Initial determinants of Mexican Mass migration

David Escamilla-Guerrero

D.Escamilla-Guerrero@lse.ac.uk

London School of Economics

August 19, 2019

Abstract

Exploiting novel micro data consisting of individual border crossings from 1906 to 1908, this paper 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 market potentials and Mexican immigrant networks in the United States were the main pull factors rather than the US-Mexico wage gap. Differences in living standards within 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.

Keywords: labor migration, railways, institutions, Mexico
JEL Classification Numbers: N36, N76, F22, J61

Acknowledgments: I am especially grateful to my PhD supervisors Eric Schneider and Joan Rosés for their guidance and invaluable comments. I thank Chris Minns, Zachary Ward, Neil Cummins, Ran Abramitzky, Alfonso Cabreros-Zurita, León Fernández-Bujanda and Mattia Bertazzini for their extensive feedback. I am grateful to Daniela Gutiérrez and Marco Villeda for their assistance on the data transcription. I benefited from presenting at the Graduate Seminar in Economic History (London School of Economics), the Oxford Graduate Seminar in Economic and Social History (Nuffield College), and the 2019 Annual Conference of the Economic History Society. This research was developed with the financial support of: the Mexican National Council for Science and Technology (2015-2018) - Scholarship No. 409165; the Mexican Ministry of Education Scholarship (2015-2016); the Radwan Travel and Discovery Fund (2016) - London School of Economics; the Pre-Dissertation Exploratory Grant (2017) - Economic History Association (US); and the Research Fund for Graduate Students (2017) - Economic History Society (UK). Additionally, I benefited from my academic visit at El Colegio de México in 2016. I am grateful to Professor Graciela Márquez and Professor Gerardo Esquivel for their support during this period. Finally, I am especially grateful to Dr. Fernando Pérez for his mentorship during my PhD fellowship at The Bank of Mexico in 2018. All errors are mine.

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, in the beginnings of the flow (1884-1910), most immigrants came from the densely populated central plateau of Mexico (known as the Bajio) (see Figure 1). In addition, previous scholarship has focused on the Mexico-US wage gap and the expansion of Mexico's railways system to explain the emigration of Mexicans during this period (Durand, 2016; Gratton & Merchant, 2015; Henderson, 2011; among others). Thus, the story of Mexican migration was similar to the general mass migration story (Hatton & Williamson, 1993, Hatton & Williamson, 1994; Hatton, 1995; and Williamson, 1998).[1] For the first time, I provide a quantitative assessment of the initial determinants of the Mexico-United States migration 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 10,895 immigrants that crossed the border through nine ports of entrance (see Figure 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

[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.

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 (1910-1920), 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. Escamilla-Guerrero (2019) provides a full description of the MBCRs and the sampling plan I followed.

Figure 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.

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, 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 did not explain the migration flow. However, differences in living standards and population size across Mexican municipalities mattered. This result reveals that, on average, locations with low living standards and 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) found 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 population size in the origin and destination influenced the decision to migrate. In contrast, migration from the Bajio was completely determined by immigrant networks and population pressures in Mexico.

Second, Cardoso (1980), Durand (2016), among 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 paper to the cliometric literature on international migration are the following. 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 paper 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 migration corridors and its main characteristics. The fifth and sixth sections address the empirical strategy and results. The last section concludes.

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 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 8 and Table 9 in the Appendix) and an unfavorable exchange rate.[7] Furthermore, the incentive to emigrate was rooted in the deep impoverishment of the rural population, whose living conditions worsened during the dictatorship of Porfirio Diaz, a period known as the *Porfiriato* (1877–1911) (Cardoso, 1980, p. 11; Henderson, 2011; Oñate, 1991).[8] Thus, the sustained US labor

[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).

[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.

demand; the wage differentials between the countries; and the deteriorated living standards in Mexico are considered 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 operations (Coatsworth, 1979). Similarly, Cardoso (1980); Durand (2016); Gamio (1930) among others argue that railways were fundamental to understand the migration flow, since they reduced importantly the migration cost at the time: distance.

While the economic impact of railways in Mexico was higher than in other countries,[10] the stagnated salary of the working class questions the social savings on passenger transportation associated to this technology (Coatsworth, 1979, p. 960). For this reason, the *enganche*, a search-matching labor institution, was implemented by Mexican and American recruiters to transport and allocate seasonal laborers in the US (Clark, 1908; Durand, 2016; Gamio, 1930). Hence, the *enganche* was a mechanism that eased emigration because recruiters covered the transportation costs in exchange of future labor.

