinaloa. **South Pacific:** Colima*, Chiapas, Guerrero and Oaxaca. **Centre:** Mexico City, Morelos, Aguascalientes*, Puebla, Querétaro, Tlaxcala, Hidalgo, Estado de México, Guanajuato*, Jalisco* and Michoacán*. * *Bajio* states.

C. County-level wages in the United States

Table 3.10: US wages by county (1900–14). US dollars per day (current prices)

Location	County	Sector	Source	Type[1]	Wage	Year
			Arizona			
Flagstaff	Coconino	Manufactures	Clark (1908, p. 494).	M	2	1907
Salt River	Gila	Agriculture	Clark (1908, p. 485).	M	2	1908
Gila	Gila	Agriculture	Hudson (1914, p. 8)	G	2.25	1914
Maricopa	Maricopa	Agriculture	Hudson (1914, p. 8)	G	2.25	1914
Patagonia	Santa Cruz	Mining	Clark (1908, p. 490).	M	2	1908
			California			
Alameda	Alameda	Agriculture	Flint (1900, p. 22).	G	1.5	1900
Southern California	Imperial	Agriculture	Clark (1908, p. 485).	M	2	1908
Bakersfield	Kern	Manufactures	Clark (1908, p. 479).	M	1.25	1907
Southern California	Los Angeles	Agriculture	Clark (1908, p. 485).	M	2	1908
Los Angeles	Los Angeles	Manufactures	Clark (1908, p. 495).	M	1.25	1907
Los Angeles	Los Angeles	Manufactures	DCL (1907), pp. 143–4.	G	1.78	1906
Mendocino	Mendocino	Agriculture	Flint (1900, p. 22).	G	1.5	1900
Southern California	Orange	Agriculture	Clark (1908, p. 485).	M	2	1908
Sacramento	Sacramento	Agriculture	Flint (1900, p. 22).	G	1.5	1900
Southern California	San Bernandino	Agriculture	Clark (1908, p. 485).	M	2	1908
Southern California	San Diego	Agriculture	Clark (1908, p. 485).	M	2	1908
San Francisco	San Francisco	Manufactures	Clark (1908, p. 494).	M	2	1907
San Francisco	San Francisco	Manufactures	DCL (1907), p. 143–4.	G	1.99	1906
San Joaquin	San Joaquin	Agriculture	Flint (1900, p. 22).	G	1.5	1900
Southern California	Santa Barbara	Agriculture	Clark (1908, p. 485).	M	2	1908
Sonoma	Sonoma	Agriculture	Flint (1900, p. 22).	G	1.5	1900
Yolo	Yolo	Agriculture	Flint (1900, p. 22).	G	1.5	1900
Yuba	Yuba	Agriculture	Flint (1900, p. 22).	G	1.5	1900
Southern California	Ventura	Agriculture	Clark (1908, p. 485).	M	2	1908
			Colorado			
Denver	Denver	Manufactures	Clark (1908, p. 479).	M	1.75	1907
Sugar City	Crowley	Agriculture	Clark (1908, p. 483).	M	1.2	1908
Rocky Ford	Otero	Agriculture	Clark (1908, p. 483).	M	1.2	1908
Trinidad	Las Animas	Mining	Clark (1908, p. 488).	M	1.5	1908

Continued

Location	County	Sector	Source	Type[1]	Wage	Year
			Illinois			
Chicago	Cook	Manufactures	DCL (1907), pp. 143–4.	G	1.52	1906
			Iowa			
Fredonia	Louisa	Manufactures	Clark (1908, p. 479).	M	1.5	1907
			Kansas			
Topeka	Shawnee	Manufactures	DCL (1907), p. 143–4.	G	1.44	1906
			Lousiana			
New Orleans	Orleans	Manufactures	DCL (1907), p. 143–4.	G	1.4	1906
			Missouri			
Kansas City	Jackson	Manufactures	Clark (1908, p. 479).	M	1.4	1907
Kansas City	Jackson	Manufactures	DCL (1907), p. 143–4.	G	1.5	1906
St. Louis	St. Louis	Manufactures	DCL (1907), p. 143–4.	G	1.48	1906
			New Mexico			
Silver City	Grant	Mining	Clark (1908, p. 486).	M	2	1908
San Antonio	Socorro	Mining	Clark (1908, p. 486).	M	2	1908
Garfield	Garfield	Mining	Clark (1908, p. 488).	M	1.83	1907
Gallup	Mckinley	Mining	Clark (1908, p. 489).	M	2.9	1908
Alburquerque	Bernalillo	Manufactures	Clark (1908, p. 494).	M	1.25	1907
			New York			
New York	New York	Manufactures	DCL (1907), p. 143–4.	G	1.42	1906
			Ohio			
Cleveland	Cuyahoga	Manufactures	DCL (1907), p. 143–4.	G	1.34	1906
Cincinnati	Hamilton	Manufactures	DCL (1907), p. 143–4.	G	1.28	1906
Toledo	Lucas	Manufactures	DCL (1907), p. 143–4.	G	1.51	1906
			Oklahoma			
Oklahoma	Oklahoma	Manufactures	Clark (1908, p. 479).	G	1.5	1907
Oklahoma	Oklahoma	Manufactures	DCL (1906a), p. 32.	G	1.87	1905
			Texas			
Eastern Texas	Anderson	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Angelina	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Cherokee	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Northern Texas	Clay	Manufactures	Clark (1908, p. 479).	M	1.15	1907
Northern Texas	Collin	Manufactures	Clark (1908, p. 479).	M	1.15	1907

Continued

Location	County	Sector	Source	Type[1]	Wage	Year
Northern Texas	Cooke	Manufactures	Clark (1908, p. 479).	M	1.15	1907
Northern Texas	Denton	Manufactures	Clark (1908, p. 479).	M	1.15	1907
Northern Texas	Grayson	Manufactures	Clark (1908, p. 479).	M	1.15	1907
Eastern Texas	Gregg	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Palestine	Anderson	Manufactures	DCL (1906b), p. 46.	G	1.82	1905
Eastern Texas	Harrison	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Henderson	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Houston	Manufactures	Clark (1908, p. 479).	M	1.5	1907
San Antonio	Bexar	Manufactures	DCL (1906b), p. 46.	G	1.77	1905
San Antonio	Bexar	Manufactures	DCL (1907), p. 143–4.	G	1.16	1906
Eastern Texas	Jasper	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Kaufman	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Marion	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Northern Texas	Montague	Manufactures	Clark (1908, p. 479).	M	1.15	1907
Eastern Texas	Nacogdoches	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Dallas	Dallas	Manufactures	DCL (1906b), p. 46.	G	1.63	1905
Dallas	Dallas	Manufactures	DCL (1907), p. 143–4.	G	1.52	1906
Eastern Texas	Newton	Manufactures	Clark (1908, p. 479).	M	1.5	1907
El Paso	El Paso	Manufactures	DCL (1906b), p. 46.	G	2.01	1905
Eastern Texas	Panola	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Galveston	Galveston	Manufactures	DCL (1906b), p. 46.	A	1.72	1905
Eastern Texas	Polk	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Denison	Grayson	Manufactures	DCL (1906b), p. 46.	G	2	1905
Sherman	Grayson	Manufactures	DCL (1906b), p. 46.	G	1.85	1905
Eastern Texas	Rusk	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Sabine	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	San Agustine	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Houston	Harris	Manufactures	DCL (1906b), p. 46.	G	1.83	1905
Eastern Texas	San Jacinto	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Shelby	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Smith	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Fort Worth	Terrant	Manufactures	Clark (1908, p. 479).	M	1.75	1907
Eastern Texas	Trinity	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Beaumont	Jefferson	Manufactures	DCL (1906b), p. 46.	G	1.86	1905
Eastern Texas	Upshur	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Eastern Texas	Van Zandt	Manufactures	Clark (1908, p. 479).	M	1.5	1907

Continued

Location	County	Sector	Source	Type[1]	Wage	Year
Eastern Texas	Walker	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Paris	Lemar	Manufactures	DCL (1906b), p. 46.	G	1.6	1905
Northern Texas	Wichita	Manufactures	Clark (1908, p. 479).	M	1.15	1907
Corsicana	Navarro	Manufactures	DCL (1906b), p. 46.	G	1.64	1905
Northern Texas	Wise	Manufactures	Clark (1908, p. 479).	M	1.15	1907
Waco	Mclennan	Manufactures	DCL (1906b), p. 46.	G	1.48	1905
Eastern Texas	Wood	Manufactures	Clark (1908, p. 479).	M	1.5	1907
Laredo	Webb	Manufactures	Clark (1908, p. 482).	M	1	1907
Southeastern Texas	Chambers	Agriculture	Clark (1908, p. 483).	M	0.6	1908
Southeastern Texas	Galveston	Agriculture	Clark (1908, p. 483).	M	0.6	1908
Southeastern Texas	Hardin	Agriculture	Clark (1908, p. 483).	M	0.6	1908
Southeastern Texas	Harris	Agriculture	Clark (1908, p. 483).	M	0.6	1908
Southeastern Texas	Jefferson	Agriculture	Clark (1908, p. 483).	M	0.6	1908
La Salle	La Salle	Agriculture	Clark (1908, p. 483).	M	0.2	1908
Southeastern Texas	Orange	Agriculture	Clark (1908, p. 483).	M	0.6	1908
Tyler	Smith	Manufactures	DCL (1906b), p. 46.	G	1.67	1905
Southeastern Texas	Tyler	Agriculture	Clark (1908, p. 483).	M	0.6	1908
Ward	Ward	Agriculture	Clark (1908, p. 483).	M	1	1907
Fort Worth	Terrant	Manufactures	DCL (1906b), p. 46.	G	1.89	1905
Austin	Travis	Manufactures	DCL (1906b), p. 46.	G	1.24	1905
Laredo	Webb	Agriculture	Clark (1908, p. 483).	M	0.26	1908
Eagle Pass	Maverick	Mining	Clark (1908, p. 489).	M	1.4	1908
Laredo	Webb	Mining	Clark (1908, p. 489).	M	1.4	1908
Laredo	Webb	Manufactures	DCL (1906b), p. 46.	G	1.12	1905
El Paso	El Paso	Mining	Clark (1908, p. 493).	M	1.1	1908
San Antonio	Bexar	Manufactures	Clark (1908, p. 495).	M	1.25	1907
		Wisconsin				
Milwaukee	Milwaukee	Manufactures	DCL (1907), p. 143–4.	G	1.29	1906

Notes:
DCL refers to US Department of Commerce and Labor.
1: *M* refers to wages for Mexicans and *G* to general wage levels.

D. Temporal effects

To rule out the possibility that the main results may be influenced by specific year effects, I estimate Equation 3.2 for each year of the sample. Table 3.11 shows that results in Table 3.5 hold without major discrepancies. However, the results for 1907 capture additional effects. It is important to consider that for 1906 only the last six months of the year were available. For this reason, I also present—as an additional robustness check—the results for the last six months of each year (columns 4–6 of Table 3.11).[26]

Column 2 and 5 of Table 3.11 shows that in 1907, municipalities sharing border with the United States had 54 percent more emigration than inland municipalities.[27] In March of that year, the American financial market crashed as a result of a strong speculative process. This financial crisis is known as the *Panic of 1907* and it caused the bankruptcy of at least 25 banks and 17 trust companies (Bruner & Carr, 2007).[28] The railway and mining industries, where most Mexican immigrants were employed, experienced important losses because major players, such as the railway company Union Pacific and the United States Steel Corporation, saw their shares devalued by 25 dollars in a single day and suspended temporarily the payment of dividends (Markham, 2002). González (2010, p. 11) argues that as a consequence of this financial crisis, in the beginnings of 1907, around 250 Mexicans were rejected at the entrance ports and thousands were returned to the border. Thus, it is likely that this returned migrants stayed at Mexican border municipalities, from where they crossed again the border to work in activities less affected by the crisis. In this sense, the effect of *the Panic of 1907* is observed in the significance of *Contiguity*.

[26] Emigration was more intense during the spring planting in the United States and during August and September (Clark, 1908, p. 473 & 474; Cardoso, 1980, p. 26). Consequently, the six-month results might not capture the complete mechanics of the flow.

[27] $exp(0.432) = 1.54$

[28] For reviews of this event see Andrew (1908), Bruner & Carr (2007) and Markham (2002).

Table 3.11: Determinants of Mexican migration to the United States by year. Dependent variable: Gross emigration (lnM_{ij})

	Jul - Dec	**Jan - Dec**		**Jul - Dec**		
	1906	*1907*	*1908*	*1906*	*1907*	*1908*
Independent Variables	1	2	3	4	5	6
ln Distance (km)	-0.195	-0.493***	-0.555***	-0.195	-0.428***	-0.499***
	(0.167)	(0.156)	(0.093)	(0.167)	(0.120)	(0.109)
ln MX Population	0.222**	0.265***	0.278***	0.222**	0.274***	0.226***
	(0.102)	(0.083)	(0.031)	(0.102)	(0.049)	(0.057)
ln US Population	0.097	0.012	0.131**	0.097	-0.024	0.084
	(0.099)	(0.058)	(0.057)	(0.099)	(0.101)	(0.050)
ln US/MX Wage	0.559	0.406***	-0.101	0.559	0.307**	-0.015
	(0.552)	(0.108)	(0.113)	(0.552)	(0.127)	(0.134)
Contiguity	0.156	0.432***	-0.011	0.156	0.268*	-0.013
	(0.158)	(0.054)	(0.131)	(0.158)	(0.141)	(0.055)
ln Migrant Stock 1900	0.239*	0.101	0.131***	0.239*	0.150*	0.110***
	(0.120)	(0.063)	(0.021)	(0.120)	(0.087)	(0.032)
Drought (1905)	0.046			0.046		
	(0.113)			(0.113)		
Drought (1906)		0.085**			0.083**	
		(0.030)			(0.036)	
Drought (1907)			0.051			0.015
			(0.058)			(0.107)
MX Development Index	-13.365***	-5.548***	5.670***	-13.365***	-3.982***	3.041***
	(2.432)	(0.945)	(0.431)	(2.432)	(0.761)	(0.723)
Constant	-0.618	1.470**	-0.440	-0.618	0.839	.725
	(2.175)	(0.530)	(0.517)	(2.175)	(0.698)	(0.844)
Observations	223	441	490	223	313	345
R-squared	0.391	0.217	0.235	0.391	0.210	0.239

Source: Mexican Border Crossing Records. Microfilm publication number A3365.
Notes: * = Significant at 10% level; ** = Significant at 5% level; *** = Significant at 1% level.
Robust standard errors in parenthesis clustered at a state-level.
lnM_{ij} = weighted migration flow.
Contiguity = Dummy variable for Mexican municipalities sharing border with the US.
Drought = $DSI_{i,t-1}$ with $t \in \{1906, 1907, 1908\}$. The values were imputed to each immigrant according to the year of the crossing and last permanent residence, and then collapsed by migration corridor.
MX Development Index = $QHDI_i$. Quasi-Human Development Index at the state-level (Mexico) in 1910. The variables used in the QHDI dimensions are: health (number of physicians per 10 thousand people), education (school enrollment and literacy rates) and income (urbanization rates - proportion of population living in places with more than 2,500 people.

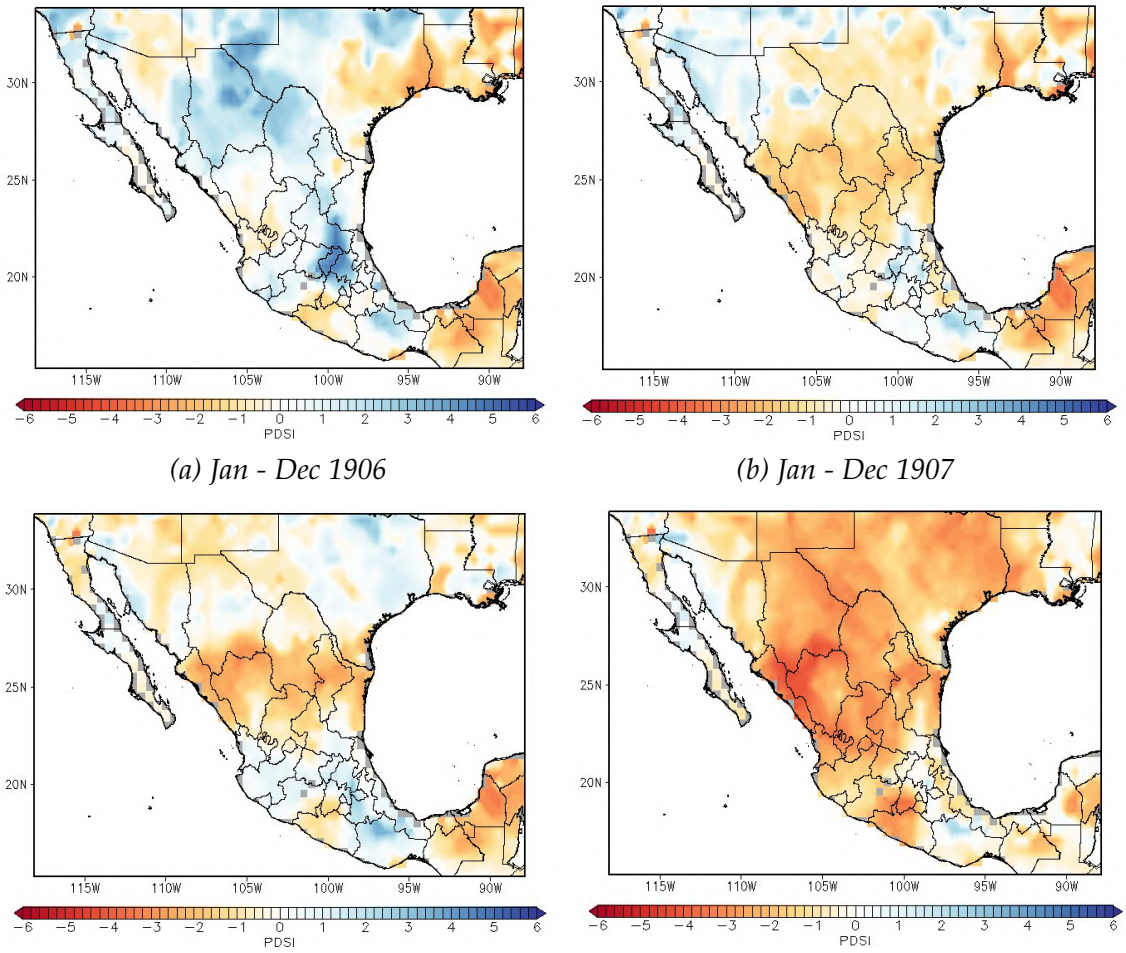
The significant effect of the relative wage ($Wage_{ij}$) can be explained by the composition of the flow. In 1907, 54 percent of the migrants moved to a county in Arizona, while in 1906 and 1908 this figure was less than 15 percent. This unusual pattern might be capturing return migration from diverse mining towns in Arizona. Due to the *Panic of 1907*, Mexican migrants might have suddenly become unemployed and had to return home. Once the labor demand resumed, they might have emigrated again. Indeed, in 1907, 65 percent of the migration flow to Arizona came from Cananea, a municipality 61 kilometers away from the

border. However, emigration from Cananea represented only 38 and 14 percent of the total flow to Arizona in 1906 and 1908, respectively. Since counties in Arizona offered the highest expected wages, the relative wage effect may be picking up a return migration effect as well (see Table 3.10).[29] Finally, the six-month results are very similar in significance and magnitude to complete sample results.

[29]The average expected wage differential in Arizona was 1.82 dollars per day, while in Texas was about 1 dollar per day.

E. Droughts

Figure 3.9: Droughts in Mexico (1906–09).
Palmer Drought Severity Index (PDSI)



(a) Jan - Dec 1906

(b) Jan - Dec 1907

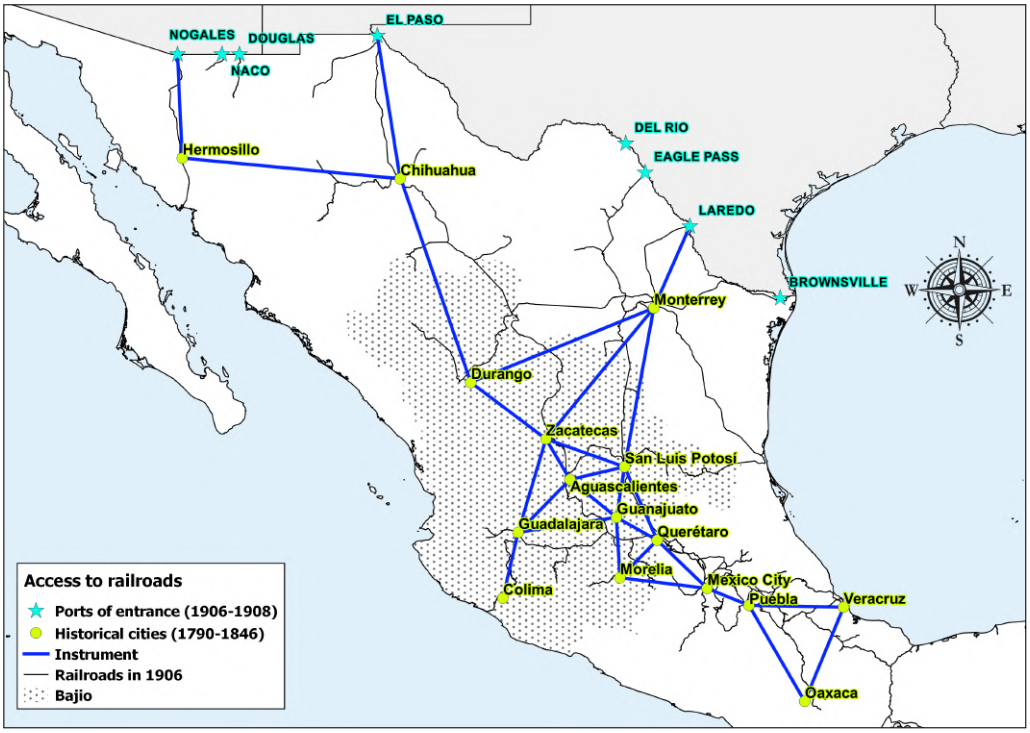
(c) Jan - Dec 1908

(d) Jan - Dec 1909

Source: Stahle et al. (2016).
Note: The Palmer Drought Severity Index (PDSI) uses temperature and precipitation data to estimate relative dryness. It is a standardized index that spans from -6 (dry) to +6 (wet). However, values below -4.0 represent extreme droughts while values above +4.0 represent extreme wet spells. The panel shows that a drought period started in 1907, which intensified in 1909 and lasted until 1910. The drought severity index in 1907 was -1.84 (Nuevo León), -0.92 (San Luis Potosí) and -0.75 (Zacatecas). This phenomena continued in 1908 with values of -1.13 (Chihuahua), -2.43 (Nuevo León), -0.78 (San Luis Potosí) and -0.46 (Zacatecas). According to the PDSI scale, these states experienced, on average, mild droughts but they were more severe in Nuevo Leon and Chihuahua. Jointly, 41% of the migration flow in 1908 had its origins in municipalities from these two states.

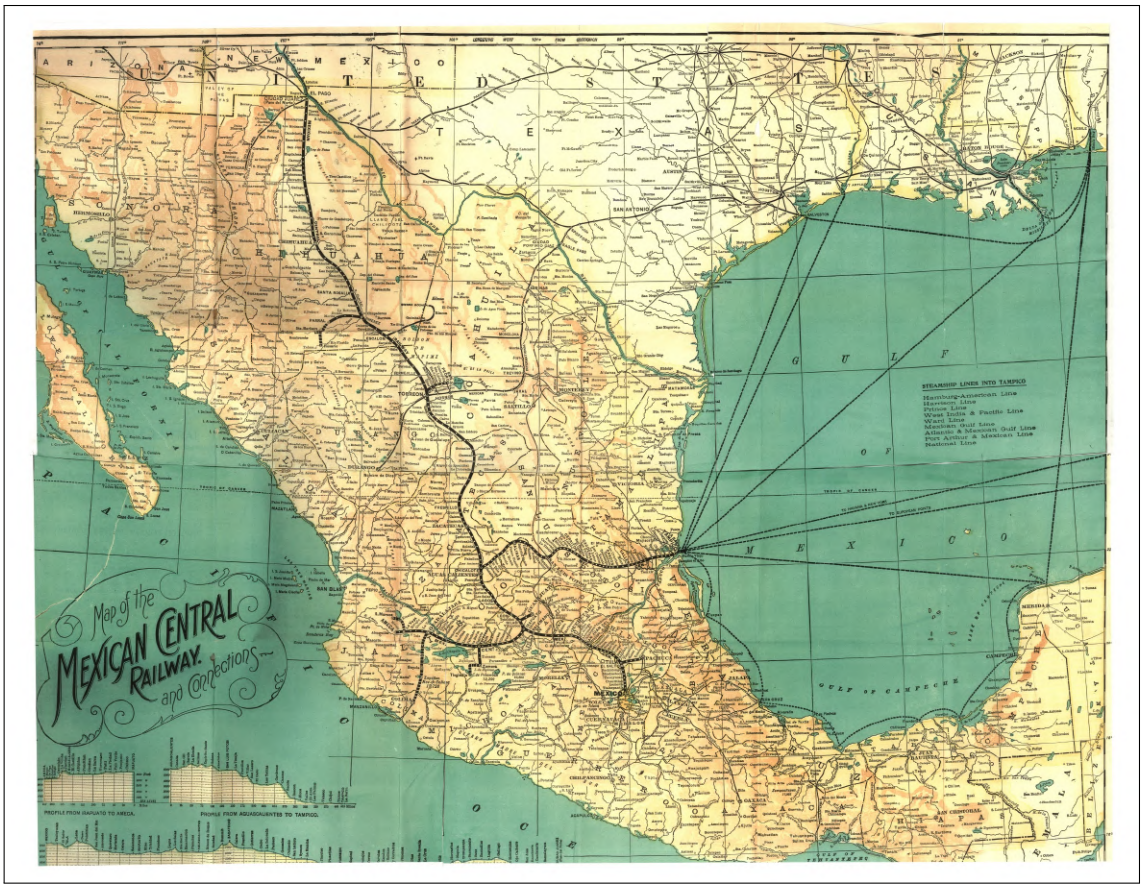
F. Instrumental variable strategy: straight lines

Figure 3.10: Historical cities (1790–1846) and straight lines as instrument (100 km criteria)



Source: The historical cities were identified with the First Colonial Population Census of 1790, also known as the *Revillagigedo* Census (Castro Aranda, 2010) and with the Historical Statistics of Mexico (Instituto Nacional de Estadística Geografía e Informática, 1986), tables 1.4.1 through 1.4.27. The ports of entrance along the US border are the ones identified in the Mexican Border Crossing Records. Microfilm publication number A3365.
Note: The black line represents the railways system in Mexico c.a. 1906. The shaded area covers the states of the *Bajio* region. I draw the straight lines (instrument) using the following decision rule. I draw a straight line from each historically important city to the nearest entrance port at the US border and/or to the nearest historically important city. If there were two cities or ports where the difference in distances is less than 100 km, I draw a line to both. As expected, the instrument (straight lines) coincide well with the railways network except in the regions of the Lower California and the Yucatán Peninsula, which were relatively isolated by the Gulf of California and the Gulf of Mexico, respectively.

Figure 3.11: Access to Railways. Train stations in 1900



Source: Map used by Woodruff & Zenteno (2007) to estimate the distance from the centroid of each municipality to the nearest train station.
Note: "Three rail lines built between 1884 and 1900 were the major means of transporting labor recruiters south into Mexico and transporting workers north to the United States. The first, the Central Mexican Railroad went south from what is now Ciudad Juarez to Irapuato in the state of Guanajuato, where it branched east to Mexico City and west through Guadalajara to Colima near the Pacific Coast. In the north, the Central Mexican Railway connected to the Southern Pacific and Texas Pacific Railroads in Texas. A second line, the Mexican International Railroad, ran a shorter distance from Durango to Piedras Negras, where it connected with the Southern Pacific Railway in Eagle Pass, Texas. Finally, the Mexican National Railroad traveled north from Mexico City through San Luis Potosi and Monterrey, reaching the border at Nuevo Laredo and Brownsville in eastern Texas. This third line was less well connected to rail lines in the United States" (Woodruff & Zenteno, 2007).

Chapter 4

Self-selection of Mexican migrants in the presence of random shocks

Evidence from the Panic of 1907

Abstract

Little is known about the speed that migrant self-selection may adjust to changes in the economic environment. Using height as a proxy for physical productivity of labor, this chapter estimates the selectivity of Mexico-US migration at the beginning of the flow (1906–08) and evaluates if self-selection patterns can change in the short run. We focus on the role of private institutions in shaping and adjusting the composition of migrants. We find that the first Mexican migrants were not negatively selected on the basis of height. Additionally, the US financial crisis of 1907—a large, unexpected and temporary shock to the demand of migrant workers—significantly modified selection into migration. Before the crisis, migrants were positively selected relative to the military elite of the time. During the crisis, migrants became negatively selected, but returned to a stronger positive self-selection after the crisis. Migrant self-selection was influenced by the *enganche*, a private labor institution that pushed toward a positive selection by neutralizing migration costs, but only for the *best* Mexicans and during *good* economic times.

4.1 Introduction

Migrants are not selected randomly from the sending population. To explain migrant self-selection, previous literature focuses predominantly on systemic drivers that are fixed in the short run and tend to change slowly over time: earnings inequality in sending and destination countries (Borjas, 1987), migration costs (Chiquiar & Hanson, 2005; Chiswick, 1999) and factors relaxing credit/liquidity constraints—for example, migrant networks (McKenzie & Rapoport, 2010; Orrenius & Zavodny, 2005). To our knowledge, changes in immigration restrictions (immigrant quotas and skill-based admission systems) are the only disruptive factors that have been studied to explain shifts in migrant self-selection (Bellettini & Ceroni, 2007; Greenwood & Ward, 2015; Massey, 2016; Spitzer & Zimran, 2018). However, immigration policy interventions may be implemented with long lags, allowing migrants to anticipate changes and adjust to them.[1] Overall, we know little about the responsiveness of migrant self-selection to changes in the economic environment.

This chapter fills this void in the literature by asking two questions. Can migrant self-selection patterns change in the short-run? And if so, through which mechanisms? To answer these questions, we study Mexican migration in the early twentieth century (1906–08) and exploit a large but temporary shock that affected unexpectedly the demand of Mexican workers. We focus on the role of private institutions in shaping and adjusting migrant self-selection. Our research is framed at the beginning of the Mexico-United States migration, because this period represents a unique setting to study selection into migration. During this period Mexican mass migration started (Cardoso, 1980; Feliciano, 2001; Gratton & Merchant, 2015) and Mexican migrants could cross the border without restrictions (Durand, 2016; Fogel, 1978; Samora, 1982). Hence, our findings are unlikely to

[1] For example, it took 25 years to pass the 1917 Immigration Act, which banned the entry of illiterate immigrants to the United States (Spitzer & Zimran, 2018, p. 236).

be influenced by under-enumeration of undocumented immigrants, and do not capture effects of immigration restrictions that artificially modify the composition of immigrant workers. Knowing the speed that migrant self-selection may change and under which conditions is as relevant as knowing the direction of self-selection. Perhaps most important, short-run shifts in migrant self-selection may impact the economic well-being of societies that depend on remittances income from abroad.

To quantify the selectivity of Mexican migrants, we use height as a proxy for physical productivity of labor.[2] Our migrant sample consists of novel historical micro data: individual border crossings registered at nine entrance ports (see Figure 4.5 in Section 4.10). We represent individuals that chose to remain in Mexico using military recruitment records of ordinary soldiers and elite forces, and passport application records. These comparison samples represent the lower, intermediate and upper ranks of Mexico's height distribution, respectively. Our empirical strategy estimates differences in height between migrants and each comparison sample to determine from which part of the height distribution the first migrants were drawn. To obtain the best results possible, we control for the individuals' region and year of birth: factors that may influence height across space and over time.

The results suggest that the first Mexican migrants were not negatively self-selected: they belonged to the upper half of the Mexican height distribution. Indeed, migrants were 2.2 centimeters taller than the average soldier; 0.5 centimeters taller than the military elite forces; and 2.1 centimeters shorter than the passport holders. Additionally, we observe considerable variation in the degree of regional selection across Mexico: migrants from poorer regions were disproportionately drawn from the upper half of the height distribution. Therefore, the

[2]Adult stature is indicative of income, health and returns to strength, especially in contexts where large sectors of the economy are not mechanized (Juif & Quiroga, 2019).

beginnings of the Mexico-US migration were characterized by an intermediate or positive selection of Mexican migrants on the basis of height.

To evaluate if migrant self-selection patterns can change in the short run, we use the Panic of 1907 as a natural experiment of history. Following Odell & Weidenmier (2004), the Panic of 1907 was determined by the 1906 San Francisco earthquake, and it was one of the most severe financial crises in the United States before the Great Depression (Frydman et al., 2015, p. 928; Moen & Tallman, 1992, p. 611; Odell & Weidenmier, 2004, p. 1003). During this nine-month crisis, the credit system of the American economy was severely affected. Banks and financial institutions of many cities limited or suspended their cash payments (Andrew, 1908, p. 497), and around two thousand firms and over one hundred state banks failed (Markham, 2002, p. 32). Although the crisis became a world-wide affair (Johnson, 1908, p. 455), in Mexico no bank collapsed or went bankrupt and there were no losses for bill holders or depositors (Gómez, 2011, p. 2095). Therefore, the Panic of 1907 provides us with exogenous variation in height across three periods: pre-Panic, Panic and post-Panic. The results suggest that, in the pre-Panic period, migrants were positively selected relative to the military elite (0.7 centimeters taller). During the Panic, migrants became negatively selected (0.2 centimeters shorter), but returned to pre-Panic levels after the crisis.

To explain the observed short-run changes in self-selection, we focus in a private institution involved in the immigration process at that time: the *enganche*. It was a system of labor recruiting used by American companies to transport and allocate laborers in the United States. We argue that Mexican migrants became less positively selected during the Panic, partially because the *enganche* was drastically reduced during this period. Indeed, the share of migrants recruited in Mexico went from 36 percent in the pre-Panic to one percent during the Panic. We provide evidence that American recruiters were rational agents that chose the tallest laborers, and thus influenced the selection of Mexican migration. On

average, *enganche* migrants were 0.9 centimeters taller than migrants that crossed the US border without using this labor institution. We show that, in the pre-Panic period, the *enganche* effect accounted for 23 percent of the difference in height between migrants and the military elite. When the Panic of 1907 hit the American financial system, companies were not able to finance the recruitment of laborers. The absence of this labor institution in combination with other factors may explain the less positive selection during the Panic. When we control for unobserved factors across states, our results reveal that in the pre-Panic period, the *enganche* effect could have accounted for 46 percent of the local (state level) self-selection pattern; and that post-Panic migrants became more positively selected relative to their pre-Panic peers.

The contributions of this chapter are three-fold. First, we extend our knowledge about the selectivity of Mexican migration to the United States. In contrast to migrants from the European periphery, who were negatively selected in the Age of Mass migration (Abramitzky et al., 2012; Cohn, 1995),[3] the first Mexican migrants were positively selected relative to the average soldier/laborer. This finding lines up with literature arguing that Mexican migrants were not drawn from the lower half of the educational, ability or height distribution (Chiquiar & Hanson, 2005; Orrenius & Zavodny, 2005; Kosack & Ward, 2014). In other words, in the beginnings of the twentieth century, Mexico sent its *best* laborers—individuals with greater physical productivity—to the United States, who in fact might have been key for the American Southwest economic expansion (Gratton & Merchant, 2015, p. 528).

Second, to our knowledge, we are the first to show that the selectivity of international migration can change in the short run in the absence of immigration restrictions. The adjustment to unexpected economic factors could be very fast, we observe significant changes in the composition of migrants in a matter of months. Shifts in our measure of selection are greater when controlling for

[3]See Abramitzky & Boustan (2017) for a review.

unobserved factors across states. In other words, the effect of the Panic of 1907 operated at the local level. This result supports previous research arguing that selection into migration is determined within sub-national environments (Abramitzky & Boustan, 2017; Spitzer & Zimran, 2018), and it suggests that the decision to migrate could be very responsive to unanticipated modifications in local conditions.

Third, we provide evidence confirming that institutions sufficiently involved in the immigration process can shape migrant self-selection (Abramitzky & Boustan, 2017, p. 1325). We speculate that the persistence of institutions neutralizing migration costs or removing liquidity constraints may explain the stickiness of migrant selection patterns over time, despite changes in immigration policy. In this sense, Jenkins (1978, p. 526) argues that after the bracero program (1942–64),[4] the Mexico-US migration flow was characterized by a contractor system with the same objective as the *enganche*: recruit, transport and allocate Mexican laborers. However, this system was based on undocumented migration. As the early twentieth century recruiters, contractors had the incentive to choose the best laborers.[5] Therefore, to assess the *quality* of documented and undocumented migrants, we should consider the persistence of private institutions, which have been shaping the selection of Mexican migrants for over 100 years.

The remainder of the chapter proceeds as follows. The next section addresses the historical context of the research and the Panic of 1907. Section 3 reviews the factors that may influence selection into migration, focusing on potential effects of unexpected changes in the demand of immigrant workers. Section 4 discusses related literature on migrant self-selection and stature as a measure of selection. Section 5 describes the data we use. Sections 6 presents our empirical strategy

[4]The bracero program was an administrative network of public and private organizations which, between 1942 and 1964, coordinated the seasonal movement of over 6 million Mexican workers for short-term agricultural employment (Jenkins, 1978, p. 525). Like the *enganche*, the program recruited, transported and allocated Mexican braceros or field hands.

[5]Using data of 52 communities in Mexico, Orrenius & Zavodny (2005) argue that undocumented immigrants are not negatively selected with regard to education.

to quantify migrant self-selection and Section 7 analyzes the results. Section 8 evaluates the role of the *enganche* in shaping migrant self-selection. Section 9 concludes.

4.2 Historical background

By the end of the nineteenth century, the United States emerged as the world´s leading manufacturing nation and its economy experienced average growth rates of 4.5 to 5 percent (Balke & Gordon, 1989; Romer, 1989; Rhode, 2002). Two factors were key to understand this robust economic growth.

First, the United States received a constant labor supply from international migration flows. From 1820 to 1920, about 36 million Europeans migrated to the United States looking for better living conditions. Mexicans joined this mass migration from 1880s, but during the 1900s Mexican migration increased sharply and expanded its geographic range of settlement (Gratton & Merchant, 2015, p. 521 & 528).[6] This was possible due to the immigration policy at the time: Mexicans were not considered immigrants who sought to settle permanently, but temporary laborers who moved back and forth supplying labor without restrictions (Fogel, 1978, p. 10; Samora, 1982).[7]

Mexican migrants were employed mainly in farms, mines and the construction of railroads across Arizona, New Mexico and Texas (Clark, 1908).[8] They were mostly rural unskilled laborers moving from the northern states and the central plateau of Mexico. The initial push and pull factors of the flow varied across regions, but living standards at the origin and immigrant networks at the

[6] In 1910, the stock of Mexican migrants in the US was about 222 thousand, and it doubled by the end of the Mexican Revolution (1910-1920) (García y Griego, 1983).
[7] Mexican migration was unrestricted before 1917.
[8] By 1909, Mexicans migrants represented 17% of the labor force of the American railway industry (Verduzco, 1995).

destination were the main determinants at the time.[9] Overall, Mexican migrants represented an inexhaustible source of cheap labor for the US economy.

Second, the American financial system was expanding rapidly by the end of the nineteenth century, facilitating investment for the creation of new firms in all economic sectors. Indeed, in 1907 there were 16 thousand financial institutions, which facilitated capital for the economy´s buoyant growth (Bruner & Carr, 2007, p. 116).[10] These institutions were small unit banks, fiduciary trust companies, clearing houses and exchange houses that provided financial services at the local level. Most of these financial institutions were supported by small companies and some others by the Bank of England and/or the United States Treasury. Furthermore, this fractioned financial system operated without a Central Bank (Bruner & Carr, 2007). This condition along with the increasing optimism, engendered by the robust performance of the economy, fueled the tendency of the public to take on more risk and invest in companies from dynamic industries, e.g. the railways and mining. Therefore, the access to capital was relatively unconstrained for the US economy.

4.2.1 The Panic of 1907: a natural experiment of history

The US economic growth was accompanied by fierce financial speculation in the first years of the twentieth century. As a sign of this phenomenon, the Dow Jones index doubled from 1904 to 1906, and at the end of 1905, the call money rates were 25 percent which were foreseen to increase to 60 percent the following year (Markham, 2002, p. 29). This speculative process occurred within a period of increasing long-term investments. National and state banks increased their bond and stock assets from 50 million in 1892 to 487 million in 1907 (Johnson, 1908,

[9] See Chapter 3.
[10] To dimension the size of the US financial system at the time, in 2007 existed 7,500 financial institutions.

p. 457).[11] This environment made the financial system fragile and limited the liquidity of the economy (Bruner & Carr, 2007, p. 115).

In April 1906, an earthquake devastated the city of San Francisco causing damages equal to more than 1 percent of the American GNP. As a consequence, extraordinary large amounts of gold flowed from London to the United States, because most of the city's insurance policies were underwritten by British companies. This was followed by defensive measures (increase of discount rates and discrimination against American finance bills) by the Bank of England and other European banks to sharply reduce the flows of gold to the United States (Odell & Weidenmier, 2004, p. 1003 & 1021). This sequence of events along with the increasing fragility of the American financial system made a market crash almost inevitable.

In March 1907, the demand for liquidity produced a wave of panic, leaving losses of 2 billion in stocks. Major players like the railway company Union Pacific saw their shares devalued by 29 percent.[12] The panic also caused the temporarily suspension of dividend payments by major mining companies such as the United States Steel Corporation (Markham, 2002, p. 29). To neutralized the panic, companies and city governments increased their bonds' interest rates. However, the wave of selling continued, pushing down stock prices.[13] As the process developed, most fiduciary institutions saw their 10 percent require reserve deposits reduced. In October, the Knickerbrocker Trust Company, the third largest trustee in New York, went into bankruptcy. This event increased the panic among the public and finally sank the financial market. Throughout August to December

[11] Similarly, trust institutions increased their holdings of debt securities by more than 500%, reaching a value of 785 million dollars in 1907 (Johnson, 1908, p. 457).

[12] In January 1906, the average price of the railroad stock was 138 dollars. In March 1907, the price fell to 98 dollars (Johnson, 1908, p. 456).

[13] This phenomenon was record by the American press throughout 1907. For instance: *"New York. Aug. 12 - The wildest break in the stock market since the present wave of selling occurred today. I carried stocks down from 1 to 17.5 points. In some cases to new low records. About one-half of the entire number of issues dealt on the exchange rate were sold at new low prices for the year."* (The Washington Post, 1907).

1907, two thousand companies went bankrupt as did 100 state banks and 30 national banks (Markham, 2002, p. 32).

This financial crisis, known as *The Panic of 1907*, was one of the most severe financial crises in the United States before the Great Depression (Frydman et al., 2015, p. 928; Moen & Tallman, 1992, p. 611; Odell & Weidenmier, 2004, p. 1003). In fact, contemporaries argued that it was "probably the most extensive and prolonged breakdown of the country´s credit mechanism which has occurred since the establishment of the national banking system" (Andrew, 1908, p. 497). The suspension of payments constrained basic transactions in all sectors, and as a consequence some industries curtailed operations and trade was considerably depressed (Frydman et al., 2015, p. 912; Johnson, 1908, p. 454).[14] To contain the impact of the Panic, substitutes for cash were emitted and rationalized to the population (Andrew, 1908), but full convertibility of deposits by the nation´s banks was not restored until January 1908 (Frydman et al., 2015, p. 912).

Although the crisis became a world-wide affair (Johnson, 1908, p. 455), in Mexico no bank collapsed or went bankrupt and there were no losses for bill holders or depositors (Gómez, 2011, p. 2095). The fact that the Panic of 1907 was influenced by the 1906 San Francisco earthquake, and that it did not affect the Mexican financial system, provides us a unique opportunity to understand how random shocks affecting the demand for immigrant labor could modify migrant self-selection in the short-run. Additionally, these events occurred in a period when Mexicans could migrate to the United States without restrictions, making it possible to quantify selection patterns without capturing any immigration policy effect.

[14]The American industrial production peaked in July 1907 and then fell 30% in the second half of the year (Hansen, 2014, p. 555).

4.3 Selection into migration

Migrant self-selection has become a prolific topic since Borjas (1987) formalized the Roy (1951) model, because it provides a framework to predict the direction of self-selection: it informs our perception about migrant *quality*. Borjas argues that migrants from countries with relatively high returns to skill and earnings inequality will be negatively self-selected: drawn from the lower half of the country-of-origin skill distribution. This is because countries with high earnings dispersion are unattractive to low-earnings workers. Therefore, workers with less-than-average productive skills would have the most to gain from moving to countries with low earnings inequality.

In addition to earnings distribution, migration costs may explain the direction of self-selection. Chiquiar & Hanson (2005) extend the Borjas (1987) model by considering that in practice migration costs vary by skill level. Bureaucratic, transportation, job-search and information costs involved in migration represent fewer hours of work for the more skilled, who can finance migration with no or lower borrowing costs. Hence, migrants should be positively self-selected if migration costs are large enough and credit constraints sufficiently binding. This approach has motivated the assessment of some factors that can shape migrant self-selection by affecting migration costs—for example, migrant networks (McKenzie & Rapoport, 2007, 2010) and wealth (Abramitzky et al., 2013; Connor, 2019). Also, in the past and present, guest worker programs, immigrant quotas and skill-based admission systems have been implemented to artificially increase or reduce the supply of workforce (Clemens et al., 2018; Massey & Pren, 2012; Timmer & Williamson, 1998). These immigration policy interventions modify migration costs and/or entry requirements, and therefore the skill-composition of migrants and degree (direction) of selection into migration (return migration) (see, for example, Antecol et al., 2003; Bianchi, 2013; Greenwood & Ward, 2015; Massey, 2016; Mayda et al., 2018; Spitzer & Zimran, 2018; Ward, 2017).

A common characteristic among the previous drivers is that they are more or less fixed in the short run (Chiquiar & Hanson, 2005, p. 243). Convergence/Divergence in absolute earnings between origin and destination countries, and changes in earnings inequality are long-run processes. Also, immigration reforms can take years or even decades to materialize due the political clout of immigrants (Goldin, 1994). As can be noticed, we know little about the responsiveness of migrant self-selection to changes in the economic environment.

4.3.1 Short-run shifts in migrant self-selection

Contrary to nowadays, by 1910 the United States was slightly more unequal than Mexico. The Gini index of income was 0.54 in the United States (Lindert & Williamson, 2016, p. 174) and 0.51 in Mexico (Moatsos et al., 2014, p. 206). Therefore, the basic Roy-Borjas model—in which migration costs are assumed to be the same for everyone—would predict intermediate or positive selection of Mexican migrants.

To assess if migrant self-selection can change in the short run, we consider a large but temporary shock that unexpectedly affected the demand of Mexican workers in the early twentieth century: the Panic of 1907. This crisis was a nine-month breakdown of the US banking system that limited payments and access to credit in all sectors of the American economy. As a consequence, thousands of companies went bankrupt or curtailed operations, which in turn sharply reduced the demand of workforce and compressed the earnings distribution in the United States—that is, the Panic could have induced a temporary reduction in earnings inequality. In this sense, the basic Roy-Borjas model would predict fewer individuals from the upper half of Mexico's skill distribution choosing to migrate: the Mexico-US migration flow should become less positively selected during the crisis.

In the presence of random shocks, short-run shifts in the direction of selection would also depend on the relationship between the shock and the costs of migration—that is, migrant self-selection could change if the shock impacts factors affecting migration costs. For example, payments and basic transactions were constrained during the Panic of 1907. This could have limited the effect of migrant networks on relaxing credit constraints for low skill-migrants, and therefore the crisis pushed toward a positive self-selection. Conversely, financial crises may neutralize factors influencing positive migrant selection, and thus push toward a negative self-selection. We provide evidence suggesting that the latter happened during the Panic of 1907, and consequently Mexican migrants became less positive selected during the crisis.

In the past and today, institutions involved in the immigration process such as immigrant banks, immigrant aid societies and visa sponsors could be impacted by shocks similar to the *Panic of 1907* (Abramitzky & Boustan, 2017, p. 1325). The degree of change in the composition of migrants will depend on the shock's capacity to modify significantly earnings inequality, migration costs and factors relaxing credit constraints for the population at risk of migration. Finally, this could be true only in the absence of immigration restrictions. Barriers to immigration aim to keep the skill-mix of migrants and/or the size of migration constant in the short run, which would prevent us to observe short-run shifts in migrant self-selection.

4.4 Selection of Mexican migrants in the past and present

Knowing the selectivity of Mexican migration—whether Mexicans tend to come from the bottom or top of the Mexican skill distribution—has important implications in the short and long run. A persistent migration stream composed of less-skilled Mexicans will tend to reduce the relative scarcity of high-skilled labor in Mexico over time and reduce earnings disparities between high and low-skilled

workers. The composition of migrants could also influence remittances sent from abroad, affecting the economic well-being of families in Mexico (Ibarraran & Lubotsky, 2007, p. 160).

There is no consensus about the selection direction of Mexican migration in contemporary times. Chiquiar & Hanson (2005) argue that the negative selection predicted by the Roy-Borjas model does not hold when comparing counterfactual wage densities for migrants and residents of Mexico. They find that, under a common price for observable skills, Mexican migration was characterized by an intermediate or positive selection from 1990 to 2000. Similarly, Orrenius & Zavodny (2005) argue that during the 1980s and 1990s undocumented Mexican migrants were drawn from the middle of the educational distribution, and stricter border enforcement was associated with higher average skill levels. McKenzie & Rapoport (2010) observe the same selection pattern in 1997, but only in communities with weak migrant networks: stronger networks influence negative migrant self-selection.

In contrast, using the 2000 Mexican Census, Ibarraran & Lubotsky (2007, p. 190) argue that Mexican migrants tend to be less educated than the average resident in Mexico. In addition, they find that the degree of negative selection is magnified in Mexican municipalities that have relatively higher returns to education. Exploiting Mexican panel data, Ambrosini & Peri (2012), Kaestner & Malamud (2014), and Fernandez-Huertas (2011) find that migrants earned lower wages than their non-migrant peers—that is, they support the negative-selection hypothesis of Borjas. As argued by Abramitzky & Boustan (2017), differences in the observed self-selection patterns may be rooted in the measure of selection used and/or under-enumeration of undocumented Mexican migrants.

To some extent, undocumented migration bias could be overcome by studying Mexican migration in the past. Before 1921, Mexican migrants did not have a clear incentive to avoid official entrance ports since they could cross the border

without restrictions (Cardoso, 1980, p. 98). However, the historical character of Mexican migration has not been exploited to study selection into migration. The exception is Kosack & Ward (2014), who use height to proxy migrant quality and estimate the self-selection of Mexican migrants into and out of the United States in the 1920s. Their empirical strategy estimates differences in height between migrants and Mexican soldiers controlling for diverse factors that may influence an individual's height. Their findings suggest that Mexican migrants were positively self-selected. In addition, they link their migrant sample to 1930 US and Mexican census to obtain samples of permanent and return migrants. They argue that return migrants were neutrally selected relative to permanent migrants.

Kosack and Ward's research is our closest reference in methodology and time period. However, their results regard to the *Restrictions and Deportations Era (1921–41)*, and their findings may capture effects of the Mexican Revolution (1910–20). We also use height as measure of selection, but our findings regard to the *Beginnings (1884–1910)* of the migration stream.[15]

4.4.1 Height as a measure of selection

Average height reflects genetic factors as well as nutritional and health conditions during early childhood and youth. Since wealthier people have better access to food, hygienic conditions and medical resources, they tend to be taller than the poorer population (see, for example, Borrescio-Higa et al., 2019; Deaton, 2007; Komlos & Baten, 2004; Komlos & Meermann, 2007; Steckel, 1995). Hence, human stature is indicative of income, wealth and life chances. Taller individuals also develop better cognitive abilities, reach higher levels of education and thus higher incomes as adults (Case & Paxson, 2008; Ogórek, 2019; Schultz, 2002). Moreover, for some occupations there are returns to strength, which is correlated with height as well (Juif & Quiroga, 2019, p. 116).