The previous factors operated under a favorable policy environment. Before 1910, Mexicans were not considered immigrants who sought to settle permanently, but temporary aliens 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

[9]Figure 8 depicts the fast expansion of the Mexican railways during this period.

[10]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).

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.

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] migration 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. 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.

3.1 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

[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%.

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 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. Escamilla-Guerrero (2019) 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 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).

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

[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.

[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.

[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.

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).

Table 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 1 for the location of entrance ports.

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 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 (agriculture, manufactures and mining) and type of wage earner (Mexican or average laborer).

Table 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.

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 (US Department of Commerce and Labor, 1906a,b) and statistics of wages in manufacture industries (US Department of Commerce and Labor, 1907).

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 10 in the Appendix presents this data in detail.

These sources provide wage levels in 114 counties across 14 states. Table 3 shows that wages in manufacture activities were relatively similar across states. Texas had

[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).

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.

Table 3: Average wage in the US 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 10.
Note: Estimates of wages at the state-level. Data from wages at US destination counties.

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 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.

Table 4: Twenty most important migration flows from Mexican municipalities to US counties (1906–08)

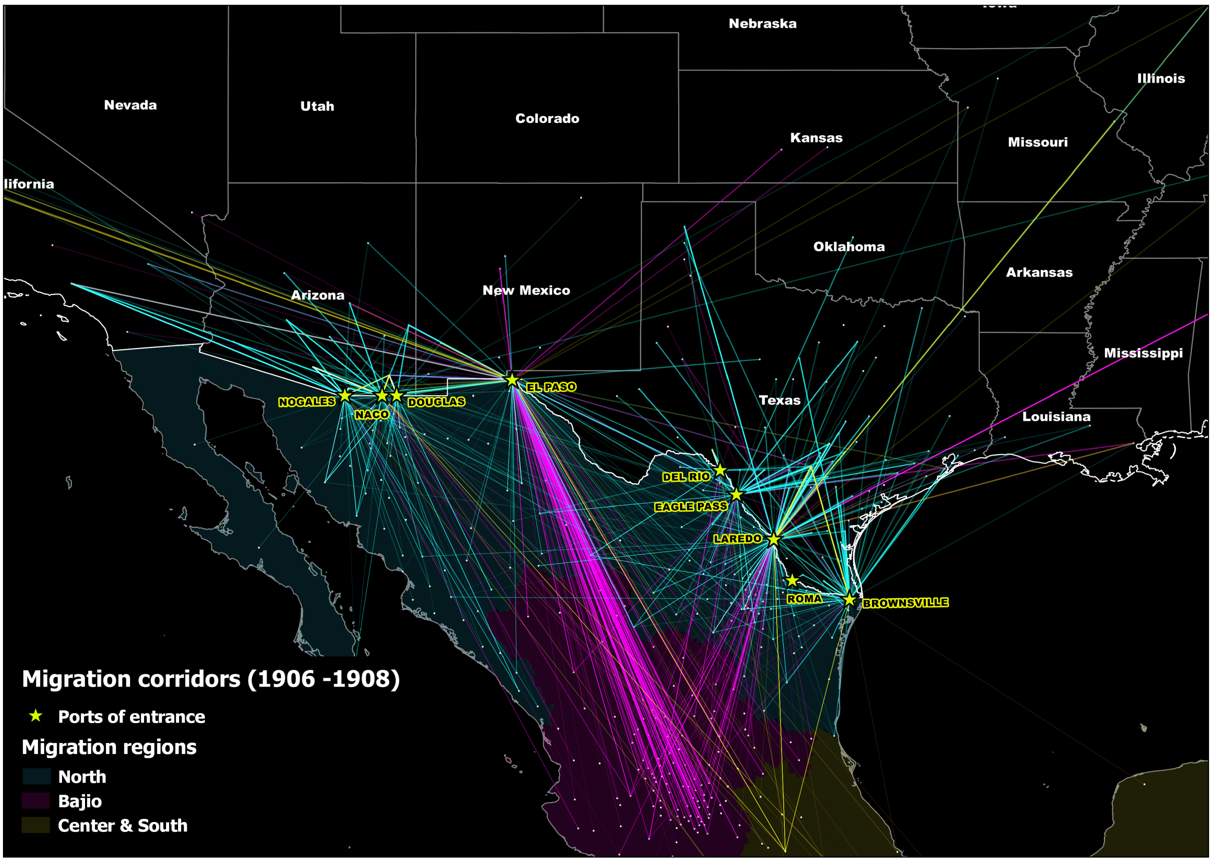
Last Permanent Residence		**Final Destination**			
State	*Municipality*	*State*	*County*	*Weighted Flow*	*Share (%)*
Sonora	Cananea	Arizona	Cochise	1,264	8.3
Nuevo Leon	Monterrey	Texas	Webb	908	6.0
Chihuahua	Chihuahua	Texas	El Paso	451	3.0
Guanajuato*	Penjamo	Texas	El Paso	424	2.8
Nuevo Leon	Monterrey	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.