[15]See Durand (2016, p. 7) for a periodization of Mexican migration.

Average heights are useful to study selection into migration, especially in contexts where income measures are scattered and large sectors of the economy rely on physical productivity of labor. This is the case of Mexico in the early twentieth century (López-Alonso, 2007). An additional advantage of using height as measure of migrant selection is that—for adult migrants—it cannot be manipulated in anticipation of or in response to migration (Spitzer & Zimran, 2018, p. 229).

Few studies have used heights to analyze self-selection. Spitzer & Zimran (2018) find that Italian migration was negatively selected at the national level from 1907 to 1925. They also show that national-level estimates could mask substantial variation at the local level. In fact, Italian migrants were positively selected at the local level and selection varied systematically within Italy, with more positive local selection from poorer provinces. Blum & Rei (2017) analyze the health human capital of Jewish migrants from 1940 to 1942. They suggest that—on the basis of stature—both refugees and nonrefugees were positively selected. In addition, Humphries & Leunig (2009) and Juif & Quiroga (2019) address the self-selection of internal migrants in England (1844–48) and Spain (1893–1954), respectively. Both studies find that migrants were taller than those who chose to remain.

4.5 Data

4.5.1 Immigrant sample

The registration of aliens arriving at Mexican-US land border ports began in 1906. American authorities used different types of documents to collect information about immigrants. These documents are known as *Mexican Border Crossing Records*, and at the time were conducted by the Bureau of Immigration and Naturalization. The immigrant sample used in this research comes from the

publication N° A3365,[16] which contains two-sheet manifests that provide rich and diverse information about immigrants. Characteristics such as age, sex, marital status, occupation, ability to read and write, citizenship and race were reported. The manifests include anthropometric data of the immigrant (height, complexion and color of eyes and hair), and geographical information regarding his/her birthplace, final destination and last residence. In addition, these documents recorded information regarding the immigrant's current and previous migration spells.

We reviewed the five rolls of the publication to identify the total data (population size) in the microfilms.[17] Data from 1909 was not considered to guarantee capturing only labor migrants and not refugees from the Mexican Revolution (1910–20). Therefore, we limited the transcription process to the period from July 1906 to December 1908. Using as reference the port of entrance, the data was transcribed following a stratified random sampling strategy. The final sample size is 9,083 individuals. Chapter 2 provides a full description of the MBCRs and the sampling plan I followed. Figure 4.5 in Section 4.10 shows the nine ports of entrance along the Mexico-US border that were identified in the microfilms.

A concern about these manifests is that they record only immigrant crossings in official entrance ports. Estimations of undocumented Mexican migration are scarce for the period, mainly because Mexicans were not consider immigrants at the time (Durand, 2016).[18] Yet, these data can be considered unique since it is the only migration data at individual level for the Porfirian period (1876-1911).[19]

To estimate the selection of Mexican migrants, this study uses as core data the immigrant's age, height, birthplace, and occupation. The anthropometric data

[16]The title of the publication is: Lists of Aliens Arriving at Brownsville, Del Rio, Eagle Pass, El Paso, Laredo, Presidio, Rio Grande City, and Roma, Texas, May 1903-June 1909, and at Aros Ranch, Douglas, Lochiel, Naco, and Nogales, Arizona, July 1906–December 1910.
[17]We did not find data for years prior 1906 and for entrance ports in California.
[18]Chapter 2 provides evidence suggesting that the sample is representative for the period under analysis.
[19]From 1877 to 1911, Mexico was ruled by General Porfirio Díaz (Cosío Villegas & Bernal, 1973). This dictatorship is known as the *Porfiriato*.

was recorded by a sworn physician and surgeon, who examined each immigrant at the port of entrance. However, the immigrant's age, birthplace and occupation were self-reported, and consequently subject to bias.

4.5.2 Comparison samples: military and passport holders

We use military recruitment files and passport records to compare migrants with the population that chose to remain in Mexico. These data are the result of extensive archival work completed by López-Alonso (2015), who uses height to study secular trends of living standards in Mexico from 1850 to 1950.[20] We believe that these comparison samples capture different parts of the height distribution of the Mexican population, allowing us to identify to which part migrants might have belonged.

The military recruitment files consist of two samples that capture two extreme points of the height distribution of the Mexican working class. On the one hand, the federales were average soldiers of the Mexican army (cavalry, infantry, and artillery), who served and retired, lost their lives in the line of duty, or left their service without authorization before the ending of their contracts (deserters). At the time, there were minimum requirements to enlist in the army. Recruits had to be between 18 and 45 years of age, be at least 160 centimeters tall, be able to understand Spanish language, be a Mexican citizen, and other health requirements. While these requirements might have introduced systematic biases to the sample, López-Alonso (2015, p. 112) shows that neither of these requirements were enforced during the period.[21] The sample size is 7,088 males born between 1840 and 1950 that proxy for the average laborer/peasant in Mexico, i.e. the lower ranks of the Mexican working class. The source of this data are

[20] López-Alonso (2015, p. 107) provides a detailed description of the archival worked involved.
[21] Moreover, it is not clear if the enrollment into the army was completely voluntary. Although the military did not required service until 1939 (Kosack & Ward, 2014), there is evidence that forced recruitment mechanisms might have been implemented at the time (Durand, 2016)

the archives of the Ministry of National Defense (*Secretaría de la Defensa Nacional*–SEDENA).

On the other hand, the rural police, known as the rurales, was a militia created in 1860 as an armed group loyal to the president. The members of this militia received a higher salary than the federales, and in its beginnings they had to bring their own horse and weapons. The rurales often received additional monetary rewards and political favors to maintain the stability in the country. The sample size is 6,820 individuals born between 1840 and 1900.[22] This sample covers all the enlistment records of this militia, and the source of this data is the General National Archive - Public Administration Section (*Archivo General de la Nación*–AGN).

We considered the previous samples separately because clearly the rurales were not representative of the average Mexican soldier. The fact that they received a higher salary and had to bring their own equipment suggests that they might have been relatively richer than the average soldier. Moreover, they received extra monetary and non-monetary rewards for their service. Hence, the rurales could be considered as the military elite of that time, and they proxy for the upper ranks of the Mexican working class (López-Alonso, 2015, p. 156).

Finally, the passport records consist of all the passport applications made from 1910 to 1942 that contain the height of the applicant. We believe that this sample represents the Mexican upper social class since the passport holders might be individuals with the economic means to travel abroad for business, leisure or education purposes (López-Alonso & Condey, 2003). Nevertheless, two important characteristics of this data should be noticed. Firstly, the height was self-reported by the applicant. Secondly, the records capture all the issued passports but not all the travel permits issued by other regional offices for applicants that could not travel to Mexico City. The sample size is 6,746 male individuals born between

[22]The desertion rates in this militia were high since its members could sell their equipment at any time and locating deserters was costly López-Alonso (2015, p. 117-121).

1860 and 1922. The source of this data are the archives of the Ministry of Foreign Affairs (*Secretaría de Relaciones Exteriores*–SRE) (López-Alonso, 2015, p. 121-22).

Besides the potential sources of bias described above, we acknowledge that the three comparison samples could be selected for different reasons.[23] For example, since the military samples record volunteers rather than conscripts, they contain only individuals who choose to enlist in the military. Following Bodenhorn et al. (2017, p. 201), the decision to join the army reflects the individual's evaluation of his future based on his accumulated human health capital (height). Thus, the decision to enlist in the military becomes less attractive for taller individuals, especially in a growing economy.

To test for the presence of selection, we performed the Bodenhorn et al. (2017) test for the three comparison samples.[24] We found that our samples might present height-based selection considering the individual's year of birth and year of registration. This might imply that our samples are not representative for the Mexican population. However, the objective of this research is not to estimate secular trends in height of the Mexican population, but to use the comparison samples to identify whether the individuals that emigrated were taller or shorter relative to the average soldier (federales), the average elite military (rurales), and the average passport holder.

4.5.3 Data refinements and descriptives

To obtain the best results possible, we impose a series of refinements to the data. We keep only males reporting their town and state of birth in Mexico.[25] This allows us to estimate accurately the migrant selection at the regional level. In

[23] Despite its limitations, these data arise as the most suitable information to measure living standards during the Porfiriato. For the Mexican case, long term data series on income, wages, prices and mortality is unreliable and scattered (López-Alonso, 2007).

[24] The test consists of including all possible interactions of birth-year and recruitment-year effects in a regression in which height is the dependent variable.

[25] The sample was constrained to males because the female military sample (comparison group) does not come with any geographic information.

addition, we keep individuals that have reached their terminal height at the moment of registration: individuals between 22 and 65 years old. This avoids capturing growing and shrinkage effects (Spitzer & Zimran, 2018, p. 231).

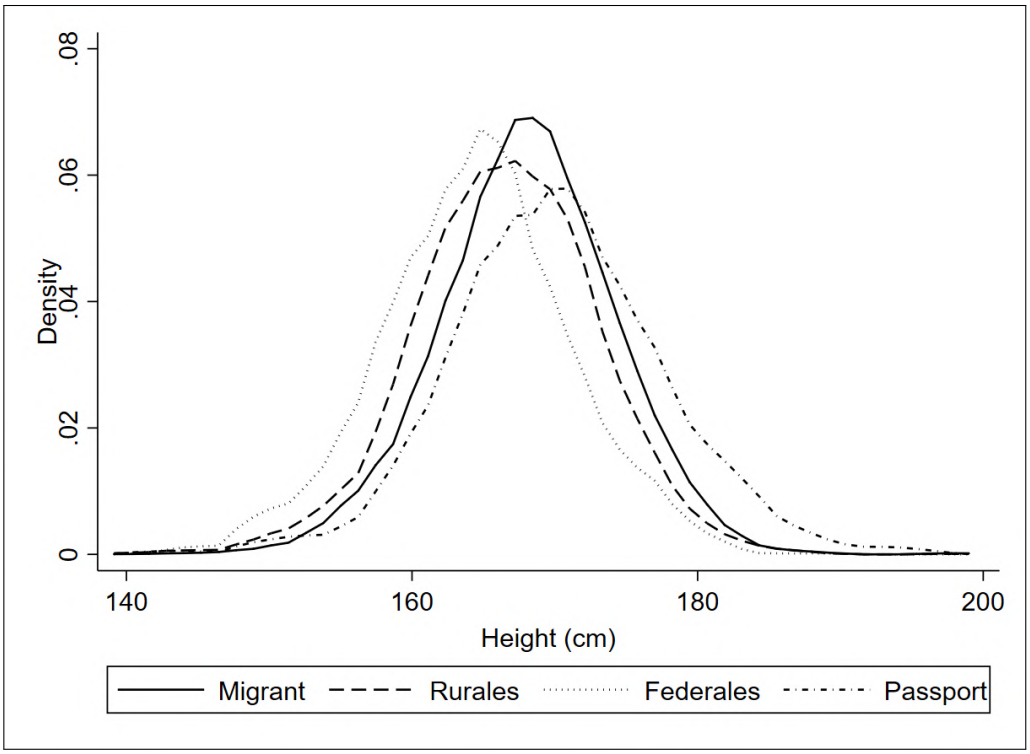
To avoid capturing effects of the Mexican Revolution in the comparison samples, we keep military and passport holders that had passed their pubertal growth spurt before the Mexican Revolution regardless of their year of registration: individuals 18 years old or older before 1911. In other words, we keep individuals that had reached their peak growth velocity before the conflict. We decided to apply this partial refinement because keeping only those individuals registered before the revolution (the ideal comparison sample) reduces significantly the sample sizes. Therefore, some effects of the revolution might be captured.

Table 4.1: Summary statistics. Migrant, military and passports samples (males)

	Migrant	**Federales**	**Rurales**	**Passport**
Average Height (cm)	168.0	164.4	166.6	170.1
Average Age (years)	31.2	35.3	29.7	48.3
Labor Class (%)				
Unskilled	89.1	73.3	47.8	3.7
Skilled	7.7	24.1	49.3	34.2
Professional	2.2	2.6	3.0	61.3
Literacy Rate (%)				
Literate	38.4	45.3	49.5	100.0
Marital Status (%)				
Married	58.9	na	na	na
Single	38.8	na	na	na
Widowed	1.8	na	na	na
Region of Birth (%)				
North	45.5	18.7	2.9	13.4
Bajio	52.5	27.3	60.6	30.0
Center	1.8	42.8	33.0	47.3
South	0.3	11.3	3.5	9.3
Cash on hand–US dollars (median)				
North	10.0	na	na	na
Bajio	1.0	na	na	na
Center	20.0	na	na	na
South	10.0	na	na	na
Observations	3,609	1,249	5,300	1,339

Source: Migrant sample from the Mexican Border Crossing Records - Microfilm publication N° A3365. Military and Passport samples from López-Alonso (2015).
Note: The migrants' regions of birth and occupations were classified following López-Alonso (2015, p. 127 & 128). The sample was constrained to males because the female military sample (comparison group) does not count with any geographic information. We consider individuals that had reached their terminal height at the moment of registration: individuals between 22 and 65 years old.

Figure 4.1: Kernel density estimates of heights



Source: Migrant sample from the Mexican Border Crossing Records - Microfilm publication No. A3365. Military and Passport samples from López-Alonso (2015).

As a result of the refinement process, the samples' sizes shrink relative to the raw data. Table 4.1 presents the main characteristics of the final samples. On average, migrants were 168 centimeters tall, 3.6 centimeters taller than the average soldier, 1.4 centimeters taller than the military elite, and 2.1 centimeters shorter than the passport holders. The kernel density estimates of the samples confirm that the federales were the shortest individuals (see Figure 4.1). This initial finding suggests that migrants did not belong to the lower tail of the height distribution of the Mexican working class. Moreover, following Schneider & Ogasawara (2018, p. 64), a similar average height might indicate that two groups (rurales and migrants) faced equivalent conditions of health care, nutrition, disease environment and work assignments some 10 to 50 years before being observed.[26] To some extent, this suggests that those individuals that decided to emigrate and those that enlisted in the military elite could have belonged to the

[26] Schneider & Ogasawara (2018) argue that disease environment, proxied by infant mortality rates, have economically meaningful effects on child height at ages 6-11.

same social class. Conversely, the average height of the federales is lower than the migrants' average height. Therefore, the average Mexican soldier was exposed to worse early nutrition/health conditions than the rest of the individuals, meaning that they might have belonged to the lowest social class in Mexico.

However, almost all migrants were unskilled laborers and had the lowest literacy rates among the samples, implying that they might have emigrated to perform activities with high returns to physical productivity. In fact, Clark (1908, p. 477 & 486) documents that most Mexican immigrants were confined to track maintenance in the railways, and that they were employed as drillers, wood choppers, coke pullers, and surface men (strip mining). All of these occupations required physical strength.[27] In contrast, 62 percent of the passport holders reported to be professionals, confirming that they belonged to the upper social class.

The regional distribution of the samples shows that migration occurred mostly from the North and Bajio regions, while military recruitment took place mainly in the Bajio and Center regions (see Table 4.1).[28] The passports sample concentrates in the Center region, reaffirming that most passport holders might have lived in Mexico City or nearby states, where the social elite resided. Table 4.2 shows that the differences in height between migrants and federales widens in the Center and South regions. Based on the amount of cash held at the crossing (see Table 4.1), migrants from the Center region were considerably richer than the rest. They reported to have 20 dollars, two times the amount reported by the migrants from the North. Bajio migrants had one dollar at the crossing, thus they might have been the poorest as argued by Durand (2016). These initial descriptives

[27] Certainly, Mexicans were employed as cotton pickers during the harvest season. This activity required nimble fingers rather than physical strength. However, complete Mexican families were employed in the cotton fields since children often picked as much as adults (Clark, 1908, p. 482).

[28] Considering the population levels in 1907, the Bajio states were among the most populated. The states of Guanajuato, Jalisco and Michoacan were more populated than Mexico City. Therefore, the recruitment of soldiers would have been common in this regions.

suggest the existence of substantial variation in selection into migration across regions (Abramitzky & Boustan, 2017, p. 20).

Table 4.2: Average heights (centimeters) across regions (males)

	North	Bajio	Center	South
Migrant	169.2	167.0	167.9	165.4
	(6.0)	(5.9)	(7.2)	(5.4)
Rurales	167.4	166.8	166.0	166.3
	(6.39	(6.3)	(6.4)	(5.7)
Federales	166.8	165.2	163.7	161.3
	(6.9)	(6.6)	(5.9)	(5.7)
Passports	171.3	171.1	169.4	168.9
	(7.3)	(7.5)	(7.3)	(7.1)
Observations	2,208	5,850	2,978	461

Source: Migrant sample from the Mexican Border Crossing Records - Microfilm publication N° A3365. Military and Passport samples from López-Alonso (2015).
Note: Standard deviations in parenthesis. The migrants' regions of birth were classified following López-Alonso (2015, p. 127). The sample was constrained to males because the female military sample (comparison group) does not count with any geographic information. We consider individuals that had reached their terminal height: individuals between 22 and 65 years old.

4.6 Empirical strategy

To estimate the selection of Mexican migration to the United States, we regress the height of individual i ($height_i$) on a dummy variable that takes the value of 1 if the individual belongs to the migrant sample and zero otherwise ($migrant_i$). We also consider a vector of individual characteristics ($\mathbf{X}_i$) that includes region of birth and skill level categories. Additionally, we control for year of birth fixed effects (α_c):

$$height_i = \beta + \Phi migrant_i + \mathbf{X}_i'\theta + \alpha_c + e_i. \tag{4.1}$$

We estimate Equation 4.1 by pooling the migrant sample with each of the comparison samples separately, hence the estimated coefficient Φ reflects the average difference in height between migrants and federales, rurales or passport holders, respectively. The region of birth categories (North, Bajio, Center and South) control for environmental factors such as food availability, dietary patterns or presence of diseases that might influence height at the regional level. Also, the

region of birth categories factor out composition effects of the sample. The regional classification was taken from López-Alonso (2015, p. 127). We include skill level categories (unskilled, skilled and professional) to control for the potential relationship between cognitive abilities acquired in early childhood and stature (Bleakley et al., 2014, p. 124).[29] Finally, the year of birth fixed effects control for any factor influencing height across years, such as structural and idiosyncratic shocks effecting the living standards of the population over time.

The estimated coefficients of Equation 4.1 are average estimates of the period October 1906–December 1908. However, as mentioned above, from August 1907 to January 1908 the US economy suffered one of the most severe financial crises before the Great Depression (Moen & Tallman, 1992, p. 611).[30] To capture shifts in selection into migration as a consequence of this crisis, we extend Equation 4.1 by interacting the migrant indicator variable with dummy variables for the Panic ($panic$) and post-Panic period ($panic^{post}$):

$$\begin{aligned} height_i = &\beta + \Phi_1 migrant_i + \Phi_2 migrant_i \times panic + \Phi_3 migrant_i \times panic^{post} \\ &+ \mathbf{X}_i'\theta + \alpha_c + e_i. \end{aligned} \tag{4.2}$$

The estimated coefficients Φ_2 and Φ_3 capture the difference in height of individuals that emigrated during the Panic period (August 1907–January 1908) or after the Panic (February 1908–December 1908), respectively. These estimates are relative to those individuals that emigrated before the Panic (October 1906–July 1907). The difference in height between pre-Panic migrants and the different comparison samples (non-migrants) is reflected in Φ_1. Holding everything else equal, the estimated selection pattern during the Panic of 1907 is $\Phi_1 + \Phi_2$.

[29] We assume that skilled and professional occupations at the time demanded more training or education relative to unskilled occupations.

[30] There is no consensus in the literature about the ending month of the crisis. Yet, the scholarship on the matter agree that normalcy in the financial market was restored in January 1908 (Frydman et al., 2015, p. 937).

4.7 Results

4.7.1 Self-selection of mexican migrants before 1910

Column 1 of Table 4.3 shows that migrants were positively selected relative to the average soldier: migrants were 2.2 centimeters taller than the federales. The difference in height between migrants and rurales was 0.5 centimeters, implying that migrants were slightly taller than the military elite forces (column 2). Regarding the passport holders, migrants were 2.1 centimeters shorter. In other words, migrants were negatively selected relative to the Mexican upper class (column 3).

Table 4.3: Self-selection of Mexican migrants. Dependent variable: height (centimeters)

	1	2	3	4	5
	Federales	Rurales	Passports	Rurales	Passports
Migrant	2.209***	0.557***	-2.143***	0.514***	-2.103***
	(0.350)	(0.187)	(0.508)	(0.194)	(0.513)
Skill level categories					
Skilled	0.928***	0.077	0.634*	0.136	0.766**
	(0.267)	(0.160)	(0.333)	(0.160)	(0.333)
Professional	0.481	1.091***	1.540***	1.172***	1.573***
	(0.552)	(0.403)	(0.440)	(0.402)	(0.442)
Region of birth					
North	5.500***	2.506***	3.365***		
	(0.528)	(0.453)	(0.660)		
Bajio	3.349***	0.515	1.372**		
	(0.521)	(0.419)	(0.651)		
Center	2.407***	-0.269	0.526		
	(0.520)	(0.433)	(0.658)		
Observations	4,822	8,860	4,901	8,860	4,901
R-squared	0.117	0.053	0.059	0.063	0.077
Birth year FE	Yes	Yes	Yes	Yes	Yes
Birth state FE	No	No	No	Yes	Yes

Source: Mexican Border Crossing Records - Microfilm publication N° A3365 and López-Alonso (2015).
Notes: * = Significant at 10% level; ** = Significant at 5% level; *** = Significant at 1% level. Robust standard errors in parenthesis. The omitted categories are individuals born in the South region and unskilled workers.

We acknowledge that the federales sample is selected because these individuals were not conscripts but volunteers; and it is expected that in a growing economy, as was the Profirian Mexico, the outside option of military service becomes less

attractive for productive and tall individuals (Bodenhorn et al., 2017, p. 173). Therefore, the federales might have belonged to the lower ranks of the population height distribution. The rurales were volunteers as well; however the estimated Φs provide strong evidence that they were well above the federales in the population height distribution. Since migrants were taller than the rurales, we can argue that migrants belonged to the upper tail of the height distribution of the Mexican working class.

As argued by Bleakley et al. (2014, p. 124) and Kosack & Ward (2014, p. 1023), height is strongly correlated with wages in countries where large sectors of the economy rely on the physical productivity of labor.[31] Even though the military samples are not representative of the Mexican population by themselves, jointly, they allow us to infer that the migrants' potential wage was very close to that of the military elite. More importantly, those individuals that decided to emigrate had unobserved individual-specific factors that reveal even higher human capital accumulation (Bodenhorn et al., 2017, p. 201). Therefore, it is unlikely that, the first Mexican migrants were negatively selected relative to the average laborer.

As a robustness check, we include state of birth fixed effects instead of region dummies in the models for which more disaggregated geographic data is available (rurales and passports). This helps us to rule out that the results are driven by unobserved factors across states of birth. Columns 4–5 of Table 4.3 show that our initial results hold in significance and magnitude.

Finally, the region categories show that individuals from the North and Bajio were considerably taller than the rest, confirming that regional environmental factors influenced height in Mexico. Did the magnitude of selection vary across regions? To answer this question, we estimate separately Equation 4.1 for each region. We only present results for the North and Bajio because these regions

[31] Schultz (2002) estimate that a one centimeter increase in height leads to an 8 to 10 percent increase in wages in Brazil and Ghana. Mexico was basically an agrarian country in the beginning of the twentieth century (Rosenzweig, 1965), thus physical strength was the principal source of human capital for the average laborer.

concentrate 98 percent of the migrant sample. Columns 4–5 of Table 4.4 show that there was considerable variation in the degree of regional selection across Mexico. The positive selection relative to the average soldiers (Panel A) was stronger in the Bajio than in the North. By 1910, salaries and living standards in the Bajio were considerably lower than in the North (Rosenzweig, 1965, p. 450; Campos-Vázquez & Vélez-Grajales, 2012, p. 613). Therefore, migrants from poorer regions were disproportionately drawn from the upper tail of the working class height distribution. This pattern does not hold relative to the rurales because the enlistment requirements to the military elite implied a more selective screening in the Bajio.

Table 4.4: Regional self-selection of Mexican migrants. Dependent variable: height (centimeters)

	1	2	3	4	5
	Complete Sample			North	Bajio
Panel A. Federales					
Migrant	3.259***	3.386***	2.209***	1.273**	2.490***
	(0.306)	(0.308)	(0.350)	(0.630)	(0.609)
Observations	4,858	4,822	4,822	1,848	2,227
R-squared	0.077	0.080	0.117	0.061	0.041
Panel B. Rurales					
Migrant	1.604***	1.633***	0.557***	1.114*	0.437**
	(0.152)	(0.163)	(0.187)	(0.608)	(0.214)
Observations	8,896	8,860	8,860	1,769	5,087
R-squared	0.038	0.039	0.053	0.049	0.033
Panel C. Passports					
Migrant	-1.993***	-0.815*	-2.143***	-2.282*	-2.849***
	(0.327)	(0.461)	(0.508)	(1.178)	(0.880)
Observations	4,948	4,901	4,901	1,793	2,286
R-squared	0.033	0.036	0.059	0.047	0.080
Skill level categories	No	Yes	Yes	Yes	Yes
Region of birth categories	No	No	Yes	No	No
Birth year FE	Yes	Yes	Yes	Yes	Yes

Source: Mexican Border Crossing Records - Microfilm publication N° A3365 and López-Alonso (2015).
Notes: * = Significant at 10% level; ** = Significant at 5% level; *** = Significant at 1% level. Robust standard errors in parenthesis. For columns 1–3, the omitted categories are individuals born in the South region and unskilled workers.