Furthermore, most migration corridors from the Bajio had El Paso as destination (see Figure 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 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.

Figure 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.

These two migration patterns are depicted in Figure 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 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 2 shows that emigration streams from the Border region ended at 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.

5. Push and pull factors of Mexican emigration

At the turn of the twentieth century, the factors influencing Mexican migration to the United States were a melting pot of development asymmetries, wage differentials, labor institutions, droughts and improvements in the means of transportation. Clark (1908), Cardoso (1980), Durand (2016), Henderson (2011), Gratton & Merchant (2015) among others argue that these factors are "key" to explain the mass migration of Mexicans. However, this literature does not disentangle the specific effect of such factors. I intend to bring light to this matter by answering two questions. First, was the US-Mexico wage gap the main incentive to emigrate? Second, did the access to railroads induce emigration?

5.1 Empirical strategy

To answer these questions, 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. The basic estimating equation is:

$$ln\, M_{ij} = \alpha + \beta_1\, ln\, D_{ij} + \beta_2\, ln\, Pop_i^{MX} + \beta_3\, ln\, Pop_j^{US} + e_{ij}. \tag{1}$$

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 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). The 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 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 1 as follows:

$$\begin{aligned} ln\, M_{ij} = \alpha &+ \beta_1\, ln\, D_{ij} + \beta_2\, ln\, Pop_i^{MX} + \beta_3\, ln\, Pop_j^{US} + \beta_4\, Wage_{ij} + \beta_5\, Cont_i \\ &+ \beta_6\, ln\, Network_j + \beta_7\, Drought_i + \beta_8\, HDI_i + e_{ij}. \end{aligned} \tag{2}$$

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 might 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, e.g. 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 effects (idiosyncratic 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. 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

[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 $t \in \{1906, 1907, 1908\}$.

[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.

(2012, p. 611), who study living standards during the Porfiriato. Table 7 in the Appendix presents summary statistics of the previous variables.

5.2 Results

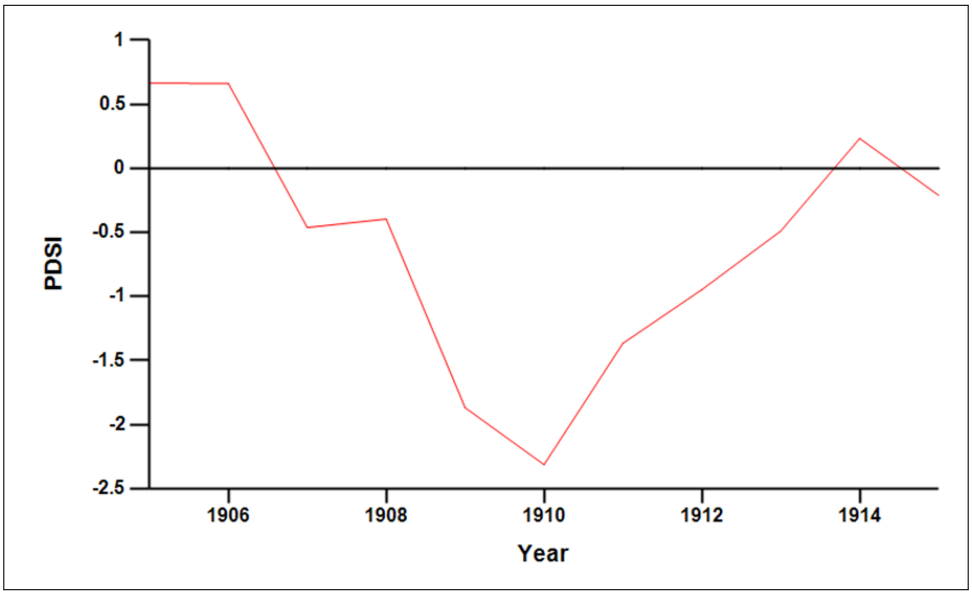
I start the analysis considering the three years as a pooled cross-sectional sample. Column 1 of Table 5 shows that the estimated coefficient β_1 is significant and has the expected sign. An increase of one percent in the distance reduces the migration flow by 0.46%. The population sizes are significant, 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.

Column 5 of Table 5 shows that the relative wage do not explain differences in size across migration corridors. However, the basic push and pull factors (population sizes and distance) remain significant. This result is consistent with Hatton (2010, p. 943) argument: 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.

In contrast, immigrant networks influenced emigration. 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 0.15%. While the marginal effect seems to be small, we need to consider the context of the research. Indeed, 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 US 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 45%. Figure 4 and Figure 5 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.

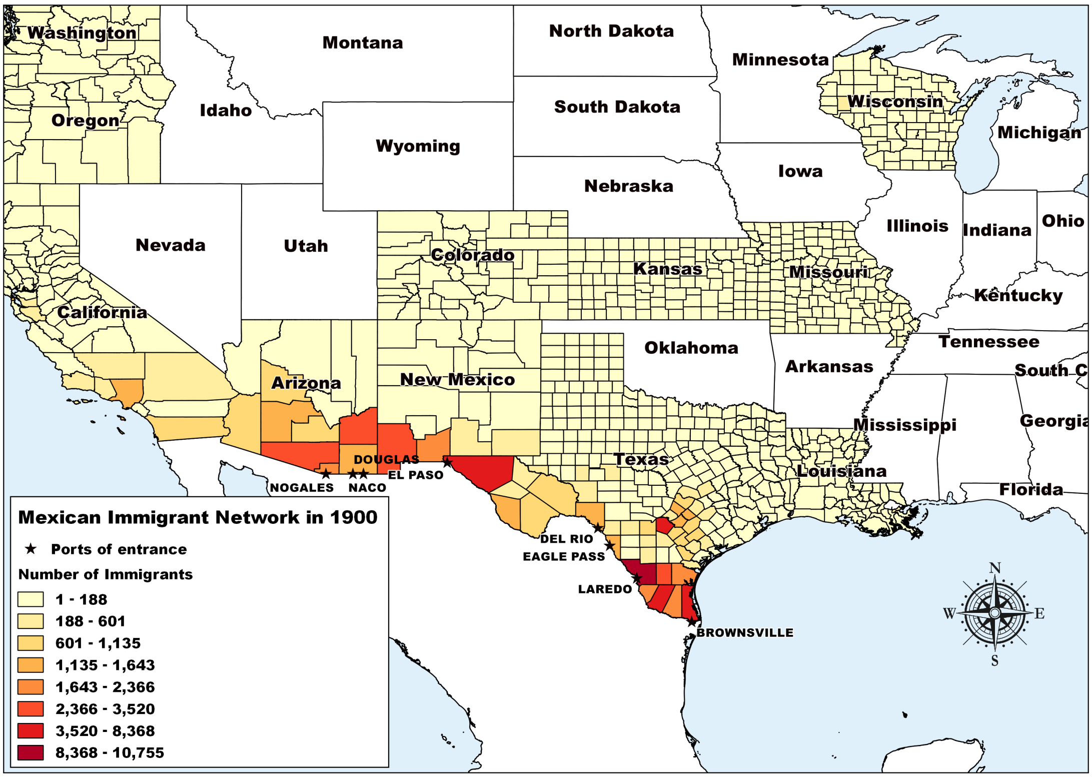
Differences in relative dryness across municipalities had a significant effect on the migration flow. An increase of one index point reduces gross emigration by 11%. Figure 3 shows the average DSI in Mexico during the period under analysis and confirms that the years of 1907 and 1908 were part of a drought that lasted until 1910. Figure 9 in the Appendix shows the regions affected by this phenomena and how it intensified at the eve of the Mexican Revolution (1910–20). The estimated coefficient β_8 reveals that living standards across Mexico—captured by the HDI—mattered. An increase of 0.01 index points decreases emigration by 1.2%. Finally, the condition of being a Mexican municipality on the US border did not influence emigration per se.