4.7.2 The effect of the Panic of 1907

Columns 1–3 of Table 4.5 shows the effect of the Panic of 1907 on migrant self-selection. Individuals that migrated during the crisis were approximately 0.9 centimeters shorter than their pre-Panic counterparts—that is, migrants became less positively selected during this period. However, the estimated selection during the post-Panic period is close to zero and not statistically significant, meaning that those who migrated after the crisis had a stature similar to pre-Panic migrants.

Table 4.5: Impact of the Panic of 1907 on migrant self-selection. Dependent variable: height (centimeters)

	1	2	3	4	5
	Federales	Rurales	Passports	Rurales	Passports
Migrant	2.400***	0.731***	-1.953***	0.412*	-2.204***
	(0.364)	(0.204)	(0.518)	(0.213)	(0.524)
Migrant × Panic	-0.976***	-0.994***	-0.958***	-0.644**	-0.675**
	(0.288)	(0.289)	(0.288)	(0.291)	(0.290)
Migrant × Post Panic	-0.111	-0.060	-0.092	0.870***	0.622**
	(0.251)	(0.246)	(0.253)	(0.279)	(0.291)
Observations	4,822	8,860	4,901	8,860	4,901
R-squared	0.119	0.054	0.061	0.065	0.079
Skill level categories	Yes	Yes	Yes	Yes	Yes
Region of birth categories	Yes	Yes	Yes	No	No
Birth year FE	Yes	Yes	Yes	Yes	Yes
Birth state FE	No	No	No	Yes	Yes

Source: Mexican Border Crossing Records - Microfilm publication N° A3365 and López-Alonso (2015).
Notes: * = Significant at 10% level; ** = Significant at 5% level; *** = Significant at 1% level. Robust standard errors in parenthesis. The omitted categories are individuals born in the South region and unskilled workers.

Column 1 of Table 4.5 reveals that before the Panic, migrants were positively selected relative to the average soldier (2.4 centimeters taller). This pattern changes during the Panic, when migrants were less positively selected (1.4 centimeters), but it returns to pre-crisis levels afterwards. Columns 2–3 show the same "U" pattern relative to the rurales and passports samples. Therefore, the findings suggest that in the beginnings of the twentieth century, when migrants

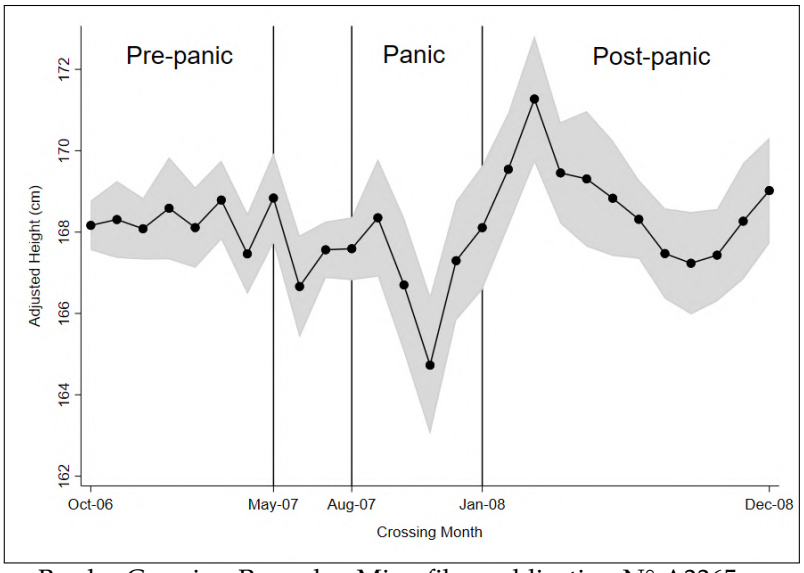
were able to cross the border without restrictions, the composition of the Mexican migration adjusted to short-run changes in the demand of migrant workers.

However, when controlling for unobserved factors across states, the selection into migration changes in the post-Panic period. Columns 4–5 of Table 4.5 show that migrants became more positively selected than their pre-Panic peers. This can be appreciated more clearly in Figure 4.2 that depicts the adjusted height of migrants during the complete period under analysis (October 1906–December 1908). To estimate the adjusted values in each month, we regress the migrants' height on skill level, state of birth, year of birth, month of crossing, and port of entrance fixed effects. There are two things to note.

First, in March 1907 the first strong drop in stock prices occurred. In the following months, the speculation and uncertainty continued and by May 1907 the US had fallen into a short but severe recession (Odell & Weidenmier, 2004, p. 1003). This might explain the fall in the adjusted height from May to August 1907. However, in August 1907 the Secretary of the Treasury announced the deposit of 28 million dollars to banks across the US for relieving the expected stringency in money supply and bring back confidence to the financial system (Markham, 2002, p. 31). This measure only delayed the financial crash of October, but along with substitutes for legal currency and the creation of "legal holidays", prevented even more bankruptcies during the Panic period (Andrew, 1908, p. 516). These events might explain why the adjusted height slightly increased after August and fell later on.

Second, the adjusted height increased significantly after January 1908, when the payments to depositors of commercial banks were fully restored. After May 1908, the adjusted height returns to pre-crisis levels. The results in columns 4–5 suggest that additional factors at the state level might have influenced the selection into migration across the periods under analysis. The next section addresses this matter by identifying a potential adjustment mechanism.

Figure 4.2: Effect of the Panic of 1907. Adjusted heights of migrants



Source: Mexican Border Crossing Records - Microfilm publication N° A3365.
Note: The estimates correspond to individuals over age 22, born in the South region and unskilled workers (state of birth, year of birth, month of crossing and port of entrance fixed effects were included). The predicted values were estimated for each individual based on year-month fixed effects. **May-07:** By May 1907, the US had fallen into a short but severe recession. **Aug-07:** In August 1907, the Secretary of the Treasury announced the deposit of 28 million dollars to banks across the US for relieving the expected stringency in money supply and bring back confidence to the financial system. **Jan-08:** In January 1908, the payments to depositors of commercial banks were fully restored.

4.8 Short-run adjustment mechanisms of migrant self-selection

We have presented evidence showing that Mexican migration was characterized by an intermediate or positive selection at the beginning of the twentieth century, and that the Panic of 1907 sparked short-run changes in the selection into migration. This section addresses one mechanism through which migrant self-selection might have adjusted during the Panic.

4.8.1 Systems of labor recruiting: the *enganche*

During the nineteenth century, Mexico was characterized by regional mismatches in the demand and supply of labor. To regulate labor markets, the *enganche* was institutionalized as a mechanism to recruit and transport workers to remote locations or with labor shortage (Durand, 2016, p. 50–1). Recruiters "hooked" workers by offering wages in advance in exchange of future labor-service, creating

a relationship of indebtedness that kept workers at the destination until the debt was cleared (Brass, 1990, p. 74). This labor-recruiting system was mainly practiced in regions characterized by population pressures and low salaries (Rosenzweig, 1965, p. 448).

At the beginning of the twentieth century, American companies and labor contractors adopted the *enganche* to satisfy the increasing demand of workers in the American Southwest and other regions. The internationalization of this labor institution was possible due to the expansion of the Mexican railways network and its connection to the US rail lines from 1884. Indeed, recruiters used railways for traveling south into Mexico and transporting workers north to the United States (Woodruff & Zenteno, 2007, p. 512). However, the recruitment of workers was not confined to places with railway access. Clark (1908, p. 475) argues that immigrants also arrived at border towns where they were met by representatives of large labor contracting companies or *enganche* agencies. Once recruited, workers crossed the border and received transportation to the destination and a subsistence allowance, both discounted from their future wage. In sum, it was search-matching labor institution used to transport and allocate seasonal laborers in the United States (Clark, 1908; Durand, 2016; Gamio, 1930).

This system of labor recruiting induced Mexican mass migration by eliminating transportation and job-search costs. However, it was characterized by the breach of contracts (changes in agreed work locations, labor tasks and wages) once the workers arrived to the United States (Durand, 2016, p. 61). The indebtedness entailed to the *enganche* also prevented immigrants from job turnover and reduced their bargaining power over working conditions. Although this labor institution was probably not attractive for everyone willing to migrate to the United States, it could have been the best option to emigrate for the poor or those facing credit or liquidity constraints. Indeed, Durand & Arias (2000) document that *enganche*

system took advantage of the precarious social conditions and limited labor options in some Mexican regions.

The *enganche* and labor recruiting systems in general can alter the skill-composition of migrants through two channels. First, by systematically changing the origins of migrants. Migrants from locations where labor recruiting is practiced could be overrepresented in the migration stream, and thus modify the skill-composition of migrants. Second, by affecting the self-selection at the local level. Systems of labor recruiting can modify the incentives to migrate. The expected discounted net return of migrating could change for individuals with access to recruiting systems. The effect of labor-recruiting institutions toward a positive or negative self-selection depends on the intensity and nature of recruiting. On the one hand, if it is implemented in low scale and workers are randomly recruited, then the overall skill-composition of migrants may not change. On the other hand, if labor recruiting is importantly involved in the immigration process, and individuals willing to migrate are sorted and recruited based on certain characteristics, then the direction and degree of self-selection can be influenced by this type of labor institutions. In subsequent sections, we provide evidence that the share of *enganche* migrants was significant and that recruiters chose the tallest workers. Thus, the *enganche* was a labor institution that pushed toward a positive migrant self-selection.

4.8.2 Identification of *enganche* migrants

The manifests in the publication N° A3365 do not identify directly migrants using the *enganche* to cross the border. Therefore, we design a methodology to identify *enganche* migrants based on the characteristics of this system of labor recruiting.

From the recruiters perspective, the *enganche* profitability depended on the number of workers delivered and transportation efficiency. Therefore, they recruited large numbers of workers and transported them using railways. Previous

literature suggests that recruiters commonly transported between 30 and 400 workers depending on the nature of the jobs and season of the year (Clark, 1908, p. 470 & 476; Durand, 2016, p. 56 & 63). We search for *enganche* advertisements in Mexican and American newspapers of the period to validate this information. We find twenty advertisements covering the period from 1902 to 1909 and diverse destinations in the American Southwest. The number of vacancies advertised range from 50 to 600, which match well the figures suggested by historical literature.

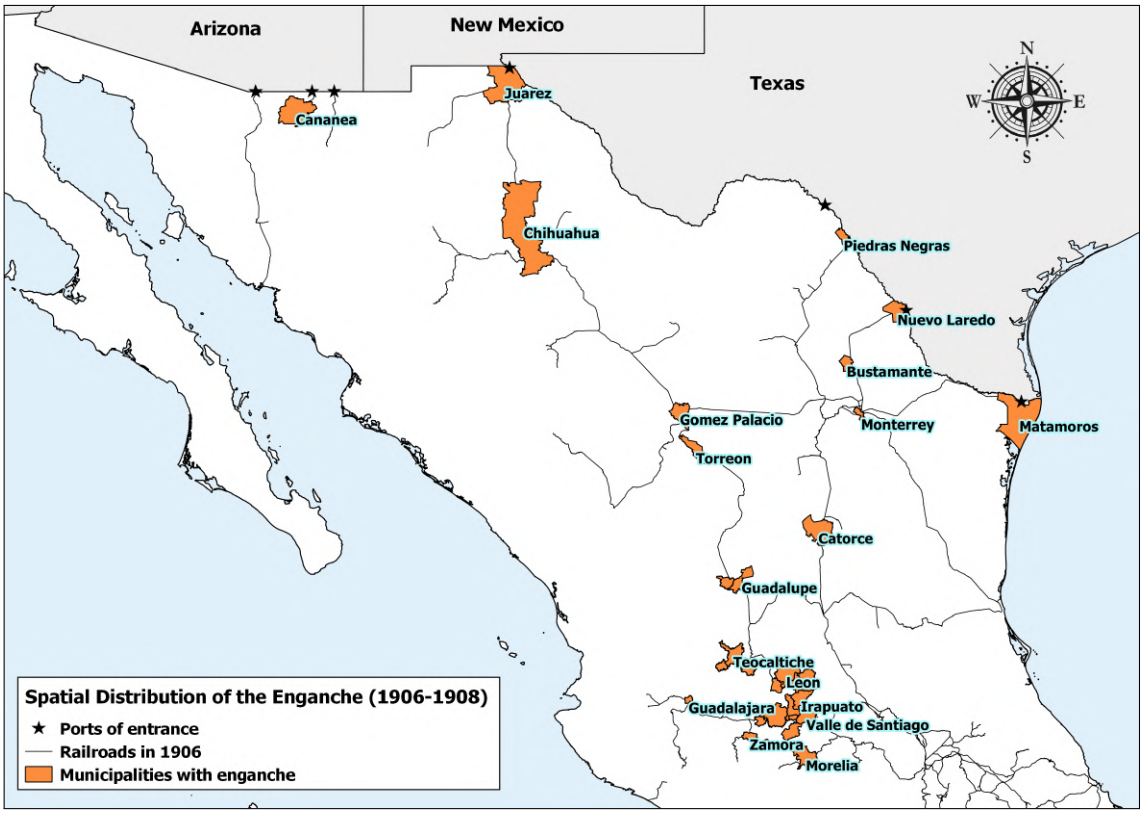
In this sense, recruited workers can be identified in the manifests as groups of migrants who crossed the border the same day and reported the same origin and destination. To identify *enganche* migrants in the sample, we first quantify the number of migrants in each migration flow—that is, number of individuals reporting the same entrance port, year-month of crossing, location of origin (Mexican municipality) and location of destination (American county).[32] Second, we standardize the size of each migration flow using the mean and standard deviation of each migration corridor (municipality-port-county combination). By estimating z-scores for each migration flow, we are able to identify unusual monthly-crossing peaks in each migration corridor. Finally, we consider as *enganche* migrants those individuals belonging to a group (migration flow) of at least 30 migrants registered at the same entrance port, in the specific month, reporting the same origin and destination; and which size was at least one standard deviation above the average size of the flows in the same migration corridor. This criteria allows us to identify groups of crossings that were actually different in size, which proxies for the presence of *enganche*. A formal expression of this methodology can be consulted in Section 4.10.

Figure 4.3 displays the spatial distribution of the origins of *enganche* migrants: locations where this system of labor recruiting was practiced. All the municipalities have direct access to railways, which was a necessary condition for the presence of the *enganche*. The micro data also support the argument that *enganche*

[32]We use the month of crossing because the manifests do not report the day.

agencies were present at border towns. Finally, the presence of the *enganche* in the central plateau of Mexico corroborates that this labor institution was practiced in locations with labor market pressures. At the time, this region was the most populated in the country and offered the lowest salaries.

Figure 4.3: Spatial distribution of the enganche (1906–08)



Source: Mexican Border Crossing Records - Microfilm publication N° A3365.
Note: The polygons display the municipalities with presence of the *enganche*, a search-matching labor institution that reduced migration costs for migrants. Recruiters or *enganchadores* covered the transportation costs of the migrant in exchange of future labor.

4.8.3 The *enganche* effect

To test if the *enganche* influenced selection into migration, we first expand Equation 4.1 as follows:

$$height_i = \beta + \Phi_1 migrant_i + \Phi_2 enganche_i + \mathbf{X}_i'\theta + \alpha_c + e_i. \tag{4.3}$$

Where $enganche_i$ is a dummy variable that takes the value of 1 if the migrant crossed the border through the *enganche* and zero otherwise. The estimated coefficient Φ_2 captures the difference in height between *enganche* and *non-enganche* migrants. Column 2 of Table 4.7 shows that American recruiters chose the tallest laborers among those willing to emigrate: on average, *enganche* migrants were 0.6 centimeters taller than migrants that crossed the US border without using this labor institution. The estimated coefficient Φ_1 is the difference in height between *non-enganche* migrants and each comparison sample. For example, column 2–Panel B of Table 4.7 shows that *non-enganche* migrants were 0.3 centimeters taller than the military elite (neutrally selected), while *enganche* migrants were 1 centimeter taller ($\Phi_1 + \Phi_2$)—that is, positively selected. These results hold when including state-fixed effects (column 6), suggesting that this labor institution had the potential to influence selection into migration at the local level.

However, the previous results are average estimates for the complete period of the sample (October 1906–December 1908). Table 4.6 shows the migrants' characteristics in the pre-Panic (October 1906–July 1907), Panic (August 1907–January 1908) and post-Panic (February 1908–December 1908) periods. We can observe that the *enganche* was almost not practiced during the Panic of 1907. The share of migrants recruited in Mexico went from 36 percent in the pre-Panic period to one percent during the Panic and partially recovers in the post-Panic period.

To practice the *enganche*, American companies and labor contracting agencies needed constant liquidity to pay train tickets, subsistence allowances and wages in advance for tens or hundreds of recruited workers. When the Panic of 1907 hit the US financial system, American companies were not able to cover the costs associated to the *enganche* because banks and financial institutions limited or suspended cash payments—that is, companies did not count with resources to finance the recruitment of Mexican workers. As the crisis developed, thousands

of firms and over one hundred banks failed (Markham, 2002, p. 32). Therefore, the demand for immigrant workers was severely reduced and the *enganche* almost suspended. Only 1.2 percent of the migrants that crossed the border during the Panic used this labor institution. The crisis particularly affected major railway companies such as Union Pacific, who limited their operations during this period (Johnson, 1908, p. 456). Therefore, the transportation of workers within the United States was constrained as well. In sum, during the Panic of 1907 the recruitment of laborers with higher physical productivity stopped, and consequently the *enganche* effect toward a positive selection.

Table 4.6: Composition of Mexican emigration across periods

	Pre-Panic Oct 1906–Jul 1907	**Panic Aug 1907–Jan 1908**	**Post-Panic Feb 1908–Dec 1908**
Panel A. Complete Sample			
Average Height (cm)	168.1	167.3	168.4
Average Age (years)	30.5	31.8	32.3
Labor Class (%)			
Unskilled	91.6	88.3	83.8
Skilled	5.4	7.8	12.8
Professional	2.0	2.8	2.6
Enganche (%)	36.2	1.2	13.2
Observations (%)	58.0	16.0	25.8
Panel B. Bajio			
Average Height (cm)	166.9	166.6	167.6
Average Age (years)	30.5	31.5	31.7
Labor Class (%)			
Unskilled	96.7	94.3	86.9
Skilled	2.2	3.6	10.7
Professional	0.7	1.4	2.1
Enganche (%)	42.7	0.7	10.2
Observations (%)	64.9	14.8	20.1
Panel C. North			
Average Height (cm)	169.8	168.2	168.9
Average Age (years)	30.4	32.1	32.8
Labor Class (%)			
Unskilled	86.2	85.0	82.6
Skilled	9.5	11.1	14.0
Professional	2.5	2.2	2.1
Enganche (%)	27.3	1.8	15.5
Observations (%)	50.0	17.0	32.5

Source: Mexican Border Crossing Records - Microfilm publication N° A3365.
Note: The migrant's regions of birth and occupations were classified following López-Alonso (2015, p. 127 & 128). We consider individuals that had reached their terminal height: individuals between 22 and 65 years old.

Table 4.7: Impact of the enganche on self-selection patterns. Dependent variable: height (centimeters)

	1	2	3	4	5	6	7	8
Panel A. Federales								
Migrant	2.209***	2.065***	2.400***	2.235***				
	(0.350)	(0.354)	(0.364)	(0.375)				
Migrant × Panic			-0.976***	-0.822***				
			(0.288)	(0.300)				
Migrant × Post Panic			-0.111	-0.007				
			(0.251)	(0.258)				
Enganche		0.631***		0.474*				
		(0.236)		(0.249)				
Observations	4,822	4,822	4,822	4,822				
R-squared	0.117	0.119	0.119	0.120				
Panel B. Rurales								
Migrant	0.557***	0.394**	0.731***	0.562**	0.514***	0.373*	0.412*	0.219
	(0.187)	(0.198)	(0.204)	(0.226)	(0.194)	(0.205)	(0.213)	(0.234)
Migrant × Panic			-0.994***	-0.841***			-0.644**	-0.474
			(0.289)	(0.301)			(0.291)	(0.302)
Migrant × Post Panic			-0.060	0.040			0.870***	0.978***
			(0.246)	(0.253)			(0.279)	(0.285)
Enganche		0.617***		0.457*		0.513**		0.513**
		(0.234)		(0.247)		(0.236)		(0.247)
Observations	8,860	8,860	8,860	8,860	8,860	8,860	8,860	8,860
R-squared	0.053	0.054	0.054	0.055	0.063	0.064	0.065	0.066
Panel C. Passports								
Migrant	-2.143***	-2.252***	-1.953***	-2.096***	-2.103***	-2.216***	-2.204***	-2.381***
	(0.508)	(0.509)	(0.518)	(0.523)	(0.513)	(0.514)	(0.524)	(0.528)
Migrant × Panic			-0.958***	-0.807***			-0.675**	-0.486
			(0.288)	(0.299)			(0.290)	(0.300)
Migrant × Post Panic			-0.092	0.010			0.622**	0.740**
			(0.253)	(0.260)			(0.291)	(0.296)
Enganche		0.618***		0.466*		0.627***		0.594**
		(0.237)		(0.249)		(0.241)		(0.251)
Observations	4,901	4,901	4,901	4,901	4,901	4,901	4,901	4,901
R-squared	0.059	0.060	0.061	0.062	0.077	0.078	0.079	0.080
Skill level categories	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region of birth categories	Yes	Yes	Yes	Yes	No	No	No	No
Birth year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth state FE	No	No	No	No	Yes	Yes	Yes	Yes

Source: Mexican Border Crossing Records - Microfilm publication N° A3365 and López-Alonso (2015).
Notes:* = Significant at 10% level;** = Significant at 5% level;*** = Significant at 1% level. Robust standard errors in parenthesis.

To assess the effect of the *enganche* on selection patterns across periods, we expand Equation 4.2 as follows:

$$\begin{aligned} height_i =& \beta + \Phi_1 migrant_i + \Phi_2 migrant_i \times panic + \Phi_3 migrant_i \times panic^{post} \\ &+ \Phi_4 enganche_i + \mathbf{X}_i'\theta + \alpha_c + e_i. \end{aligned} \tag{4.4}$$

Where $enganche_i$ is the same indicator variable previously defined. Equation 4.4 controls for the *enganche* effect (Φ_4) and provides estimates of Φ_1, Φ_2 and Φ_3 for *non-enganche* migrants. Basically, the estimated coefficient Φ_4 is the

average difference in height between *enganche* and *non-enganche* migrants in the pre-Panic and post-Panic, because the share of *enganche* migrants was very small during the Panic.

Column 3–Panel B of Table 4.7 shows that migrants were 0.7 centimeters taller than the rurales in the pre-Panic period. When controlling for the *enganche* migrants, who were the tallest individuals in the migrant sample, we observe a less positive selection relative to the military elite (column 4–Panel B).[33] This effect accounts for 23 percent of the average difference in height between migrants and rurales. A similar pattern is observed with the other comparison samples as well. Therefore, we can argue that the *enganche* influenced the selection into migration in the pre-Panic period.

During the Panic of 1907, the *enganche* effect toward a more positive selection is lost, because there were very few *enganche* migrants. Column 4 of Table 4.7 shows that *non-enganche* migrants became less positively selected: they were 0.8 centimeters shorter than their pre-Panic counterparts. This result reveals that unobserved forces in combination with the absence of the *enganche* effect could explain the less positive selection observed during the Panic. In addition, the estimated coefficient Φ_3 remains insignificant, implying that the *enganche* did not influence the selection of migrants in the post-Panic. While the share of *enganche* migrants increased in the post-Panic period (from 1 to 13 percent), it was far from pre-Panic levels (36 percent). This suggests that the reimplementation of this labor institution might have been gradual and thus, its influence was not significant in the short-run after the Panic.

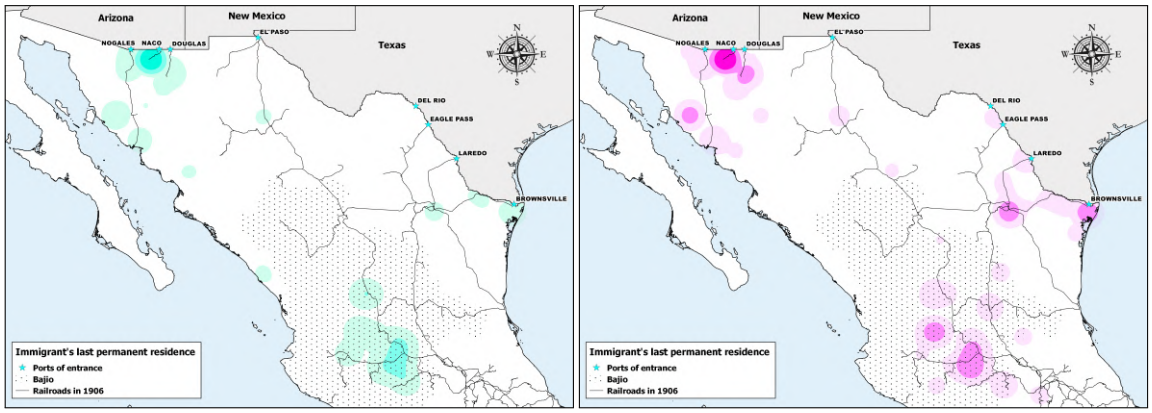
We regress Equation 4.4 including state of birth fixed effects to control for unobserved factors across states that might have influenced these shifts in selection into migration. Column 8-Panel B shows that the *enganche* effect remains strong and statistically significant. Although the estimates for the pre-Panic period

[33] Using the estimated coefficient Φ_4 to approximate the selection pattern of *enganche* migrants in each period would be inaccurate, because the share of *enganche* migrants varies across periods.

are not statistically significant, the coefficients' size suggest that the *enganche* accounted for 46 percent of the average difference in height between migrants and the military elite. The results also confirm that in the post-Panic period, migrants became more positively selected than their pre-Panic peers (column 7). This difference of about 0.6 to 0.8 centimeters is captured in Figure 4.2. In Section 4.10, we provide evidence suggesting that the presence of regional droughts may have driven this shift.