Figure 3: Droughts in Mexico. Palmer Drought Severity Index (1905–15)



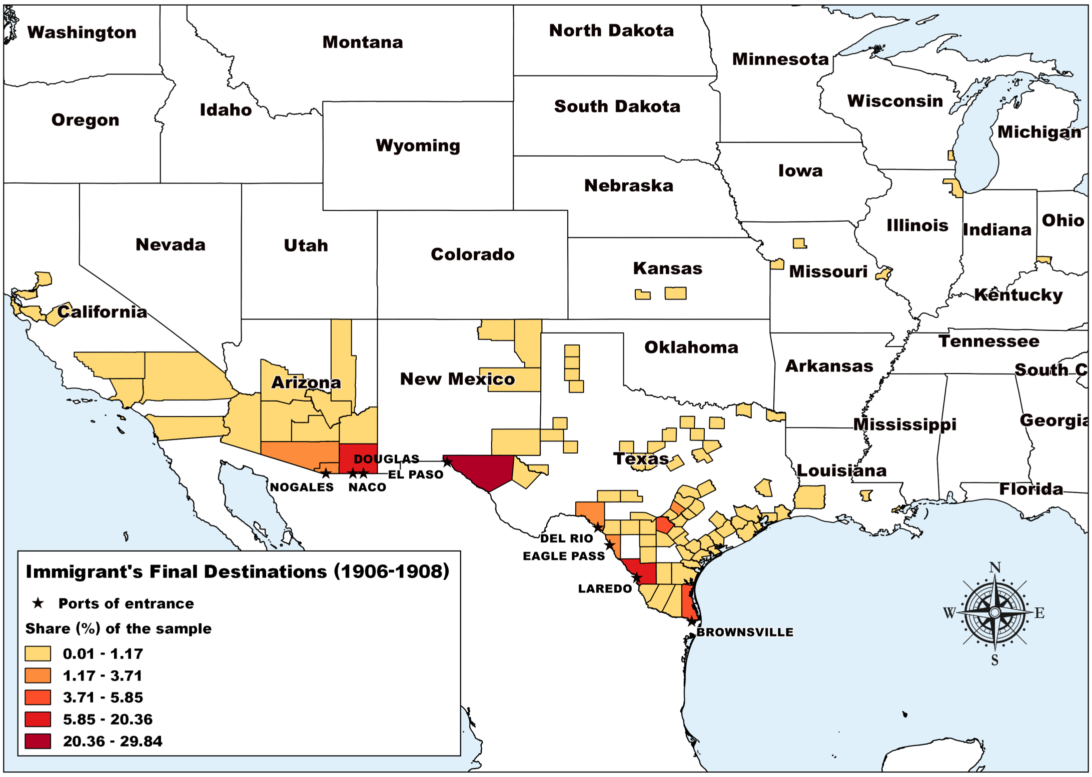
Source: Stahle et al. (2016).
Note: The Palmer Drought Severity Index (PDSI) uses temperature and precipitation data to estimate relative dryness. It is a standardized index that spans from -6 (dry) to +6 (wet). However, values below -4.0 represent extreme droughts while values above +4.0 represent extreme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

Figure 4: Mexican immigrant network in the US (1900)



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

Figure 5: Immigrant's destinations in the US (1906–08)



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

Table 5: Determinants of Mexican migration to the US (1906–08). Dependent variable: Gross emigration (lnM_{ij})