Finally, it may be the case that the observed shifts in selection patterns were induced by changes in the migrants' locations of birth during the Panic. Figure 4.4 shows that in both periods (pre-Panic and Panic), the migrant sample was concentrated in the south of the Bajio (Guanajuato, Jalisco and Michoacan) and in the northern states of Sonora, Nuevo Leon and Tamaulipas (see Figure 4.5 in Section 4.10 for guidance). Hence, the less positive selection during the crisis did not arise from changes across regions of origin.

Figure 4.4: Spatial distribution of the migrant sample



(a) Pre-Panic (October 1906 - July 1907) *(b) Panic (August 1907 - January 1908)*

Source: Mexican Border Crossing Records - Microfilm publication No. A3365.
Note: We consider individuals that had reached their terminal height: individuals between 22 and 65 years old.

4.9 Conclusion

Using a unique data set consisting of individual border crossings, military recruitment records and passport applications, we estimate the selection of Mexican migration to the United States in the beginnings of the flow (1906–08). The results suggest that, on the basis of height, the first Mexican migrants were not drawn from the lowest ranks of the height distribution of the Mexican population. On the contrary, Mexico sent its tallest and most physically productive laborers to the United States. This positive selection relative to the Mexican working class continued in the early 1920s (Kosack & Ward, 2014). Consequently, the persistent drain of high quality laborers arises as a key element to understand the outstanding economic expansion of the American Southwest in the early twentieth century (Gratton & Merchant, 2015, p. 528).

In addition, migrants from poorer regions were disproportionately drawn from the upper ranks of the height distribution. While this finding corroborates the importance of liquidity constraints in generating positively selected migration (Belot & Hatton, 2012), we focus on institutions as mechanisms that can influence migrant selection (Abramitzky & Boustan, 2017, p. 1325). We show that the *enganche*, a search-matching mechanism used by American recruiters to transport and allocate laborers in the United States, influenced the positive selection of Mexican migration, and the absence of this labor institution was associated with a less positive selection into migration. The persistence of institutions like the *enganche* may have unpredicted long-lasting effects on migrant selection patterns, and therefore on the economic development of migrant-sending regions. We believe this is a crucial area of future research.

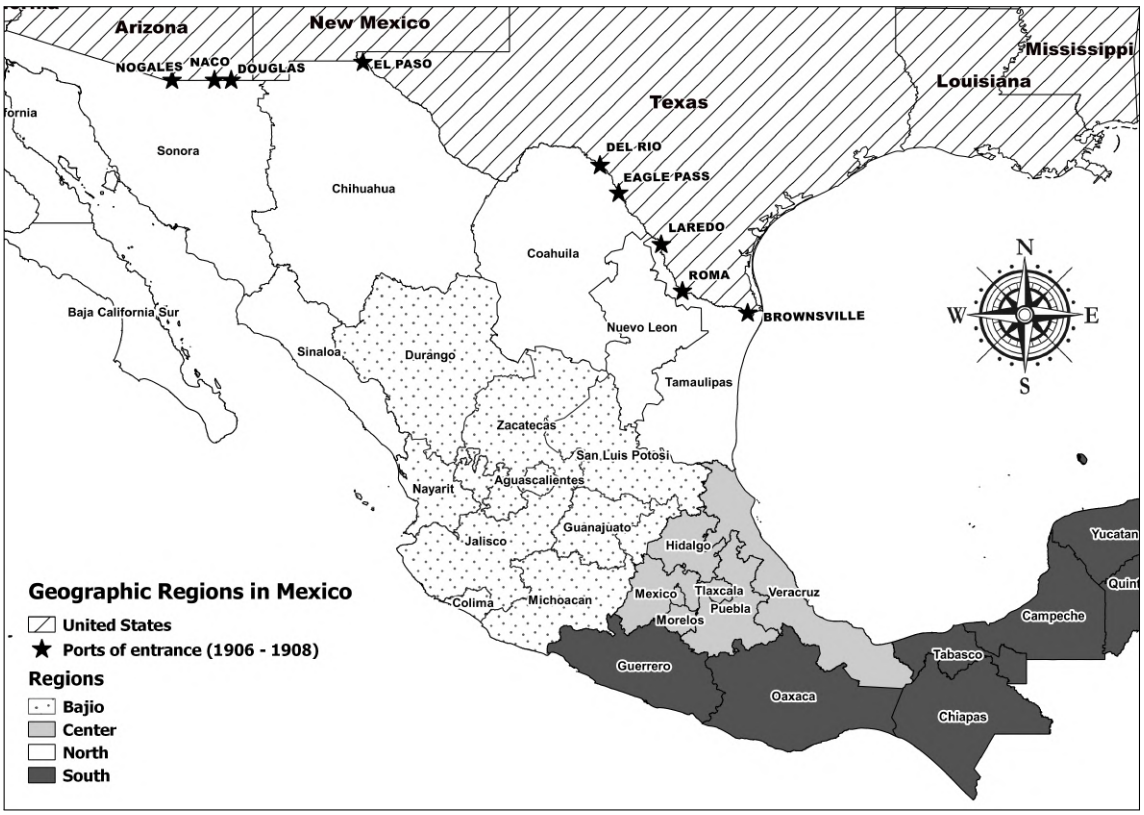
The main lesson to take from our research is that, in the absence of legal immigration restrictions, random shocks from both demand or supply side can change migrant self-selection in the short-run. This adjustment mechanism operated a the local level, confirming that the decision to migrate is influenced by

local conditions (Spitzer & Zimran, 2018). While structural changes in the national economic environment would predict shifts in selection over time, we argue that unanticipated events such as financial crises have the potential to modify selection patterns very fast at the local level. We join previous recommendations highlighting that to understand properly migrant self-selection, we must evaluate the quality of migrants relative to their local environments.

4.10 Appendix

A. Figures for guidance

Figure 4.5: Mexican regions and entrance ports (1906–08)



Source: The immigrant's regions of birth were classified following López-Alonso (2015, p. 127).

B. Identification of the *enganche*

To identify the *enganche*, we quantify the number of migrants (i) by port of entrance (p), year-month of crossing (t), municipality of origin (o), and county of destination (d):

$$w_{ptod} = \sum i_{ptod}. \tag{4.5}$$

We standardize the size of each migration flow (w_{ptod}) using the mean (μ_{pdo}) and standard deviation (σ_{pdo}) of the corridor (w_{pdo}) to which the flow belongs:

$$z_{ptod} = (w_{ptod} - \mu_{pdo})/\sigma_{pdo}. \tag{4.6}$$

The *z-scores* (z_{ptod}) allow us to identify unusual monthly crossing peaks in each migration corridor. Following Clark (1908) and Durand (2016), American recruiters commonly hired between 30 and 400 migrants depending on the nature of the jobs and season of the year. Therefore, we identified the *enganche* flows with the following criteria:

$$enganche_{ptod} = \begin{cases} 1 & \text{if } w_{ptod} \geq 30 \text{ and } z_{ptod} \geq 1 \\ 0 & \text{if } otherwise. \end{cases} \tag{4.7}$$

We identify flows of at least 30 migrants registered at the same port of entrance, in the same specific month, reporting the same origin (Mexican municipality) and destination (US county) locations; and which size was at least one standard deviation above the average size of the flows in each migration corridor. Finally, we match the identified *enganche* flows with the final migrant sample: all individuals belonging to an *enganche* flow are considered *enganche* migrants.

C. Shocks from the supply side: droughts in Mexico

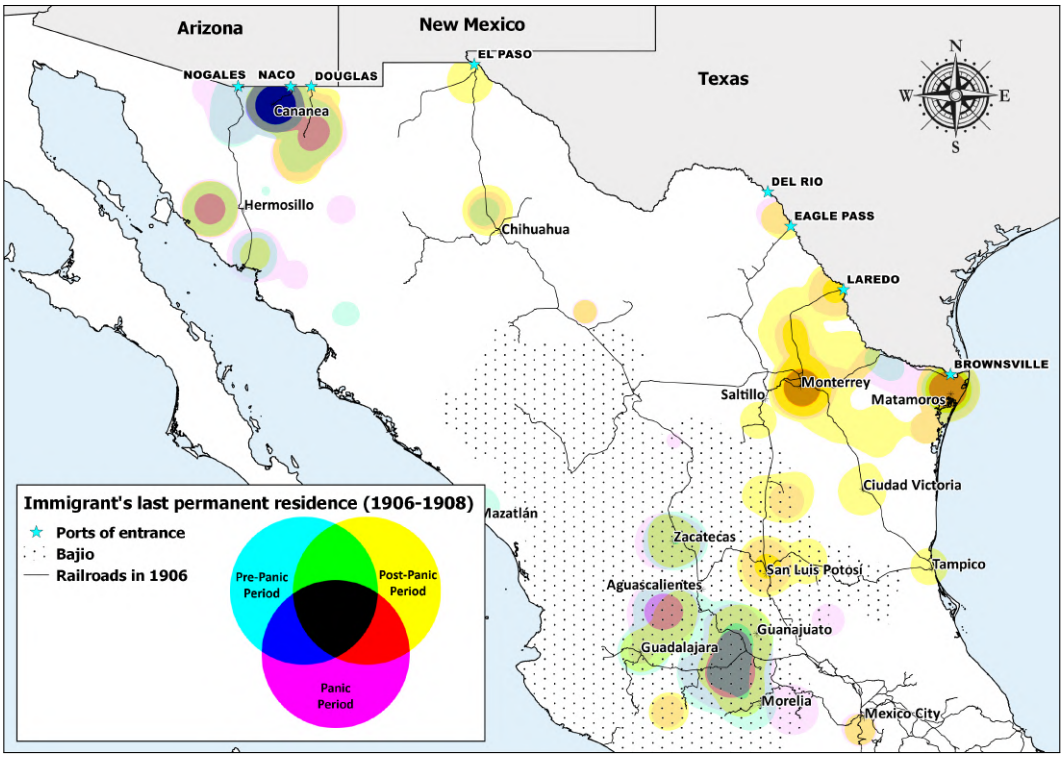
Results in Table 4.7 show that the selection into migration of Mexicans changed after the Panic of 1907, and that this shift arose from unobserved factors across states. The forces explaining why migrants became more positively selected in the post-panic period could come from both demand or supply side. On the one hand, as a result of the shock, the post-panic labor demand in the US could have been organically different. This new composition might have demanded a specific migrant profile. On the other hand, factors in Mexico might have pushed taller individuals to emigrate during the post-panic period. We explore the latter scenario looking at climate shocks that might have influenced migrant self-selection after the Panic of 1907. Contreras (2005, p. 123), Clark (1908, p. 473), and Mayet et al. (1980, p. 757) document that the states of Chihuahua, Nuevo León, Quéretaro, San Luis Potosí and Zacatecas experienced droughts in 1907 and 1908, causing important crop losses in some areas (Cardoso, 1980, p. 12). Moreover, Figure 4.6 shows that migrants came disproportionately from these states during the post-panic period.

We identify the presence of droughts at the municipality level using the Mexican Drought Atlas (Stahle et al., 2016). It provides reconstructions of a self-calibrating Palmer Drought Severity Index (PDSI) on a 0.5° latitude/longitude grid centered over Mexico from AD 1400-2012. We consider that a municipality experienced droughts if the estimated PDSI was -2.0 or lower. According to Wells et al. (2004), these values represent moderate to severe droughts.[34]

Figure 4.7 shows that droughts affected specific states within regions: the northern Bajio, the eastern states of the North region and the Yucatan peninsula. The PDSI estimates confirm the presence of droughts in the states mentioned

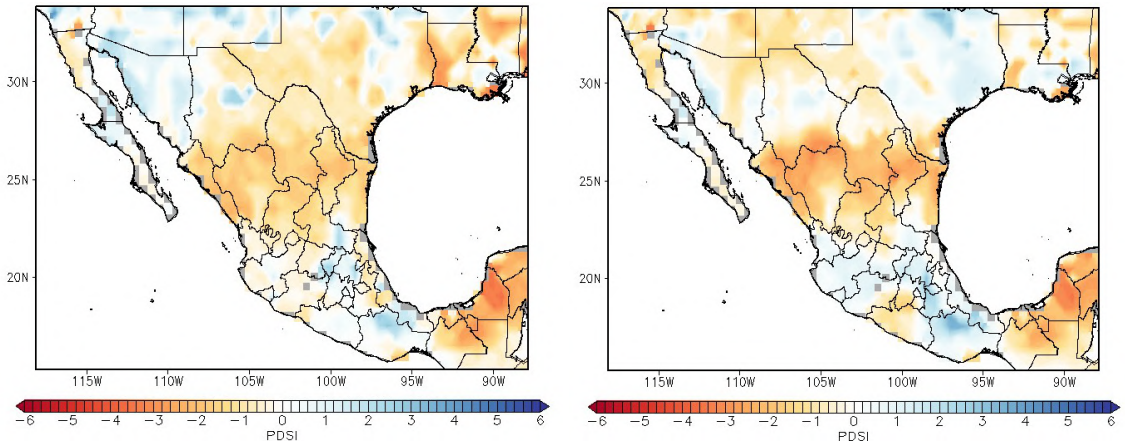
[34]The Palmer Drought Severity Index (PDSI) uses temperature and precipitation data to estimate relative dryness. It is a standardized index that spans from -6 (dry) to +6 (wet). However, values below -4.0 represent extreme droughts while values above +4.0 represent extreme wet spells (Wells et al., 2004).

Figure 4.6: Spatial distribution of the immigrant sample (last permanent residence)



Source: Mexican Border Crossing Records - Microfilm publication N° A3365.
Note: We consider individuals that had reached their terminal height: individuals between 22 and 65 years old. Pre-Panic (Oct 1906 - Jul 1907), Panic (Aug 1907 - Jan 1908) and Post-Panic (Feb 1908 - Dec 1908).

Figure 4.7: Droughts in Mexico 1907–08. Palmer Drought Severity Index (PDSI)



(a) Jan - Dec 1907

(b) Jan - Dec 1908

Source: Stahle et al. (2016). Note: The Palmer Drought Severity Index (PDSI) uses temperature and precipitation data to estimate relative dryness. It is a standardized index that spans from -6 (dry) to +6 (wet). However, values below -2.0 represent moderate droughts while values above +2.0 represent moderate wet spells. The panel shows the presence of regional droughts in 1907 and continues in 1908. The average drought severity index (at the state level) in 1908 was -2.7 (Chihuahua), -2.3 (Coahuila), -2.4 (Durango), -2.4 (Nuevo León), 1.4 (Queretaro), -0.9 (San Luis Potosi), -2.3 (Sinaloa), -2.2 (Tamaulipas), -0.8 (Zacatecas).

by the historical literature except for Queretaro.[35] The PDSI also captures that municipalities of Coahuila, Durango, Sinaloa and Tamaulipas were affected by these climate shocks. Precisely, the municipality level estimates allows us to identify droughts accurately in regions covering more than one state.

To test if droughts influenced the selection into migration, we expand Equation 4.2 as follows:

$$\begin{aligned} height_i =& \beta + \Phi_1 migrant_i + \Phi_2 migrant_i \times panic + \Phi_3 migrant_i \times panic^{post} \\ &+ \Phi_4 migrant_i \times panic^{post} \times drought + \Phi_5 enganche_i + \mathbf{X}_i'\theta \\ &+ \alpha_c + \gamma_s + e_i. \end{aligned} \tag{4.8}$$

Where *drought* is a dummy variable that takes the value of 1 if the migrant's location of last residence (municipality) experienced droughts (PDSI values of -2.0 or lower) and zero otherwise. Since the *enganche* was restored after the panic, we include an indicator variable for migrants that crossed the border through this labor institution ($enganche_i$). Equation 4.8 includes state of birth fixed effects (γ_s) instead of birth region categories, because droughts did not affect states homogeneously. Moreover, large states were partially affected, thus we want to capture the effect of droughts on local selection. Everything else equal, the estimated selection pattern in the locations experiencing droughts during the post-panic period is $\Phi_1 + \Phi_3 + \Phi_4$.

Table 4.8 shows that migrants from municipalities experiencing droughts were taller (at least 0.7 centimeters) than their counterparts from non-drought municipalities (estimated coefficient Φ_4): they were more positively selected relative to the comparison samples. We can see that post-panic migrants were 0.8 centimeters taller than their pre-panic peers (column 1), but when controlling

[35]The municipalities belonging to the states of San Luis Potosi and Zacatecas present PDSI estimates close to our threshold, and thus we considered them as municipalities that experienced droughts. However, the municipalities of Queretaro present a positive estimates (1.4), which imply the presence of mild wet spells.

for the droughts effect the coefficient size reduces (column 2). In other words, the presence of droughts accounts for 28 percent of the differences in height between pre-panic and post-panic migrants. Additionally, we control for the *enganche* effect to obtain the net effect: droughts accounted for 15 percent of the stronger positive selection observed after the Panic of 1907. The same pattern holds relative to the passport holders, however the estimates are not statistically significant, potentially due to sample size constraints.

Since droughts impact homogeneously the population of an affected location, it is likely that laborers/peasants relatively taller than pre-panic migrants were pushed to emigrate during the post-panic due to the poor harvests. Hence, the observed positive selection might be a result of two overlapping forces: the reactivation of the American financial system and the persistence of regional droughts in Mexico.

Table 4.8: Impact of droughts on self-selection patterns.
Dependent variable: height (centimeters)

	1	2	3	4	5	6
	Rurales	Rurales	Rurales	Passports	Passports	Passports
Migrant	0.412*	0.409*	0.224	-2.204***	-2.212***	-2.382***
	(0.213)	(0.213)	(0.234)	(0.524)	(0.523)	(0.528)
Migrant × Panic	-0.644**	-0.627**	-0.465	-0.675**	-0.656**	-0.475
	(0.291)	(0.291)	(0.302)	(0.290)	(0.290)	(0.300)
Migrant × Post Panic	0.870***	0.620**	0.736**	0.622**	0.396	0.524
	(0.279)	(0.314)	(0.321)	(0.291)	(0.321)	(0.327)
Migrant × Post Panic × Drought		0.826**	0.789*		0.785*	0.736*
		(0.410)	(0.412)		(0.420)	(0.422)
Enganche			0.492**			0.574**
			(0.248)			(0.252)
Observations	8,860	8,860	8,860	4,901	4,901	4,901
R-squared	0.065	0.066	0.066	0.079	0.080	0.081
Skill level categories	Yes	Yes	Yes	Yes	Yes	Yes
Region of birth categories	No	No	No	No	No	No
Birth year FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth state FE	Yes	Yes	Yes	Yes	Yes	Yes

Source: Mexican Border Crossing Records - Microfilm publication N° A3365 and López-Alonso (2015). Notes: * = Significant at 10% level; ** = Significant at 5% level; *** = Significant at 1% level. Robust standard errors in parenthesis. The omitted categories are unskilled workers and locations without droughts.

Chapter 5

One step at the time: wage premium for stage migration

Abstract

I estimate the wage premium for stage migration exploiting a novel data set reporting birthplace, last residence and destination of first-time Mexican immigrants. The micro data reveal that, in the beginnings of the Mexico-US migration (1906-08), 36% of the immigrants moved within Mexico before crossing the border. I estimate correlations between stage migration and potential wage at the destination controlling for the individual's age, literacy, marital status, sex and birthplace. My results show that stage migrants may have experienced a wage premium of 3 to 9%, suggesting that migration methods could explain disparities in the immmigrant's labor-market performance at the destination.

5.1 Introduction

In both past and present, immigrants work toward earnings convergence with natives. However, the immigrant labor-market performance varies substantially by sending country and across generations. Previous research acknowledges that the process of assimilation into the labor market is influenced by factors of diverse nature that can increase, maintain or reduce the earnings gap between immigrants and natives (Abramitzky & Boustan, 2017, p. 1329–30). Although we know that language proficiency (Bleakley & Chin, 2004, 2010; Chiswick & Miller, 2010; Lleras-Muney & Shertzer, 2015; Ward, 2019b), residential segregation (ethnic enclaves) and networks (Collins & Margo, 2000; Cutler et al., 2008; Damm, 2009; Edin et al., 2003; Eriksson & Ward, 2018; Eriksson, 2018; Munshi, 2003), labor-market discrimination (Moser, 2012; Orraca-Romano & García-Meneses, 2016; Reimers, 1983), and self-presentation (Abramitzky et al., 2016; Biavaschi et al., 2017; Carneiro et al., 2016) influence the immigrants' labor-market assimilation, there other forces that remain unexplored. In this Chapter, I show that differences in the chosen migration method may explain disparities in the labor-market performance of Mexican immigrants. Specifically, I evaluate if stage migration (migrating in steps) improves the immigrant's wage in the United States.

To study the effect of stage migration, I exploit original micro data that consist of Mexican immigrants that crossed the Mexico-US land border from July 1906 to August 1908. The data source is the Mexican Border Crossing Records (MBCRs) publication N° A3365, which contain microfilms reproducing two-sheet manifests that list Mexican immigrants at nine entrance ports (see Figure 5.1). These documents report rich demographic, socioeconomic and geographic data for each immigrant. I use the immigrant's reported birthplace, last permanent residence and final destination to identify two types of migration: stage migration (immigrants whose reported birthplace and last permanent residence in Mexico is different) and direct migration (immigrants whose reported

birthplace and last permanent residence in Mexico is the same). In other words, I identify individuals who migrated within Mexico before crossing the border and individuals who migrated directly to the United States. The micro data reveal that 36% of the working-age individuals migrated in stages. In addition, stage migration occurred mostly within regions—that is, stage migration followed local dynamics. Immigrants moving between regions before crossing the border were underrepresented in the total outflow.

Figure 5.1: Migration regions and entrance ports (1906–08)

Source: Based on Durand (2016, p. 28) and Mexican Border Crossing Records. Microfilm publication N° A3365.

I complement the immigrant sample with data on wages in the county of destination. The local-level data allow me to capture different factors across space or labor markets that may generate disparities in immigrant performance (Abramitzky & Boustan, 2017, p. 1333). I collect wage data for Mexican immigrants and average immigrants from miscellaneous sources, and I impute it to each immigrant based on the reported county of final destination. In addition, I adjust the imputed wage by sex, age cohort (14 to 18 years and 18 years and over)

and region of destination in the United States. I obtain the adjustment factors by region from industries that reported participation of Mexican immigrants from both sexes and/or age cohorts (US Immigration Commission, 1911b). The adjusted wages capture age-earning patterns and gender penalty gap within and across labor markets. To some extent, this method overcomes some of the limitations of the occupation-based income scores approach, in which the income measure is fixed for all individuals reporting the same occupation (Inwood et al., 2019, p. 114).

Migrating in stages may affect the migrant's wage at the final destination through different channels. For example, stage migrants may gain access to networks and information than otherwise they would not have. Recent literature argues that the assistance provided by immigrant networks varies across regions of migrant-sending countries (Giulietti et al., 2018; Munshi, 2003). Also, stage migration can influence the accumulation of productive skills relevant to the labor market of final destination. I use ordinary least squares (OLS) to estimate correlations between stage migration and potential wage at the final destination. To obtain the best results possible, I control for the migrants' sex, age, literacy and marital status. I also include region-of-birth fixed effects to capture wage differences between stage and direct migrants from the same origin.

My results suggest that individuals that migrated in stages before crossing the border may have experience a wage premium of 3 to 9%. The size of the premium for stage migration is similar the premium for English fluency during the same period (Ward, 2019b, p. 3). Although this finding do not provide information about the effect of stage migration on immigrant assimilation *per se*, it do shows that among Mexican immigrants stage migration may have influenced significantly the starting point of the assimilation process. In other words, stage migrants may have faced an smaller earnings gap relative to natives. Abramitzky et al. (2014) argue that the labor-market performance of first-generation immigrants

can persist across generations. Therefore, the chosen migration method could have long-lasting consequences for the earnings convergence with natives. The main contribution of the Chapter is to show that stage migration—as part of the international migration process—could explain disparities in the immigrant labor-market performance at the destination. However, I acknowledge that it is not always possible to observe stage migration spells within the home country.

The remainder of the Chapter is organized as follows. I describe the data in Section 2. In Section 3, I address the characteristics of stage migration in Mexico by the end of the 1900s. I present the empirical strategy and results in Section 4 and Section 5, respectively. I conclude in Section 6.

5.2 Data

To disentangle the effect of stage migration on the immigrant labor-market performance, I exploit a sample of Mexican immigrants that crossed the US border during 1906–08 and data on daily wages at US counties.

5.2.1 Immigrant Sample

The immigrant sample comes from the Mexican Border Crossing Records (MBCRs) publication N° A3365. This data source consists of microfilms containing manifests listing arriving aliens at nine ports of entrance along the Mexico-US border.[1] The sample consists of 10,895 individuals that crossed the border from July 1906 to December 1908. To transcribe the data, I follow a stratified sampling plan in which the strata are the entrance ports. I select the subsamples of each stratum using an equal probability systemic sampling. Chapter 2 offers a formal description of the sampling plan. It describes the criteria that I follow to identify immigrants, and it presents evidence suggesting that the sample is representative

[1]Publication Title: Lists of Aliens Arriving at Brownsville, Del Rio, Eagle Pass, El Paso, Laredo, Presidio, Rio Grande City, and Roma (Texas) from May 1903 to June 1909; and at Aros Ranch, Douglas, Lochiel, Naco, and Nogales (Arizona) from July 1906–December 1910.

for the period under analysis. In addition, it addresses the potential limitations of the data.

The manifests report rich individual-level data, but in this Chapter I exploit the immigrant's demographic characteristics (sex, age, literacy, occupation and marital status) and geographic information (birthplace, last permanent residence and final destination). The geographic data allow me to identify immigrants that moved within Mexico before crossing the border—individuals whose reported birthplace and last permanent residence is different—and those who migrated directly to the United States—that is, individuals whose reported birthplace and last permanent residence is the same. In other words, it is possible to characterize immigrants based on their chosen migration method: direct or stage migration. I classified the geographic data as Mexican municipalities and American counties to capture the characteristics of stage migration at the local level. Clearly, to address stage migration, it is compulsory to count with full geographic information for each immigrant. Table 5.1 shows that 67% of the transcribed crossings (7,313 observations) count with full classified geographic information. As can be noticed, underreporting of locations of birth is relatively high (9.9%), which could be an additional source of bias. The final weighted sample consists of 13,455 immigrants.

5.2.2 Wage Data

As noted by Inwood et al. (2019, p. 114), most literature addressing the immigrant's performance in the United States rely on occupations and occupation-based income scores. This methodology presents diverse limitations since the income measure is fixed for all individuals reporting the same occupation. Indeed, individual characteristics that influence wages such as age are neglected in occupation-based income scores (see Inwood et al. (2019) for a full discussion). Although the MBCRs report the immigrant's last occupation in Mexico, in my

Table 5.1: Sample refinement

	Obs.	*Share (%)*
Transcribed crossings	10,895	100
Return immigrants	718	6.6
Non-immigrants	1,045	9.6
Immigrants	9,083	83.4
Last residence in Mexico		
Unreported	405	3.7
Not classified	10	0.1
A. Classified as Mexican municipalities	**8,668**	**79.6**
Final destination in the United States		
Unreported	203	1.9
Not classified	82	0.8
B. Classified as American counties	**8,798**	**80.8**
C. A ∩ B	**8,420**	**77.3**
Place of birth in Mexico		
Unreported	1,087	9.9
Not classified	20	0.2
D. Full classified geographic information	**7,313**	**67.1**

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: *Return immigrants* refer to Mexican individuals whose final destination was in Mexico. *Non-immigrants* refer to Mexican individuals whose final destination and last permanent residence was in the United States. *Immigrants* refers to Mexican individuals whose last permanent residence was in Mexico and final destination was in the United States. C = Mexican immigrants whose last permanent residence and final destination was reported and classified in Mexican municipalities and US counties, respectively. D = immigrants with full classified geographic information (birthplace, last residence and destination).

sample 60 percent of the individuals reported *laborer* as occupation; and 21 percent reported not having an occupation. Thus, variation across occupations is almost nonexistent. To overcome this problem, I construct potential minimum-daily wages for Mexican immigrants exploiting the geographic variation in the US county reported as final destination. This approach relies on differences in wage levels across and within regions arising from the location of industries or economic specialization at the local level. To some extent, this methodology captures unobservable skills of immigrants revealed by their chosen destination.

I collect wage data at the county level from three sources. First, from the historical research of Clark (1908), I obtain wages for Mexican immigrants in diverse counties of Arizona, California, New Mexico, and Texas. Second, I obtain from the Immigration Commission Reports of 1910 wage levels for cities and counties in California, Illinois, Kansas, New York, Ohio, Oklahoma and Wisconsin

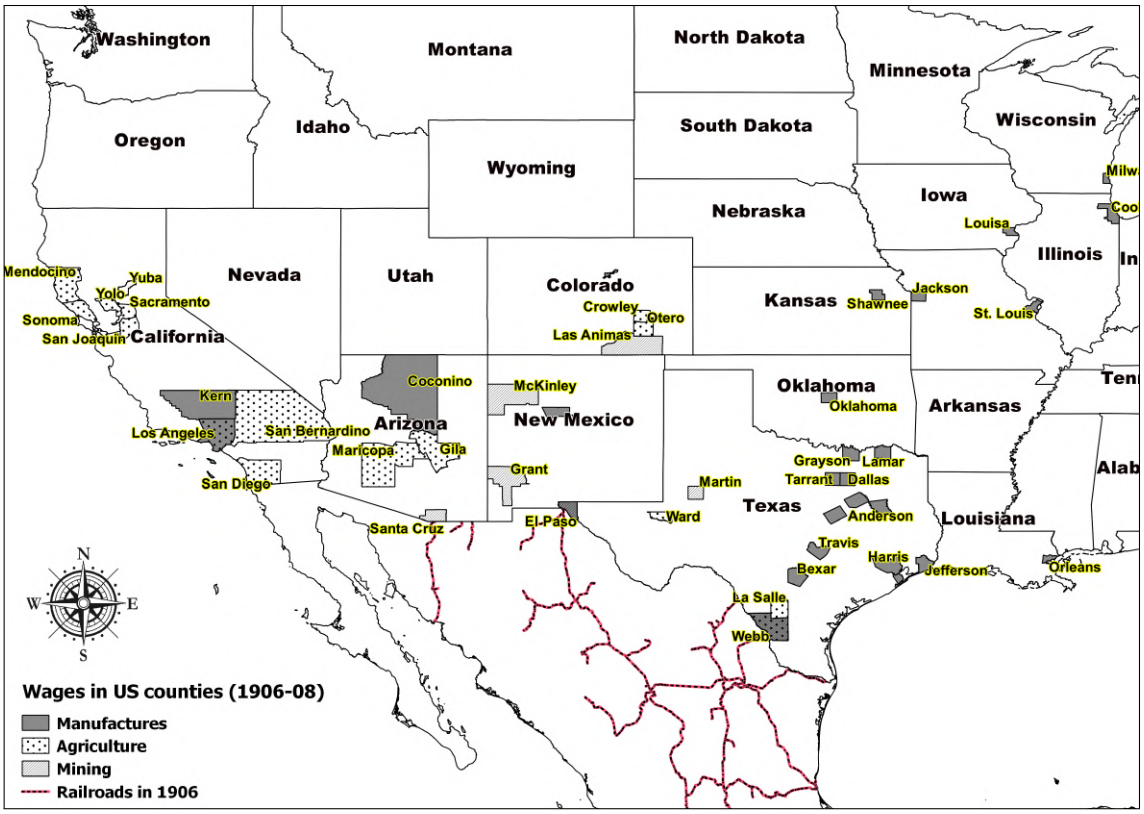
(US Immigration Commission, 1911b). Third, I use the Census of Manufactures of 1905 and Bulletins of the Bureau of Labor of 1907 to obtain wages for counties in California, Louisiana, Missouri, Ohio, Oklahoma, Texas, and Wisconsin (U.S. Department of Commerce and Labor, 1906a,b, 1907). The criteria to impute fixed wage levels by county of destination was the following.

I prioritize wage data for Mexican immigrants over average-immigrant wage data. I also keep the lowest wage level reported to proxy for minimum wages. In destinations for which I did not obtain data, I assigned the wage level from the nearest county with available data. To account for crucial differences at the individual level that affect wages, I adjust the imputed wage levels by sex and age cohort. The adjustment was based on regional wage differences by sex and age cohort (14 to 18 years and 18 and over) reported in the US Immigration Commission (1911b). I obtain the factors of adjustment from industries that reported participation of Mexican immigrants from both sexes and/or age cohorts. This guarantees that the adjustment factors capture wage differentials derived from sex and age within the same industry.[2] Figure 5.2 depicts the counties and economic sectors for which I obtained wage data.

This approach presents two characteristics that overcome the limitations of occupation-based income scores. First, all wages are for the years 1905 to 1908: they correspond to the period of analysis. Thus, estimates using these data will not be biased due to changes in labor markets over time. Second, to some extent it considers age-earning patterns and gender penalty gaps. I use these data to explore if stage migrants experienced higher or lower wages—at the destination—relative to their peers that crossed the border in one move.

[2]The industries used to adjust the wage levels are: clothing manufacture (Northeast); clothing manufacture, and butcher and meat packing (Midwest); clothing manufacture, and iron and mining (South); and canning (West).

Figure 5.2: Wages in US counties (1906–08)



Source: Based on data reported by Clark (1908); US Immigration Commission (1911b); U.S. Department of Commerce and Labor (1906a,b, 1907).

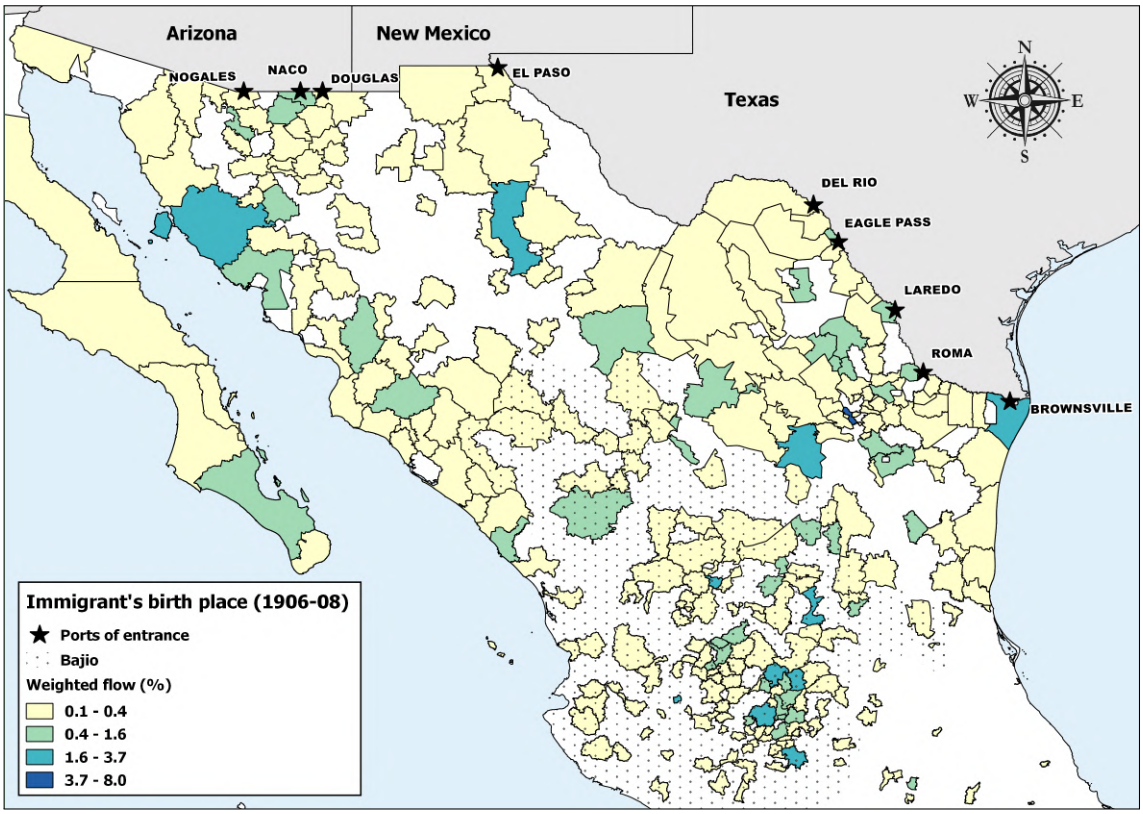
5.3 Stage migration in Mexico before 1910

The beginnings of the Mexico-US migration (1884–1910) overlaps with a constant northward drift of labor within Mexico. This internal migration has been documented by previous literature, arguing that the economic development of northern Mexico, understood as the expansion of railways and export-oriented industries (agriculture and mining), pulled a constant flow of workers from the central plateau (Bajio) to the border states. However, labor market pressures jointly with higher wages in the American Southwest might have motivated immigrants to continue moving north (Clark, 1908, p. 470; Cardoso, 1980, p. 17; Durand, 2016, p. 61; Rosenzweig, 1965, p. 448).

Although the presence of internal migration suggests that migration to the United States might have been a two-step process, we know little about its

characteristics and relevance for the Mexico-US migration flow. According to Durand (2016, p. 31), the reason is the complexity to identify accurately the origin of those individuals that ultimately crossed the border during this period (Durand, 2016, p. 31). To fill this gap in the literature, I use the immigrant's reported birthplace and last permanent residence to identify two types of migration: stage migration (individuals who migrated within Mexico before crossing the border) and direct migration (individuals who migrated directly to the United States). This section presents the main characteristics of stage migration at the local level.

Figure 5.3: Immigrant's birthplace (1906–08)



Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Spatial distribution of the Mexican-American migration flow from July 1906 to December 1908. The polygons display the immigrant's birth place (municipalities) and their shares in the overall weighted flow (quartiles calculated with Jenks natural breaks classification method). The shaded area covers the states of the Bajio region.

Figure 5.3 shows that the immigrants' municipalities of birth were more diversified than the locations of last residence (see Figure 2.4 in Chapter 2), suggesting the presence of migration within Mexico. Table 5.2 presents the composition of the weighted migration flow and confirms the presence and

importance of stage migration. In fact, 40% of the immigrants moved within Mexico before migrating to the United States. Most of this stage migration took place within regions, but patterns differed. In the Bajio, it occurred mostly within states (between municipalities), whereas in the Border region stage migration within and between states was equally common. In contrast to previous literature, the micro data reveals that only 14% of the immigrants moved between regions before migrating to the United States. Most of the interregional migration (82%) did occur between the Bajio and the Border region.

Table 5.2: Decomposition of Mexican emigration to the United States (1906–08)

	Weighted Flow	*Share (%)*
Panel A. Migration Flow		
Total (A+B)	**13,455**	**100**
A. Direct migration	**8,122**	**60.4**
B. Stage migration	**5,333**	**39.6**
C. Within regions	3,485	25.9
D. Between regions	1,848	13.7
Panel B. Stage migration		
C. Within regions	**3,485**	**100**
Border	2,953	84.7
Bajio	523	15.0
Center	4	0.1
Southeast	5	0.2
D. Between regions	**1,848**	**100**
Bajio – Border	1,516	82.0
Center – Border	88	4.8
Southeast – Border	27	1.5
Other	217	11.7
Panel C. Local stage migration		
E. Border region (F+G)	**2,953**	**100**
F. Between states	1,488	50.4
G. Between municipalities (within states)	1,465	49.6
H. Bajio region (I+J)	**523**	**100**
I. Between states	162	31.0
J. Between municipalities (within states)	361	69.0

Source: Mexican Border Crossing Records. Microfilm publication number A3365.
Note: A = immigrants that were born in the municipality reported as last permanent residence. C = immigrants that were born in the region reported as last permanent residence, but in a municipality or state different from the reported as last permanent residence. D = immigrants that were born in a region different from the reported as last permanent residence. E and H = immigrants that were born in the region reported as last permanent residence (Border or Bajio, respectively), but in a municipality or state different from the reported as last permanent residence.

To characterize stage migration, I use as reference the municipalities of Monterrey, Cananea, Chihuahua, and Matamoros. Based on the last residence data, these municipalities in the Border region have the highest shares in the total outflow (12.2, 10.8, 3.6, and 3.4%, respectively). Looking at the immigrants whose last

place of residence was Monterrey, 56% were natives of the city; 21% were born in municipalities within the Border region; 22% were born in Bajio municipalities; and the rest were born in farther locations (see Table 5.3). The cases of Chihuahua and Matamoros present a similar pattern: most of the population that migrated from these cities were locals (did not use stage migration) or were natives from municipalities within the Border region. Migrants born in Bajio municipalities represent only 30.8% and 6.7% of the outflow from Chihuahua and Matamoros, respectively. The outflow from Cananea presents a similar pattern, 75% were born in diverse municipalities of the Border region, and only 16% were born in the Bajio (see Table 5.3).

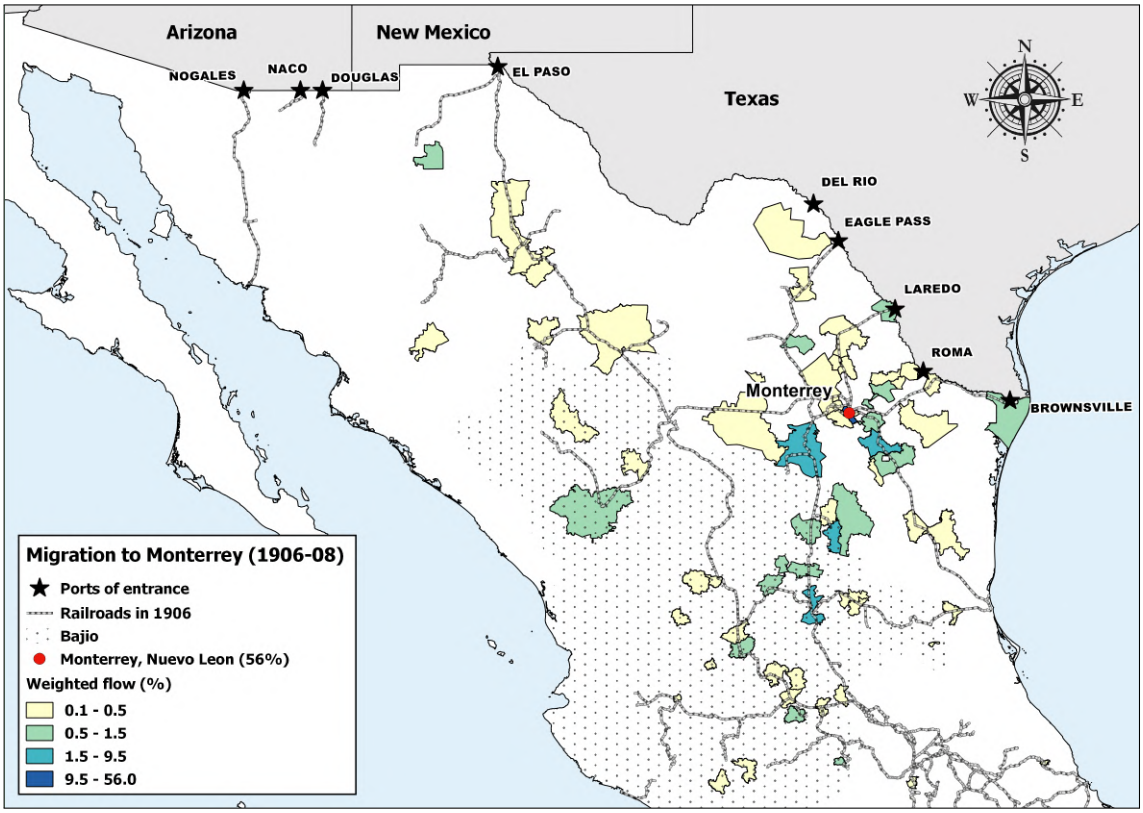
Table 5.3: Decomposition of stage migration (1906–08). Selected cities.

	Weighted Flow	*Share (%)*
Panel A. Immigrants from Monterrey		
Total	**1,565**	**100**
A. Born in Monterrey	876	56.0
B. Born in Border locations	324	20.7
C. Born in Bajio locations	349	22.3
D. Born in the Center or South	16	1.0
Panel B. Immigrants from Chihuahua		
Total	**507**	**100**
A. Born in Chihuahua	234	46.2
B. Born in Border locations	117	23.1
C. Born in Bajio locations	156	30.8
D. Born in the Center or South	0	0
Panel C. Immigrants from Matamoros		
Total	**478**	**100**
A. Born in Chihuahua	326	68.2
B. Born in Border locations	106	22.2
C. Born in Bajio locations	32	6.7
D. Born in the Center or South	14	2.9
Panel D. Immigrants from Cananea		
Total	**1,462**	**100**
A. Born in Cananea	101	6.9
B. Born in Border locations	1,089	74.5
C. Born in Bajio locations	239	16.3
D. Born in the Center or South	33	2.3

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.

Figure 5.4 depicts the spatial distribution of the migration flows to Monterrey. It suggests a potential relationship between stage migration and the access to transportation technologies: several municipalities of birth are crossed by the railway network. However, some of these municipalities are large in extension, questioning if the population had effective access to trains. Looking at origin of the migration flows to Cananea (see Figure 5.6 in the Annex), it is likely that migrants moved towards northern Sonora by boat, but a closer inspection reveals that some non-coastal municipalities were more than 150 km away from the sea. Similarly, Figure 5.7 in the Annex shows that for some Bajio municipalities, the nearest railway was 80 km away. Considering that—in the absence of railways—overland transportation was very precarious (Bignon et al., 2015, p. 1279), it is not clear if the access to transportation technologies induced stage migration.

Figure 5.4: Migration to Monterrey (1906–08)



Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Spatial distribution of the Mexican-American migration flow from July 1906 to December 1908. The polygons display the immigrant's birth place (municipalities) and their shares in the overall weighted flow to Monterrey, Nuevo Leon (quartiles calculated with Jenks natural breaks classification method). The shaded area covers the states of the Bajio region.

Overall, stage migration was present in the Mexico-US migration, and it occurred mostly within regions. Immigrants moving between regions before crossing the border were underrepresented in the total outflow. Indeed, the massive exodus of laborers from central plateau states to northern Mexico, which according to Cardoso (1980, p. 14) started circa 1884 with the railways expansion, is not captured by the MBCRs. The micro data suggest that—by the end of the 1900s—sequential migration followed local dynamics, and interregional migration might have been losing strength as an instrument to reach the United States.

5.4 Wage premium for stage migration

Little is known about the effects of stage migration on the migrant's wage at the final destination. The micro data reported in the MBCRs reveal that 36% of the working-age migrants moved within Mexico before crossing the border. Table 5.4 characterizes the profile of first-time migrants based on their chosen migration method: direct or stage migration. There are clear differences between these groups. Stage migrants present higher shares of female and literate individuals, and they had on hand 2.5 dollars more than their peers migrating directly. These statistics suggest that migrating in stages may be associated with higher levels of human capital. Table 5.4 also shows that on average the potential wage in the United States is the same for both groups. However, these figures could be an artifact derived from the composition of the samples. The following sections offer a systematic analysis to understand the influence of stage migration on the migrant's potential wage at the final destination.

5.4.1 Empirical strategy

Migrating in stages may affect the migrant's wage at the final destination through different channels. For example, stage migrants may gain access to networks and information than otherwise they would not have. Recent literature argues that

Table 5.4: Summary statistics.
First-time migrants over 12 years of age (weighted flow)

	Complete Sample	*Direct Migration*	*Stage Migration*
Average Age (years)	28	27	29
Sex (%)			
Females	23.9	19.5	31.6
Literacy (%)			
Literate	30.8	25.9	39.6
Marital Status (%)			
Single	44.0	43.4	45.0
Married	51.5	53.1	48.6
Widowed	4.5	3.4	6.3
*Cash on hand**	10.9	9.9	12.5
*Potential daily wage in the US**	1.2	1.2	1.2
Total	**8,021**	**5,127**	**2,894**
Total (%)	**100**	**63.9**	**36.1**

Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: I consider only first-time immigrants to identify accurately the socioeconomic characteristics associated with stage migration and not to previous migratory experience. I consider only immigrants of 12 years of age or older, because at the time this was the criteria to identify the working age population in Mexico. In the manifests 3,048 individuals were under twelve years of age or reported previous migratory experience. * US dollars (mean).

the assistance provided by immigrant networks varies across regions of migrant-sending countries. Weak networks provide some information about jobs (wages and working conditions) at the destination, whereas strong networks provide high-quality information and support to migrate (Giulietti et al., 2018; Munshi, 2003). In the context of this research, migrants moving in stages before crossing the border could have gained access to information and support available only at these first-stage places. Therefore, stage migrants could have experienced a wage premium because they counted with more elements to maximize their potential wage at the final destination.

My empirical strategy explores the correlation between stage migration and the migrant's potential wage at the destination. I use ordinary least squares (OLS) to estimate the following equation:

$$ln\,wage_i = \sigma + \Phi\, mig_i^{stage} + \mathbf{x}_i'\phi + \psi_b + \varepsilon_i. \tag{5.1}$$

I regress the wage at the destination ($wage_i$) on an indicator variable for stage migrants (mig_i^{stage}) and a vector of individual characteristics ($\mathbf{x}_i$). The vector of individual characteristics include variables observed at the crossing: sex, age, literacy and marital status. These characteristics control for gender pay penalty, differences in human capital and dependency burden. Age is a proxy for work experience, but also controls for wage differentials between adult and child labor. I also include its quadratic form to capture diminishing returns to work experience. Literacy provides access to information sent by previous migrants, family or friends through written means. The access to experiences of other migrants allow to depend less on word of mouth information, and consequently take better decisions aiming to maximize wages at the destination. Literacy also enables individuals to learn and perform skilled jobs, which on average are paid better.[3] In addition, migrants with low dependency burden (single individuals) can travel longer distances and spend more time searching for higher wages, because they can minimize migration costs more easily than migrants moving as a family (Hatton & Williamson, 1994, p. 535–43). Finally, I include region-of-birth fixed effects (ψ_b) to capture wage differences between two types of migrants from the same origin: those who migrate in stages and those who migrate directly to the final destination.

5.5 Results

Table 5.5 reports the effect of stage migration on the migrant's potential wage. Column 1 shows that on average stage migrants experienced a wage premium of 6.1% relative to their peers from the same region that migrated directly. The stage migration premium increases to 12.8% when controlling for the gender pay penalty (column 2). Age, literacy and marital status have modest effects on the potential wage relative to stage migration (columns 3–5).

[3] Although the manifests report the migrant's labor class (unskilled, skilled and professional), I do not use it since 93% of the migrants were unskilled workers.

Table 5.5: Influence of stage migration.
Dependent variable: log of potential wage at the destination

	1	2	3	4	5	6	7	8
$Migrant^{stage}$	0.061***	0.128***	0.130***	0.122***	0.121***	0.029*	0.026	0.091***
	(0.021)	(0.018)	(0.018)	(0.018)	(0.018)	(0.015)	(0.017)	(0.017)
Female		-0.693***	-0.689***	-0.681***	-0.678***	-0.619***	-0.616***	-0.632***
		(0.021)	(0.021)	(0.021)	(0.022)	(0.020)	(0.018)	(0.021)
Age			0.016***	0.016***	0.017***	0.013***	0.011**	0.014***
			(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Age^2			-0.000***	-0.000***	-0.000***	-0.000***	-0.000**	-0.000***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Literate				0.072***	0.072***	0.043***	0.045***	0.071***
				(0.018)	(0.018)	(0.016)	(0.015)	(0.017)
Single					0.010	0.011	0.001	0.020
					(0.019)	(0.017)	(0.017)	(0.018)
Distance (km)								0.001***
								(0.000)
$Distance^2$ (km)								-0.000
								(0.000)
Constant	-0.071***	0.132***	-0.117**	-0.149**	-0.173**	0.147**	0.169**	-0.390***
	(0.021)	(0.020)	(0.059)	(0.060)	(0.073)	(0.073)	(0.083)	(0.072)
Observations	4,265	4,261	4,241	4,239	4,227	4,227	4,227	4,227
R-squared	0.089	0.390	0.397	0.400	0.401	0.599	0.693	0.468
Region FE	Yes	Yes	Yes	Yes	Yes	No	No	Yes
State FE	No	No	No	No	No	Yes	No	No
Municipality FE	No	No	No	No	No	No	Yes	No

Notes: * significant at 10% level; ** significant at 5% level; *** significant at 1% level. Robust standard errors in parenthesis. The omitted categories are males, illiterate, and married or widowed individuals. Weighted estimates.