	Complete Sample					Bajio	Border
Independent Variables	1	2	3	4	5	6	7
ln Distance (km)	-0.457***	-0.493***	-0.553***	-0.463***	-0.573***	0.145	-0.670***
	(0.117)	(0.086)	(0.072)	(0.091)	(0.112)	(0.307)	(0.128)
ln MX Population	0.261***	0.264***	0.264***	0.282***	0.322***	0.238***	0.404***
	(0.077)	(0.077)	(0.077)	(0.067)	(0.052)	(0.038)	(0.063)
ln US Population	0.147***	0.136***	0.140***	0.104***	0.111**	-0.002	0.111*
	(0.046)	(0.043)	(0.043)	(0.034)	(0.043)	(0.100)	(0.051)
ln US/MX Wage		0.156	0.166	0.087	0.059	0.264	-0.065
		(0.134)	(0.125)	(0.113)	(0.083)	(0.207)	(0.102)
Contiguity			-0.284	-0.100	0.111		0.050
			(0.184)	(0.176)	(0.112)		(0.099)
ln Migrant Stock 1900				0.172***	0.153***	0.368***	0.111***
				(0.037)	(0.030)	(0.050)	(0.028)
Drought					-0.116*	-0.095	-0.138
					(0.068)	(0.169)	(0.080)
MX Development Index					-1.195***	-0.543	6.678
Constant	0.560	0.585	0.962	-0.586	0.384	-4.993	-1.606
	(0.758)	(0.785)	(0.833)	(0.826)	(0.546)	(2.699)	(0.837)
Observations	892	892	892	866	866	258	570
R-squared	0.073	0.076	0.080	0.123	0.202	0.164	0.247

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.
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 assigned 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.

However, the previous results capture average effects across emigration regions, which in fact followed different migration patterns at the time. Columns 6 and 7 of Table 5 uncover the particular push and pull factors in the Bajio and Border region, respectively. In the Bajio results, distance 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 up in El Paso, i.e. there is little heterogeneity in distance across migration corridors. However, the concentration of immigrants at El Paso is due to the presence of the *enganche*. This labor institution consisted in recruiting laborers in the Bajio and transport them to El Paso, from where they were distributed to farther locations. To some extent, this labor institution constrained or eliminated completely the effect of distance. Similarly, the access to dynamic labor markets in the United States (populated counties) did not explain differences in gross emigration from this region.

What certainly was important for the Bajio migrants were the immigrant networks in the US. Networks might have contributed as well to lower distance costs by providing financial assistance to first-time migrants. This effect is captured by the stock of Mexicans in 1900, which has a positive and significant effect. An increase of one percent in the stock of Mexicans in the US increases the emigration in the Bajio by 0.36%, three times more than in the border states. This result might explain why emigration from the Bajio increased significantly from 1910, overcoming emigration from the Border region.

Column 7 of Table 5 confirms that emigration from the Border region was influenced by different factors. Border immigrants rely less on networks, and more on labor market potentials. The proximity to the border might have allowed them to observe labor market dynamics in the American Southwest and obtain information more promptly to maximize their probability of employment. In addition, this geographical condition might have had two effects. First, the possibility to emigrate facing short distances favored Border immigrants to develop long-term relations with American employers to supply labor every season in the same location.[21] 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 the flow by 0.40%.

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 capture two migration models that might have existed at the time. In the Bajio, emigration was mainly determined by immigrant networks and the *enganche*. This pull

[21] On average, immigrants from the Bajio traveled 1,460 km while their Border region counterparts traveled 658 km only (Table 7).