In addition, I perform two sensitiveness tests. First, migration regions in Mexico cover large territories (see Figure 5.1), which suggests that the migrants' socioeconomic profile may vary considerably within migrant-sending regions. For this reason, I estimate the effect of stage migration narrowing the place of birth fixed effects. Column 6 reports the estimates considering migrants born in the same state rather than in the same region: state-of-birth fixed effects. The wage premium for stage migration reduces to 3% as well as its statistical significance (10% level). This confirms the presence of heterogeneity in productive skills within regions, and therefore the stage migration effect controlling for birthplace at the region level may be overestimated. Column 7 reports the estimates considering municipality-of-birth fixed effects, which help to isolate the stage migration effect for migrants from the same local context (rural or urban). To some extent, this specification guarantees like with like comparisons. Although the size of the stage migration coefficient is unaffected, it loses statistical significance perhaps

due to the low number of observations used in estimations at the municipality level.

Second, the estimated wage premium associated to stage migration may be rooted in selection issues. For example, migrants from locations far away from the border—traveling long distances—may be more likely to migrate in stages, but at the same time, they may be positively selected relative to those individuals from locations closer to the border—who migrate in one move. To compare stage and direct migrants from places at similar distance from the border, I expand Equation 5.1 by adding as independent variable the minimum distance (km) from the migrant's municipality of birth to the Mexico-US border.[4] I also include its quadratic form to control for diminishing returns to distance and region-of-birth fixed effects.

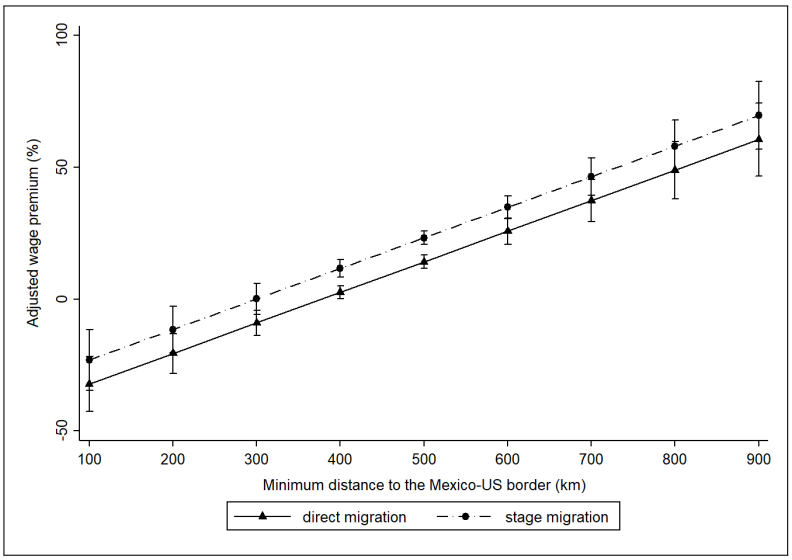
Column 8 shows a positive correlation between distance and the migrant's potential wage at the destination. A 10 km increase in distance is associated with a 1% increase in the potential wage. Figure 5.5 depicts the adjusted wage premium for stage and direct migration considering different bands of distance to the border. Since the standard errors overlap at the extremes, migrants from locations relatively close (less than 300 km)/far (more than 800 km) to the border may have obtained similar wages at the United States regardless of whether they migrated in stages or directly. However, the average migrant was born 473 km away from the border, and Figure 5.5 shows that for bands of distance between 300 and 700 km there is a statistically significant difference at means: stage migration is associated with a 9% wage premium once controlling for distance (see column 8 of Table 5.5).

The results that I have presented should be interpreted with caution since they are only correlations. My estimates may suffer from omitted variable bias

[4]The distance estimates were kindly shared by Woodruff & Zenteno (2007).

(OVB).[5] Additionally, the data do not allow me to identify if stage migrants become literate at their birthplace or at their first-stage destination, and thus may be an additional source of endogeneity. Yet, the results suggest that the method of migration may be relevant to explain the migrants' labor-market performance. A suitable instrumental variable strategy could uncover the causal effect of stage migration on the migrant's potential wage at the destination.

Figure 5.5: Influence of distance on wages. Adjusted wage premium



Source: Mexican Border Crossing Records–Microfilm publication N° A3365. Note: 95% confidence intervals. The estimates correspond to first-time migrants over 12 years of age. The adjusted values were estimated considering region-of-birth fixed effects.

5.6 Conclusions

In this Chapter, I estimate correlations between stage migration and the migrant's potential wage at the final destination—that is, the premium for stage migration. I use novel micro data reporting birthplace, last residence and final destination of first-time Mexican migrants that crossed the US border from 1906 to 1908. At the beginning of the twentieth century, 36% of the working-age migrants moved within Mexico before crossing the border. The results show that migrating in stages is associated with a 3 to 9% increase in the potential wage at the destination.

[5] OVB occurs when a variable that is correlated with both the dependent and one or more included independent variables is omitted from a regression equation.

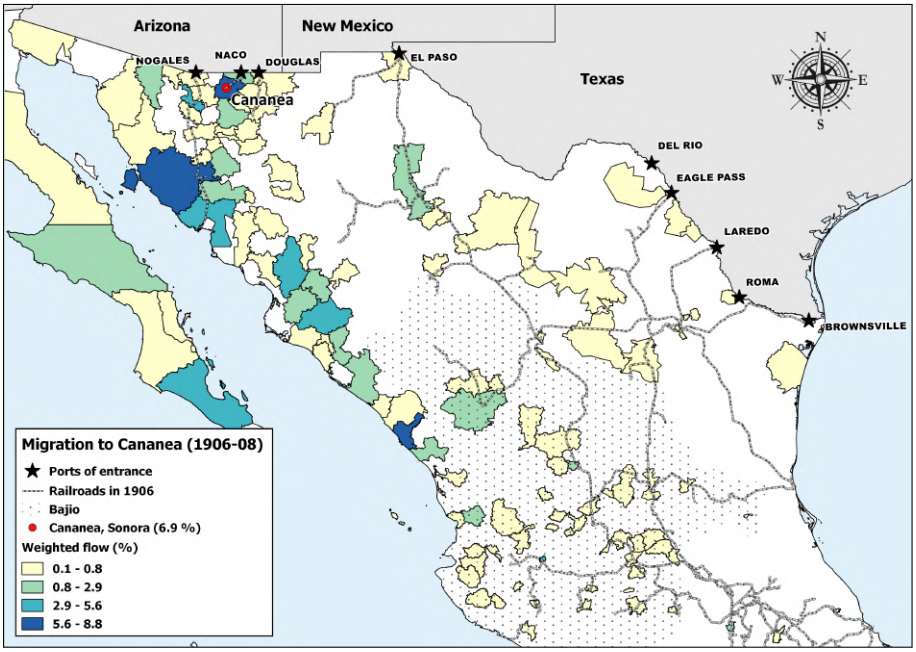
This magnitude is equivalent to the wage premium for English fluency during the same period (Ward, 2019b, p. 3). Hence, the benefits of migrating in stages (access to migrant networks and information) may compensate the lack of fluency in the destination's language. I acknowledge that in addition to sex, age, literacy and marital status, there are unobservable factors that may influence both the decision to migrate in stages and the potential wage at the destination.

To my knowledge, the migration method has not been explored as a factor explaining differences in migrant labor-market performance at the destination. Stage migration can influence the accumulation of productive skills relevant to the labor market of final destination. However, stage migration may provide access to migration networks assisting with relevant information and support to migrate as well. This could be especially true considering that the strength and quality of migration networks may vary across regions of migrant-sending countries.

An interesting next step could be estimating the causal effect of stage migration on earnings convergence within a generation and/or across generations. Although the findings represent a level effect on wages upon arrival, this initial wage premium could be enough to make a difference in assimilation trajectories over time. The main take away of the Chapter is that, as today, during the Age of Mass Migration not everyone could migrate internationally in one single move. Those that migrated in steps due to income or geographic constraints could have benefited from this experience: it may have improved their labor-market performance at the final destination.

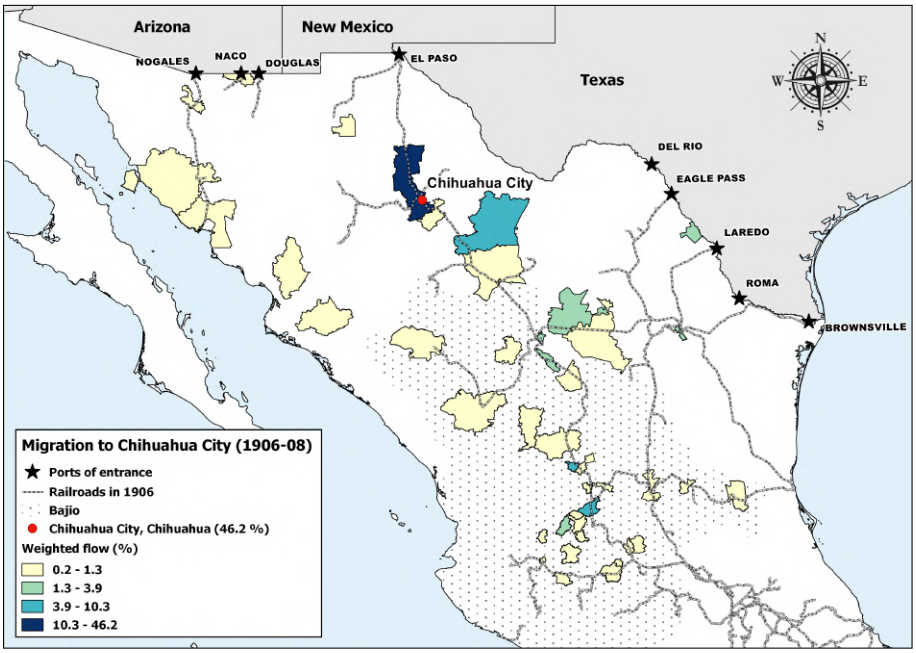
5.7 Appendix

Figure 5.6: Migration to Cananea (1906–08)



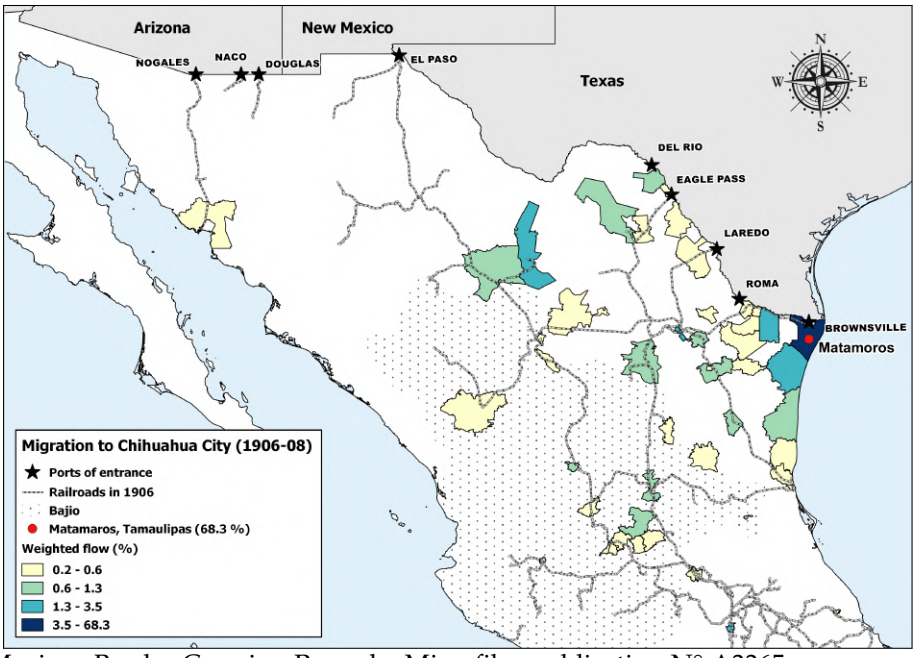
Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Spatial distribution of the Mexican-American migration flow from July 1906 to December 1908. The polygons display the immigrant's birth place (municipalities) and their shares in the overall weighted flow to Cananea, Sonora (quartiles calculated with Jenks natural breaks classification method). The shaded area covers the states of the Bajio region.

Figure 5.7: Migration to Chihuahua (1906–08)



Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Spatial distribution of the Mexican-American migration flow from July 1906 to December 1908. The polygons display the immigrant's birth place (municipalities) and their shares in the overall weighted flow to Chihuahua City, Chihuahua (quartiles calculated with Jenks natural breaks classification method). The shaded area covers the states of the *Bajio* region.

Figure 5.8: Migration to Matamoros (1906–08)



Source: Mexican Border Crossing Records. Microfilm publication N° A3365.
Note: Spatial distribution of the Mexican-American migration flow from July 1906 to December 1908. The polygons display the immigrant's birth place (municipalities) and their shares in the overall weighted flow to Matamoros, Tamaulipas (quartiles calculated with Jenks natural breaks classification method). The shaded area covers the states of the *Bajio* region.

Chapter 6

Conclusion

The previous chapters study Mexican migration to the United States by the end of the Age of Mass Migration. They intend to be the first cliometric research addressing the characteristics of the Mexican-American migration flow in its beginnings (1884–1910). They are supported by novel micro data that allows me to provide a quantitative assessment of different dimensions of the initial mechanics of Mexican migration. The following paragraphs highlight the main contributions of the thesis and outline general lessons and future research directions.

The first contribution of the thesis is the transcription of new micro data to study Mexican migration in the early twentieth century (1906–08). With these data—the Mexican Border Crossing Records publication N° A3365—it is possible to characterize the full profile and migration experience of each immigrant. After introducing and analyzing these new data, I conclude in Chapter 2 that previous historical research might have described inaccurately the initial migration patterns. This is because the data used by previous literature are not decomposable at the local level and do not cover long periods of time. The main take away from Chapter 2 is that at the onset of the Mexican Revolution (1910–20), most Mexican immigrants came from northern Mexico and migration rates presented substantial variation within and across states. In other words, Mexican migration followed

local dynamics, and this could have been the case of other migration streams during the Age of Mass Migration.

The Mexican Border Crossing Records represent a unique source of data that can be used to continue expanding our knowledge on Mexican migration in th early twentieth century. However, they capture migration flows from other source countries such as China, England, Greece, Italy or Japan. It would be interesting to know, for example, if Italian immigrants entering to the United States via Mexico were different than those registered at Ellis Island (New York), and if so, in which dimensions and why.

The second contribution of the thesis is to show that differences in wage rates between Mexico and the United States were not the main factor behind the emigration decision, and that the forces that drove Mexican emigration varied between migrant-sending regions. In Chapter 3, I provide a quantitative assessment of the push and pull factors that may explain differences in size across local-level migration corridors. I also present the spatial distribution and intensity of these migration corridors, which uncover a clear heterogeneity in migration costs and suggests the presence of different mechanisms to overcome constraints to migration.

A natural step to follow is to exploit these differences in migration intensity to study the direct or indirect effect of migration at the local level in Mexico and/or the United States. For example, it is possible to evaluate if municipalities with high migration rates—at the beginning of the migration flow—followed different development trajectories than municipalities with low migration rates (see Droller (2018) for an example of this literature). Moreover, it may be possible to assess if differences in the county of destination explain heterogeneous effects (Hatton, 2010, p. 948).

The third contribution of the thesis is to show that—on the basis of height—the first Mexican migrants were neutrally or positively selected relative to the average

laborer. In Chapter 4, I quantify the selectivity of Mexican migration to the United States before 1910. Following a similar approach, Kosack & Ward (2014) show that just after the Mexican Revolution, Mexican migrants became more positively self-selected relative to my results. Hence, it would be very interesting to assess if the Mexican Revolution changed migrant selection and through which mechanisms.

The fourth contribution of the thesis is to show that in the absence of legal immigration restrictions, random shocks such as financial crises and natural disasters have the potential to modify selection patterns in the short-run. In Chapter 4, I provide evidence suggesting that institutions can adjust selection patterns to unanticipated changes in the demand of migrant labor. A promising research area is to continue studying how institutions influence migrant selection (Abramitzky & Boustan, 2017, p. 1325), but also address the long-run effects—on migrant selection and other dimensions of migration—derived from the persistence of institutions involved in the immigration process.

Finally, the fifth contribution of the thesis is to show that stage migration may improve the immigrant's labor-market performance. In Chapter 5, I estimate the wage premium for stage migration. The implications of migrating in stages during the Age of Mass Migration are relatively unknown. I believe that this topic invites economists and economic historians to think beyond human capital accumulation and explore eclectically the mechanisms through which sequential migration can influence the assimilation of immigrants into the labor market in the short and long run.

Bibliography

Abramitzky, R., Boustan, L. P., , & Eriksson, K. (2019c). To the new world and back again: Return migrants in the Age of Mass Migration. *ILR Review*, *72*(2), 300–322.

Abramitzky, R. & Boustan, L. P. (2017). Immigration in American Economic History. *Journal of Economic Literature*, *55*(4), 1311–45.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, *102*(5), 1832–56.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2013). Have the poor always been less likely to migrate? Evidence from inheritance practices during the age of mass migration. *Journal of Development Economics*, *102*, 2–14.

Abramitzky, R., Boustan, L. P., & 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.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2016). Cultural assimilation during the Age of Mass Migration. *National Bureau of Economic Research Working Paper 22381*.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2019d). Do immigrants assimilate more slowly today than in the past? *American Economic Review-Insights (forthcoming)*.

Abramitzky, R., Boustan, L. P., Eriksson, K., Feigenbaum, J. J., & Pérez, S. (2019b). Automated linking of historical data. *National Bureau of Economic Research Working Paper 25825*.

Abramitzky, R., Mill, R., & Pérez, S. (2019a). Linking individuals across historical sources: A fully automated approach. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 1–18.

Ambrosini, J. W. & Peri, G. (2012). The determinants and the selection of Mexico–US migrants. *The World Economy*, *35*(2), 111–151.

Andrew, A. P. (1908). Substitutes for Cash in The Panic of 1907. *The Quarterly Journal of Economics*, *22*(4), 497–516.

Angelucci, M. (2012). Conditional Cash Transfer Programs, Credit Constraints, and Migration. *Labour*, *26*(1), 124–136.

Angelucci, M. (2015). Migration and financial constraints: Evidence from Mexico. *Review of Economics and Statistics*, *97*(1), 224–228.

Angrist, J. D. & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Antecol, H., Cobb-Clark, D. A., & Trejo, S. J. (2003). Immigration policy and the skills of immigrants to Australia, Canada, and the United States. *Journal of Human Resources*, *38*(1), 192–218.

Arnaut, J. L. (2018). Mexican real wages before the revolution: A reappraisal. *Iberoamericana–Nordic Journal of Latin American and Caribbean Studies*, *47*(1).

Balke, N. S. & Gordon, R. J. (1989). The Estimation of Prewar Gross National Product: Methodology and New Evidence. *Journal of Political Economy*, *97*(1), 38–92.

Banerjee, A., Duflo, E., & Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in China. (17897).

Barde, R., Carter, S. B., & Sutch, R. (2006). Table Ad 162-172: Immigrants, by country of last residence-North America: 1820-1997. In S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, & G. Wright (Eds.), *Historical statistics of the United States: Earliest times to the present*. New York: Cambridge University Press.

Basu, B. & Bang, J. T. (2013). Insurance and remittances: New evidence from Latin American immigrants to the US. *Migration Letters*, *10*(3), 383.

Bellettini, G. & Ceroni, C. B. (2007). Immigration Policy, Self-selection, and the Quality of Immigrants. *Review of International Economics*, *15*(5), 869–877.

Belot, M. V. & Hatton, T. J. (2012). Immigrant Selection in the OECD. *The Scandinavian Journal of Economics*, *114*(4), 1105–1128.

Bertazzini, M. (2018). The Long-Term Impact of Italian Colonial Roads in the Horn of Africa, 1935-2000 . *LSE-Economic History Working Papers*, *2018*(272), 1–57.

Bianchi, M. (2013). Immigration Policy and Self-Selecting Migrants. *Journal of Public Economic Theory*, *15*(1), 1–23.

Biavaschi, C., Giulietti, C., & Siddique, Z. (2017). The Economic Payoff of Name Americanization. *Journal of Labor Economics*, *35*(4), 1089–1116.

Bignon, V., Esteves, R., & Herranz-Loncán, A. (2015). Big push or big grab? Railways, government activism, and export growth in Latin America, 1865–1913. *The Economic History Review*, *68*(4), 1277–1305.

Bleakley, H. & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics*, *86*(2), 481–496.

Bleakley, H. & Chin, A. (2010). Age at arrival, English proficiency, and social assimilation among US immigrants. *American Economic Journal: Applied Economics*, *2*(1), 165–92.

Bleakley, H., Costa, D., & Lleras-Muney, A. (2014). Health, Education, and Income in the United States, 1820–2000. In *Human Capital in History: The American Record* (pp. 121–159). University of Chicago Press.

Blum, M. & Rei, C. (2017). Escaping Europe: health and human capital of Holocaust refugees. *European Review of Economic History*, *22*(1), 1–27.

Bodenhorn, H., Guinnane, T. W., & Mroz, T. A. (2017). Sample-selection biases and the industrialization puzzle. *The Journal of Economic History*, *77*(1), 171–207.

Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review*, *77(4)*, 531–553.

Borjas, G. J. (2007). *Mexican immigration to the United States*. University of Chicago Press.

Borrescio-Higa, F., Bozzoli, C. G., & Droller, F. (2019). Early life environment and adult height: The case of chile. *Economics & Human Biology*, *33*, 134–143.

Boustan, L. P. (2007). : (pp. 267–90). Cambridge, MA: MIT Press.

Brass, T. (1990). The Latin American enganche system: Some revisionist reinterpretations revisited. *Slavery and Abolition*, *11*(1), 74–103.

Brown, H. S., Wilson, K. J., & Angel, J. L. (2015). Mexican immigrant health: health insurance coverage implications. *Journal of Health Care for the Poor and Underserved*, *26*(3), 990–1004.

Bruner, R. & Carr, S. (2007). Lessons from the Financial Crisis of 1907. *The Journal of Applied Corporate Finance*, *19*(4), 115–124.

Bureau of the Census (1900). *12th Population Census of the United States*. Washington: Government Printing Office.

Bureau of the Census (1910). *13th Population Census of the United States*. Washington: Government Printing Office.

Campos-Vazquez, R. M., Flores, C. D., & Márquez, G. (2017). *Long-Run Human Development in Mexico: 1895–2010*, (pp. 89–112). Cham: Springer International Publishing.

Campos-Vázquez, R. M. & Vélez-Grajales, R. (2012). Did Population well-being improve during Porfirian Mexico? A regional analysis using a quasi-human development index. *Journal of Human Development and Capabilities*, *13*(4), 597–620.

Caponi, V. (2011). Intergenerational transmission of abilities and self-selection of Mexican immigrants. *International Economic Review*, *52*(2), 523–547.

Cardoso, L. A. (1980). *Mexican emigration to the United States, 1897-1931*. Tucson: University of Arizona Press.

Carneiro, P. M., Lee, S., & Reis, H. (2016). Please call me John: name choice and the assimilation of immigrants in the United States, 1900-1930.

Case, A. & Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy*, *116*(3), 499–532.

Castro Aranda, H. (2010). *Primer censo de población de la Nueva España 1790. Censo de Revillagigedo "Un Censo Condenado"*. Mexico City: Sociedad Mexicana de Geografía y Estadística; Instituto Nacional de Estadística y Geografía.

Chacón, S. (2009). Política migratoria: proceso negociador 1947-1954. *Foro Internacional*, 518–558.

Chalfin, A. (2015). The long-run effect of mexican immigration on crime in us cities: Evidence from variation in mexican fertility rates. *American Economic Review*, *105*(5), 220–25.

Chiquiar, D. & 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.

Chiswick, B. (1999). Are immigrants favorably self-selected? *American Economic Review*, *89*(2), 181–185.

Chiswick, B. R. & Miller, P. W. (2010). Occupational language requirements and the value of English in the US labor market. *Journal of Population Economics*, *23*(1), 353–372.

Chort, I. & De La Rupelle, M. (2016). Determinants of Mexico-US outward and return migration flows: A state-level panel data analysis. *Demography*, *53*(5), 1453–1476.

Clark, V. (1908). Mexican Labor in the US. *Bulletin of the United States Bureau of Labor*, (78), 466–522.

Clemens, M. A., Lewis, E. G., & Postel, H. M. (2018). Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion. *American Economic Review*, *108*(6), 1468–87.

Coatsworth, J. H. (1979). Indispensable railroads in a backward economy: The case of Mexico. *The Journal of Economic History*, *39*(4), 939–960.

Cohn, R. L. (1995). Occupational evidence on the causes of immigration to the United States, 1836-1853. *Explorations in Economic History*, *32*(3), 383–408.

Collins, W. J. & Margo, R. A. (2000). Residential segregation and socioeconomic outcomes: When did ghettos go bad? *Economics Letters*, *69*(2), 239–243.

Collins, W. J. & Zimran, A. (2018). The economic assimilation of irish famine migrants to the united states.

Connor, D. S. (2019). The cream of the crop? Geography, networks, and Irish migrant selection in the age of mass migration. *The Journal of Economic History*, *79*(1), 139–175.

Contreras, C. (2005). The droughts in Mexico during the XIX Century. *Investigaciones Geográficas*, (56), 118–133.

Coon, M. (2014). Financial development and the end-use of migrants' remittances. *IZA Journal of Labor & Development*, *3*(1), 7.

Coon, M. (2015). "Walking ATMs": Do Crime Rates Affect Remittances of Mexican Migrants in the United States? *Eastern Economic Journal*, *41*(1), 6–23.

Cosío Villegas, D. & Bernal, I. (1973). *Historia mínima de México*. Mexico City: El Colégio de México.

Cosío Villegas, D. (1974). *Historia moderna de México*. Mexico City: Hermes.

Covarrubias, M., Lafortune, J., & Tessada, J. (2015). Who comes and why? determinants of immigrants skill level in the early xxth century us. *Journal of Demographic Economics*, *81*(1), 115–155.

Crozet, M. (2004). Do migrants follow market potentials? An estimation of a new economic geography model. *Journal of Economic Geography*, *4*(4), 439–458.

Cutler, D. M., Glaeser, E. L., & Vigdor, J. L. (2008). Is the melting pot still hot? Explaining the resurgence of immigrant segregation. *The Review of Economics and Statistics*, *90*(3), 478–497.

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, *27*(2), 281–314.

De Cardona, S. A. (1892). *De México a Chicago y Nueva York: guía para el viajero en la que se describen las principales ciudades y ferrocarriles*. New York: Moss Engraving Co.

Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences*, *104*(33), 13232–13237.

Dell, M. (2012). Path dependence in development: Evidence from the Mexican Revolution. *Unpublished manuscript*.

Donato, K. M. (1993). Current trends and patterns of female migration: Evidence from Mexico. *International Migration Review*, 748–771.

Droller, F. (2018). Migration, population composition and long run economic development: Evidence from settlements in the pampas. *The Economic Journal*, *128*(614), 2321–2352.

Durand, J. (2007). *Braceros: las miradas mexicana y estadounidense: antología (1945-1964)*. Mexico City: Miguel Angel Porrúa.

Durand, J. (2016). *Historia Mínima de la Migración México-Estados Unidos*. Mexico City: El Colégio de México.

Durand, J. & Arias, P. (2000). *La experiencia migrante: iconografía de la migración México-Estados Unidos*. Ucol.

Edin, P.-A., Fredriksson, P., & Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment. *The Quarterly Journal of Economics*, *118*(1), 329–357.

El Economista Mexicano (1907). Emigración de Braceros Mexicanos, 6–6.

Eriksson, K. (2018). Ethnic Enclaves and Immigrant Outcomes: Norwegian Immigrants during the Age of Mass Migration. *National Bureau of Economic Research Working Paper 24763*.

Eriksson, K. & Ward, Z. (2018). The Ethnic Segregation of Immigrants in the United States from 1850 to 1940. *National Bureau of Economic Research Working Paper 24764*.

Feliciano, Z. M. (2001). The skill and economic performance of Mexican immigrants from 1910 to 1990. *Explorations in Economic History*, *38*(3), 386–409.

Fernandez-Huertas, J. (2011). New Evidence on Emigrant Selection. *The Review of Economics and Statistics*, *93*(1), 72–96.

Fernandez-Huertas Moranga, J. (2011). New Evidence on Emigrant Selection. *The Review of Economics and Statistics*, *93*(1), 72–96.

Flint, D. (1900). Hope Culture in California. *Farmers Bulletin*, (115).

Flores, M., Zey, M., & Hoque, N. (2013). Economic liberalization and contemporary determinants of Mexico's internal migration: an application of spatial gravity models. *Spatial Economic Analysis*, *8*(2), 195–214.

Fogel, W. A. (1978). *Mexican illegal alien workers in the United States*. Monograph series (University of California, Los Angeles. Institute of Industrial Relations) ; 20. Los Angeles: Institute of Industrial Relations of University of California.

Frydman, C., Hilt, E., & Zhou, L. Y. (2015). Economic effects of runs on early "shadow banks": Trust companies and the impact of the panic of 1907. *Journal of Political Economy*, *123*(4), 902–940.

Gamio, M. (1930). *Mexican immigration to the United States*. Chicago: University of Chicago Press.

Gamio, M. (2002). *El inmigrante mexicano. La historia de su vida. Entrevistas completas, 1926-1927*. The Regents of The University of California and Centro de Investigaciones y Estudios Superiores de Antropología Social.

García y Griego, M. (1983). The importation of Mexican Contract Laborers to the United States, 1942–1964: Antecedents, Operation and Legacy. In P. G. Brown & H. Shue (Eds.), *The border that joins: Mexican migrants and US responsibility*. Rowman & Littlefield Pub Inc.

Garcia, A. S. & Schmalzbauer, L. (2017). Placing assimilation theory: Mexican immigrants in urban and rural America. *The ANNALS of the American Academy of Political and Social Science*, *672*(1), 64–82.

Giulietti, C., Wahba, J., & Zenou, Y. (2018). Strong versus weak ties in migration. *European Economic Review*, *104*, 111–137.

Goldin, C. (1994). The political economy of immigration restriction in the United States, 1890 to 1921. In *The regulated economy: A historical approach to political economy* (pp. 223–258). University of Chicago Press.

Goldstein, J. R. & Stecklov, G. (2016). From Patrick to John F: Ethnic Names and Occupational Success in the Last Era of Mass Migration. *American Sociological Review*, *81*(1), 85–106.

Gómez, M. (2011). Había una vez un sistema de bancos privados emisores de billetes. México, 1897-1910. *Historia Mexicana*, 2069–2109.

González, G. G. & Fernandez, R. (2002). Empire and the origins of twentieth-century migration from Mexico to the United States. *Pacific Historical Review*, *71*(1), 19–57.

González, M. (2010). Los braceros en el Porfiriato. *Estudios Agrarios*, (44), 9–26.

Gould, J. D. (1979). European inter-continental emigration 1815-1914: patterns and causes. *Journal of European Economic History*, *8*(3), 593–679.

Gratton, B. & Merchant, E. (2015). An Immigrant's Tale: The Mexican American Southwest 1850 to 1950. *Social Sciences History*, (39), 521–50.

Gray, R., Narciso, G., & Tortorici, G. (2019). Globalization, agricultural markets and mass migration: Italy, 1881-1912. *Explorations in Economic History (forthcoming)*.

Greenwood, M. J. (1997). Internal migration in developed countries. *Handbook of Population and Family economics*, *1*, 647–720.

Greenwood, M. J. & Ward, Z. (2015). Immigration quotas, World War I, and emigrant flows from the United States in the early 20th century. *Explorations in Economic History*, *55*, 76–96.

Gutmann, M. P., McCaa, R., Gutiérrez-Montes, R., Gratton, B., & Orensanz, L. (2000). Los efectos demográficos de la Revolución Mexicana en Estados Unidos. *Historia Mexicana*, 145–165.

Hansen, B. A. (2014). A failure of regulation? Reinterpreting the panic of 1907. *Business History Review*, *88*(3), 545–569.

Hanson, G. H. & Spilimbergo, A. (1999). Illegal immigration, border enforcement, and relative wages: Evidence from apprehensions at the US-Mexico border. *American Economic Review*, *89*(5), 1337–1357.

Hardy, O. (1934). The revolution and the Railroads of Mexico. *Pacific Historical Review*, *3*(3), 249–269.

Hatton, T. & Ward, Z. (2018). International Migration in the Atlantic Economy 1850-1940. *The Australian National University–Center for Economic History Discussion Paper 24764*.

Hatton, T. J. (1995a). A model of UK emigration, 1870-1913. *The Review of Economics and Statistics*, 407–415.

Hatton, T. J. (1995b). A model of scandinavian emigration, 1870–1913. *European Economic Review*, *39*(3-4), 557–564.

Hatton, T. J. (1997). The Immigrant Assimilation Puzzle in Late Nineteenth-Centuty America. *The journal of economic history*, *57*(1), 34–62.

Hatton, T. J. (2010). The Cliometrics of International Migration: A Survey. *Journal of Economic Surveys*, *24*(5), 941–969.

Hatton, T. J. (2014). The economics of international migration: A short history of the debate. *Labour Economics*, *30*, 43–50.

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

Hatton, T. J. & Williamson, J. G. (1994). What Drove the Mass Migrations from Europe in the Late Nineteenth Century? *Population and Development Review*, *20*(3), 533–559.

Hatton, T. J. & Williamson, J. G. (1998). *The Age of Mass Migration: Causes and Economic Impact*. New York: Oxford University Press.

Henderson, T. J. (2011). *Beyond borders: A history of Mexican migration to the United States*. John Wiley & Sons.

Herranz-Loncán, A. (2014). Transport technology and economic expansion: The growth contribution of railways in Latin America before 1914. *Revista de Historia Económica-Journal of Iberian and Latin American Economic History*, *32*(1), 13–45.

Hudson, E. W. (1914). Growing Egyptian Cotton in Salt River Valley Arizona. *Farmers Bulletin*, (577).

Humphries, J. & Leunig, T. (2009). Was Dick Whittington taller than those he left behind? Anthropometric measures, migration and the quality of life in early nineteenth century London? *Explorations in Economic History*, *46*(1), 120–131.

Ibarraran, P. & Lubotsky, D. (2007). Mexican immigration and self-selection: New evidence from the 2000 Mexican census. In *Mexican immigration to the United States* (pp. 159–192).

Imbens, G. & Angrist, J. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2).

Instituto Nacional de Estadística Geografía e Informática (1986). *Estadísticas Históricas de México*. Mexico City.

Inwood, K., Minns, C., & Summerfield, F. (2019). Occupational income scores and immigrant assimilation. Evidence from the Canadian census. *Explorations in Economic History*.

Jedwab, R. & Moradi, A. (2016). The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa. *Review of Economics and Statistics*, *98*(2), 268–284.

Jenkins, J. C. (1978). The demand for immigrant workers: Labor scarcity or social control? *International Migration Review*, *12*(4), 514–535.

Johnson, J. F. (1908). The crisis and panic of 1907. *Political Science Quarterly*, *23*(3), 454–467.

Juif, D. & Quiroga, G. (2019). Do you have to be tall and educated to be a migrant? Evidence from Spanish recruitment records, 1890–1950. *Economics & Human Biology*, *34*, 115–124.

Kaestner, R. & Malamud, O. (2014). Self-selection and international migration: New evidence from Mexico. *Review of Economics and Statistics*, *96*(1), 78–91.

Knight, A. (1986). *The Mexican Revolution*. Cambridge Latin American Studies. Vol. 54–55. Cambridge: Cambridge University Press.

Komlos, J. & Baten, J. (2004). Looking Backward and Looking Forward: Anthropometric Research and the Development of Social Science History. *Social Science History*, *28*(2), 191–210.

Komlos, J. & Meermann, L. (2007). The introduction of anthropometrics into development and economics. *Historical Social Research/Historische Sozialforschung*, 260–270.

Kosack, E. (2016). The Bracero Program and Entrepreneurial Investment in Mexico. *Unpublished manuscript*.

Kosack, E. (2019). Guest Worker Programs and Human Capital Investment: The Bracero Program in Mexico, 1942–1964. *Journal of Human Resources (forthcoming)*.

Kosack, E. & Ward, Z. (2014). Who crossed the border? Self-selection of Mexican migrants in the early twentieth century. *The Journal of Economic History*, *74*(4), 1015–1044.

Kosack, E. & Ward, Z. (2019). El sueño americano? The generational progress of Mexican Americans prior to World War II. *Unpublished manuscript*.

Kuntz, S. (1982). *Un Siglo de Devaluaciones del Peso Mexicano*. Mexico City: Siglo XXI.

Kuntz, S. (1995). *Empresa extranjera y mercado interno: El Ferrocarril Central Mexicano, 1880-1907*. Mexico City: El Colégio de México.

Larsen, U. M. (1982). A quantitative study of emigration from Denmark to the United States, 1870–1913. *Scandinavian Economic History Review*, *30*(2), 101–128.

Lee, J., Peri, G., & Yasenov, V. (2017). The Employment Effects of Mexican Repatriations: Evidence from the 1930's. *National Bureau of Economic Research Working Paper 23885*.

Lindert, P. H. & Williamson, J. G. (2016). *Unequal gains: American growth and inequality since 1700*. Princeton University Press.

Lleras-Muney, A. & Shertzer, A. (2015). Did the Americanization movement succeed? An evaluation of the effect of english-only and compulsory schooling laws on immigrants. *American Economic Journal: Economic Policy*, *7*(3), 258–90.

López-Alonso, M. (2007). Growth with inequality: Living standards in Mexico, 1850–1950. *Journal of Latin American Studies*, *39*(1), 81–105.

López-Alonso, M. (2015). *Estar a la altura. Una historia de los niveles de vida en Mexico, 1850-1950*. Mexico City: Fondo de Cultura Económica.

López-Alonso, M. & Condey, R. P. (2003). The ups and downs of Mexican economic growth: the biological standard of living and inequality, 1870–1950. *Economics & Human Biology*, *1*(2), 169–186.

Lozano, F. A. & Sorensen, T. (2015). Mexican Immigrants, Labour Market Assimilation and the Current Population: The Sensitivity of Results Across Seemingly Equivalent Surveys. *International Migration*, *53*(2), 250–262.

Markham, J. W. (2002). The Panic of 1907. In *A financial history of the United States*, volume Volume 2: From JP Morgan to the Institutional Investor:(1900-1970) (pp. 118–133). Sharpe.

Massey, C. G. (2016). Immigration quotas and immigrant selection. *Explorations in Economic History*, *60*, 21–40.

Massey, D. S. (1987). Understanding mexican migration to the United States. *American Journal of Sociology*, *92*(6), 1372–1403.

Massey, D. S. & Basem, L. C. (1992). Determinants of savings, remittances, and spending patterns among US migrants in four Mexican communities. *Sociological Inquiry*, *62*(2), 185–207.

Massey, D. S. & Espinosa, K. E. (1997). What's driving Mexico-US migration? A theoretical, empirical, and policy analysis. *American Journal of Sociology*, *102*(4), 939–999.

Massey, D. S. & Liang, Z. (1989). The long-term consequences of a temporary worker program: The US Bracero experience. *Population Research and Policy Review*, *8*(3), 199–226.

Massey, D. S. & Pren, K. A. (2012). Unintended consequences of US immigration policy: Explaining the post-1965 surge from Latin America. *Population and Development Review*, *38*(1), 1–29.

Mayda, A. M., Ortega, F., Peri, G., Shih, K., & Sparber, C. (2018). The effect of the H-1B quota on the employment and selection of foreign-born labor. *European Economic Review*, *108*, 105–128.

Mayet, E. F., y Cervera, J. S., & Arias, D. P. G. (1980). Las sequías en México: historia, características y efectos. *Comercio Exterior*, *30*(7), 756.

McKenzie, D. & Rapoport, H. (2007). Network effects and the dynamics of migration and inequality: theory and evidence from Mexico. *Journal of development Economics*, *84*(1), 1–24.

McKenzie, D. & Rapoport, H. (2010). Self-selection patterns in Mexico-US migration: the role of migration networks. *The Review of Economics and Statistics*, *92*(4), 811–821.

Minns, C. (2000). Income, cohort effects, and occupational mobility: a new look at immigration to the United States at the turn of the 20th century. *Explorations in Economic History*, *37*(4), 326–350.

Moatsos, M., Baten, J., Foldvari, P., van Leeuwen, B., & van Zanden, J. L. (2014). Income inequality since 1820. In *How was life? Global well-being since 1820* (pp. 291–316). OECD.

Moen, J. & Tallman, E. W. (1992). The bank panic of 1907: the role of trust companies. *The Journal of Economic History*, *52*(3), 611–630.

Mokyr, J. & Ó Gráda, C. (1982). Emigration and poverty in prefamine Ireland. *Explorations in Economic History*, *19*(4), 360–384.

Moser, P. (2012). Taste-based discrimination evidence from a shift in ethnic preferences after WWI. *Explorations in Economic History*, *49*(2), 167–188.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *The Quarterly Journal of Economics*, *118*(2), 549–599.

NARA (2000). *Publication A3365: Lists of Aliens Arriving at Brownsville, Del Rio, Eagle Pass, El Paso, Laredo, Presidio, Rio Grande City and Roma, Texas, May 1903-June 1909, and at Aros Ranch, Douglas, Lochiel, Naco and Nogales, Arizona, July 1906-December 1910*. Washington DC: National Archives and Records Administration (NARA).

Nawrotzki, R. J. & DeWaard, J. (2016). Climate shocks and the timing of migration from Mexico. *Population and Environment*, *38*(1), 72–100.

Nawrotzki, R. J., Riosmena, F., Hunter, L. M., & Runfola, D. M. (2015). Undocumented Migration in Response to Climate Change. *International Journal of Population Studies*, *1*(1), 60.

Oñate, A. (1991). La política agraria del estado mexicano durante el Porfiriato. In A. Hernández & M. Miño (Eds.), *Cincuenta Años de Historia de México* (pp. 293–314). El Colégio de México.

Odell, K. A. & Weidenmier, M. D. (2004). Real shock, monetary aftershock: The 1906 San Francisco earthquake and the panic of 1907. *The Journal of Economic History*, *64*(4), 1002–1027.

Ogórek, B. (2019). Talented but lazy. The height-school premium among Cracow's schoolboys in the interwar period. *Economics & Human Biology*, *34*, 252–256.

O'Connor, A., Batalova, J., & Bolter, J. (2019). Central American Immigrants in the United States. Technical report, Migration Policy Institute.

Orraca-Romano, P. P. & García-Meneses, E. (2016). Why are the wages of the mexican immigrants and their descendants so low in the united states? *Estudios Económicos*, *31*(2 (62)), 305–337.

Orrenius, P. M. & Zavodny, M. (2005). Self-selection among undocumented immigrants from Mexico. *Journal of Development Economics*, *78*(1), 215–240.

Ortmeyer, D. L. & Quinn, M. A. (2012). Coyotes, migration duration, and remittances. *The Journal of Developing Areas*, 185–203.

Ortmeyer, D. L. & Quinn, M. A. (2015). The impact on health of recurring migrations to the United States. *The Journal of Developing Areas*, *49*(3), 49–64.

Passel, J. & Cohn, D. (2019). Mexicans decline to less than half the U.S. unauthorized immigrant population for the first time. Technical report, Pew Research Center.

Passel, J., Cohn, D., & Gonzalez-Barrera, A. (2012). Net Migration from Mexico Falls to Zero—and Perhaps Less. Technical report, Pew Research Center.

Pérez-Armendáriz, C. & Crow, D. (2010). Do migrants remit democracy? International migration, political beliefs, and behavior in Mexico. *Comparative Political Studies*, *43*(1), 119–148.

Perlmann, J. (2005). *Italians Then, Mexicans Now: Immigrant Origins and the Second-Generation Progress, 1890-2000*. Russell Sage Foundation.

Poot, J., Alimi, O., Cameron, M., & Maré, D. (2016). The Gravity Model of Migration: The successful comeback of an ageing superstar in regional science. *IZA Discussion Paper Series*, (10329), 1–24.

Pérez, S. (2017). The (South) American Dream: Mobility and Economic Outcomes of First- and Second-Generation Immigrants in Nineteenth-Century Argentina. *77*(4), 971–1006.

Quigley, J. M. (1972). An economic model of Swedish emigration. *The Quarterly Journal of Economics*, 111–126.

Quinn, M. A. (2005). Remittances, savings, and relative rates of return. *The Journal of Developing Areas*, 1–23.

Ramos, R. (2016). Gravity models: a tool for migration analysis. *IZA World of Labor*, (239).

Reimers, C. (1983). Labor Market Discrimination Against Hispanic and Black Men. *The Review of Economics and Statistics*, *65*(4), 570.

Rhode, P. W. (2002). Gallman's annual output series for the United States, 1834-1909. *National Bureau of Economic Research*.

Roberts, K. D. & Morris, M. D. S. (2003). Fortune, Risk, and Remittances: An Application of Option Theory to Participation in Village-Based Migration Networks. *International Migration Review*, *37*(4), 1252–1281.

Romer, C. D. (1989). The prewar business cycle reconsidered: New estimates of gross national product, 1869-1908. *Journal of Political Economy*, *97*(1), 1–37.

Ríos-Bustamante, A. (1981). *Mexican Immigrant Workers in the US*. Los Angeles, California: University of California.

Rosenzweig, F. (1965). El Desarrollo Económico de México de 1877 a 1911. *El Trimestre Económico*, *32*(127(3)), 405–454.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, *3*(2), 135–146.

Samora, J. (1982). *Los mojados: The wetback story*. Notre Dame: University of Notre Dame Press.

Sánchez-Alonso, B. (2000a). European emigration in the late nineteenth century: the paradoxical case of Spain. *Economic History Review*, 309–330.

Sánchez-Alonso, B. (2000b). Those who left and those who stayed behind: explaining emigration from the regions of spain, 1880–1914. *The Journal of Economic History*, *60*(3), 730–755.

Sánchez-Alonso, B. (2019). The age of mass migration in Latin America. *The Economic History Review*, *72*(1), 3–31.

Schluter, C. & Wahba, J. (2009). Illegal migration, wages, and remittances: Semi-parametric estimation of illegality effects.

Schneider, E. B. & Ogasawara, K. (2018). Disease and child growth in industrialising Japan: critical windows and the growth pattern, 1917-39. *Explorations in Economic History*, *69*, 64–80.

Schultz, T. P. (2002). Wage gains associated with height as a form of health human capital. *American Economic Review*, *92*(2), 349–353.

Secretaría de Agricultura y Fomento (1910). *Tercer Censo de Población de los Estados Unidos Mexicanos*. Mexico City: Oficina Impresora de Secretaría de Hacienda.

Secretaría de Comunicaciones y Obras Públicas (1906). *Carta de Ferrocarriles de los Estados Unidos Mexicanos*. Mexico City.

Secretaría de Economía (1956). *Estadísticas Sociales del Porfiriato 1877-1911*. Mexico City.

Sequeira, S., Nunn, N., & Qian, N. (2019). Immigrants and the Making of America. *Review of Economic Studies*.

Spitzer, Y. & Zimran, A. (2018). Migrant self-selection: Anthropometric evidence from the mass migration of Italians to the United States, 1907–1925. *Journal of Development Economics*, *134*, 226–247.

Stahle, D. W., Cook, E. R., Burnette, D. J., Villanueva, J., Cerano, J., Burns, J. N., Griffin, D., Cook, B. I., Acuña, R., & Torbenson, M. (2016). The Mexican Drought Atlas: Tree-ring reconstructions of the soil moisture balance during the late pre-Hispanic, colonial, and modern eras. *Quaternary Science Reviews*, *149*, 34–60.

Steckel, R. H. (1995). Stature and the Standard of Living. *Journal of Economic Literature*, *33*(4), 1903.

Takenaka, A. & Pren, K. A. (2010). Determinants of emigration: Comparing migrants' selectivity from Peru and Mexico. *The Annals of the American Academy of Political and Social Science*, *630*(1), 178–193.

Taylor, P. S. (1930–1933). *Mexican labor in the United States*. Vol. I-III. Berkeley, California: University of California Press.

The Washington Post (1907). Stocks thrown over. *Aug 13*, 3.

Timmer, A. S. & Williamson, J. G. (1998). Immigration policy prior to the 1930s: Labor markets, policy interactions, and globalization backlash. *Population and Development Review*, *24*(4), 739–742.

Ullmann, S. H. (2012). The health impacts of international migration on Mexican women. *Global Public Health*, *7*(9), 946–960.

US Congress (1903). Immigration Act, 32 Stat. 1213, 57th Congress.

US Congress (1907). Immigration Act, 34 Stat. 898, 59th Congress.

U.S. Department of Commerce and Labor (1906a). *Census of Manufactures: 1905. Arizona, Indian Territory, New Mexico and Oklahoma (Bulletin No. 30)*. Washington, DC: Government Printing Office.

U.S. Department of Commerce and Labor (1906b). *Census of Manufactures: 1905. Louisiana, Mississippi, and Texas (Bulletin No. 48)*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Commerce and Labor (1907). *Bulletin of the Bureau of Labor (No. 71)*. Washington, DC: U.S. Government Printing Office.

US Department of Labor (1965). Contracts issued to Mexican workers. In T. N. Schroth (Ed.), *Congress and the Nation 1945-1964: A Review of Government and Politics in the Postwar Years* (pp. 1343). Washington DC: Congressional Quarterly Service.

US Immigration Commission (1911a). *Abstracts of Reports of the Immigration Commission: With Conclusions and Recommendations and Views of the Minority (In Two Volumes)*, volume 2. Washington DC: US Government Printing Office.

US Immigration Commission (1911b). *Reports of the Immigration Commission. Immigrants in industries (in twenty-five parts)*. Washington, DC: U.S. Government Printing Office.

Vargas, A. J. (2016). Assimilation effects beyond the labor market: time allocations of Mexican immigrants to the US. *Review of Economics of the Household*, *14*(3), 625–668.

Verduzco, G. (1995). La migración mexicana a Estados Unidos: recuento de un proceso histórico. *Estudios Sociológicos*, 573–94.

Verduzco, G. (2000). La migración mexicana a Estados Unidos: estructuración de una selectividad histórica. In R. Tuiran (Ed.), *Migración México-Estados Unidos: Continuidad y Cambios* (pp. 11–32). CONAPO.

Ward, Z. (2017). Birds of passage: Return migration, self-selection and immigration quotas. *Explorations in Economic History*, *64*, 37–52.

Ward, Z. (2019a). The Not-So-Hot Melting Pot: The Persistence of Outcomes for Descendants of the Age of Mass Migration. *American Economic Journal: Applied Economics (forthcoming)*.

Ward, Z. (2019b). The low return to English fluency during the Age of Mass Migration. *European Review of Economic History*.

Wegge, S. A. (2002). Occupational self-selection of european emigrants: Evidence from nineteenth-century hesse-cassel. *European Review of Economic History*, *6*(3), 365–394.

Wells, N., Goddard, S., & Hayes, M. J. (2004). A self-calibrating Palmer drought severity index. *Journal of Climate*, *17*(12), 2335–2351.

Williamson, J. G. (1998). Globalization, labor markets and policy backlash in the past. *Journal of Economic Perspectives*, *12*(4), 51–72.

Wilson, F. A., Stimpson, J. P., & Pagán, J. A. (2014). Disparities in health outcomes of return migrants in Mexico. *International Journal of Population Research*, *204*.

Woodruff, C. & Zenteno, R. (2007). Migration networks and microenterprises in Mexico. *Journal of Development Economics*, *82*(2), 509–528.

World Bank (2019). Migration and Remittances. Recent Developments and Outlook. Technical report, World Bank Group.