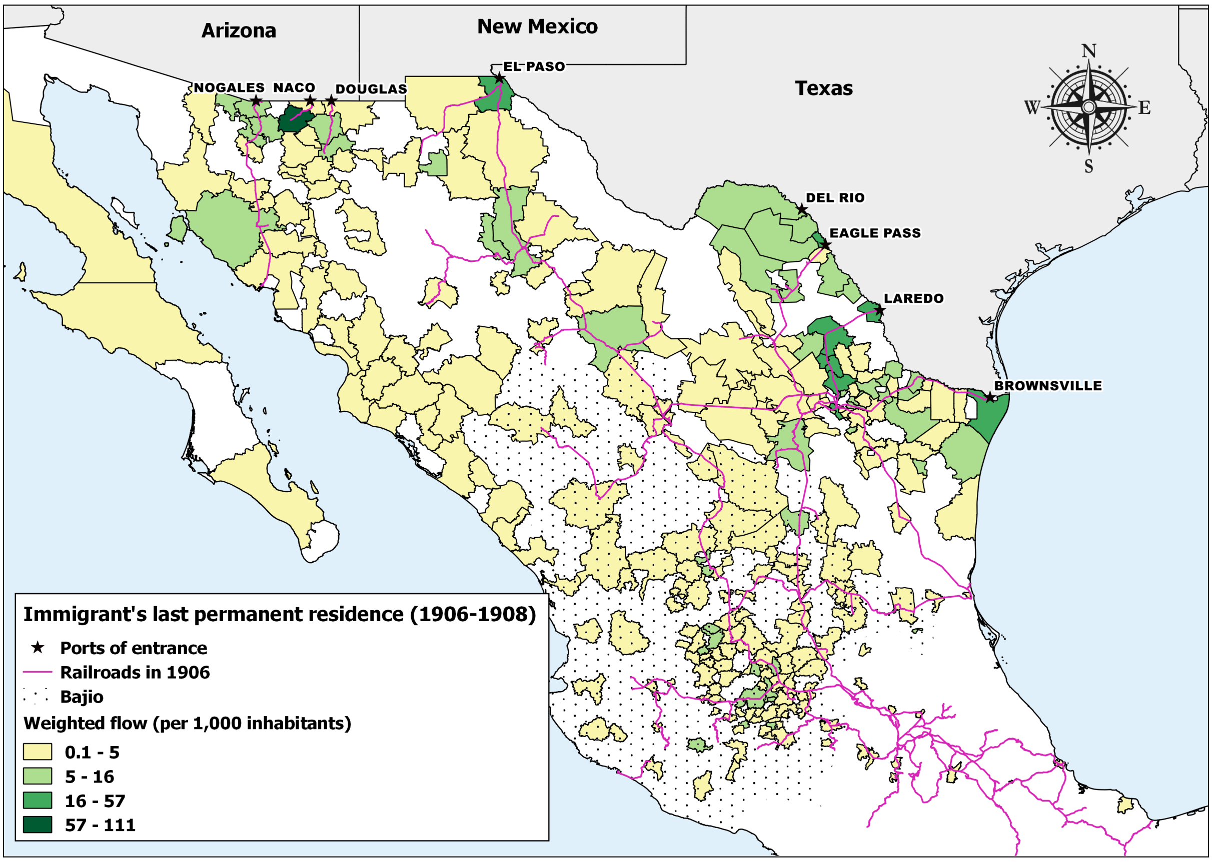
factors might have neutralized the effect of geographic distance and other migration costs. Therefore, emigration from this region might have not driven by the neoclassical cost-benefit decision, but by the social capital formation at the destination and labor institutions. In contrast, in the Border region, the emigration decision was influenced by the distance to the destination. This sensitivity implies that Border migrants might have counted with the means to finance emigration and thus, the rely less on networks or institutions. Also, the probability of employment at home and abroad determined emigration from the Border region, suggesting that migrants had access to information on the dynamics of the local American labor market.

Undoubtedly, living standards and climate shocks might have been important push factors, however they lose explanatory power in the regional regressions because heterogeneity in living standards and temperature within regions is limited. As a robustness check, I estimated Equation 2 for each year of the sample. Table 11 in the Appendix shows that the main results hold. The next section addresses an additional factor that may have shaped Mexican emigration: the access to railways.

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 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.

Figure 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.

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).

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.[22] 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 7 shows that the instrument (straight lines) coincide relatively well with the railways network except in the regions of 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.

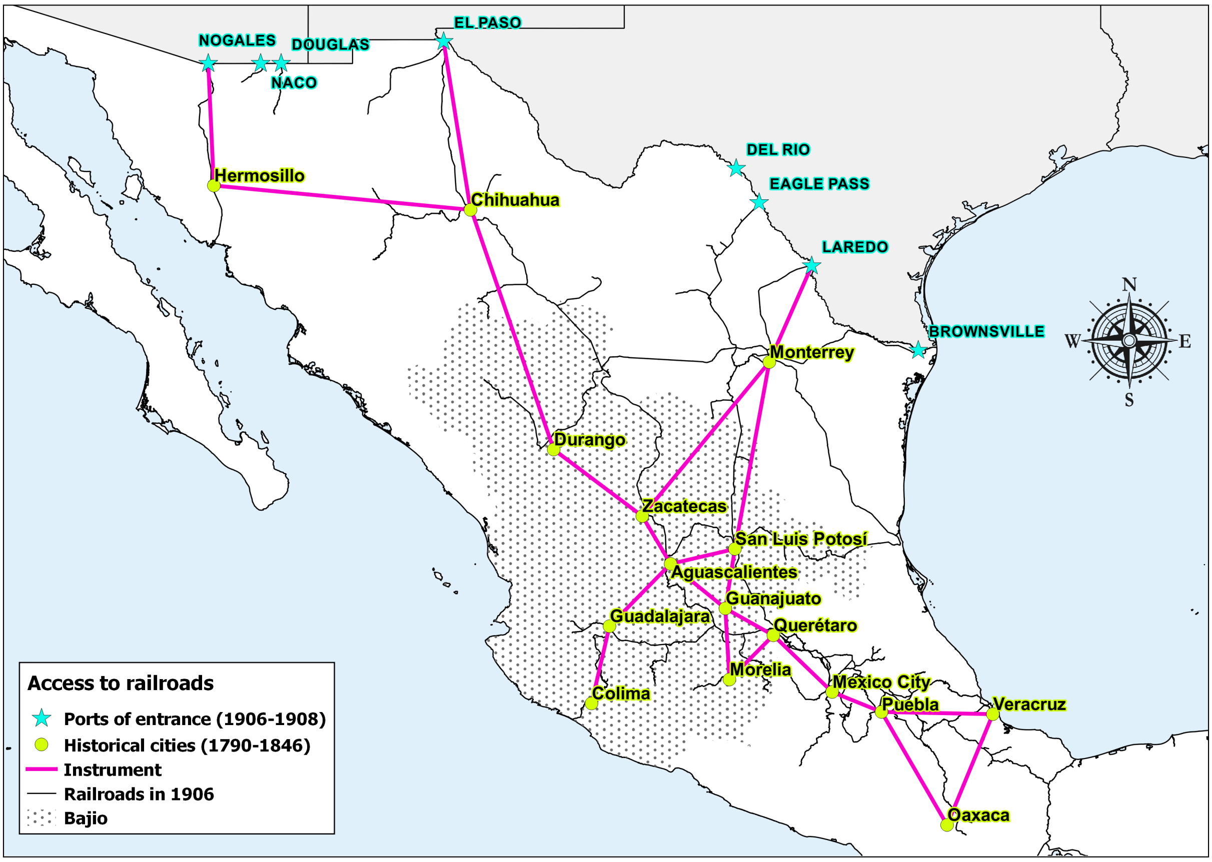
The idea behind using straight lines is that they abstract for transportation corridors that have been present since colonial times. Therefore, being on or near a straight line between two historical cities makes it much more likely that a transportation route will be present, making more feasible emigration 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.

$$ln\, S_{1900,i} = \rho\, ln\, S_{hist,i} + e_i. \tag{3}$$

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).[23] Equation 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

[22] I chose the 60 km criteria because it is the minimum distance needed to connect Puebla and Queretaro with Mexico City.
[23] Figure 11 depicts the map used by Woodruff & Zenteno (2007) in their estimates.

Figure 7: Historical cities (1790–1846) and straight lines as an instrument



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.

cities, e.g. roads. To the author's knowledge, detailed data about roads before 1910 is nonexistent.[24]

Adding $S_{hist,i}$ as explanatory variable to Equation 2 results in the equation:

$$ln\, M_{ij} = \alpha + \Phi\, ln\, S_{hist,i} + \mathbf{W}'_{ij}\beta + \mathbf{X}'_i\delta + \mathbf{Y}'_j\eta + e_{ij}. \tag{4}$$

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.

[24]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).

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).

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 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 7 in the Appendix).

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

	60 km criteria			100 km criteria		
Variables	**Complete sample** 1	**Bajio** 2	**Border** 3	**Complete Sample** 4	**Bajio** 5	**Border** 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 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 10 in the Appendix). 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.

7. Conclusions

Exploiting a novel dataset consisting of individual border crossings from 1906 to 1908, this paper refines 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 time. The results suggest that differences in market potentials and differences in immigrant networks at the United States were the main pull factors rather than differences in wage rates between countries. Consistent with previous literature, differences in living standards across Mexico were the main incentive to emigrate. The regional droughts experienced at the eve of the Mexican Revolution (1910) pushed Mexican peasants to emigrate as well. 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 population 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.

References

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

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

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).

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.

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.

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.

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

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.

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.

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

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.

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

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

Escamilla-Guerrero, D. (2019). Revisiting Mexican migration in the Age of Mass Migration. Evidence from individual border crossings. *Unpublished manuscript (London School of Economics)*.

Feliciano, Z. M. (2001). The skill and economic performance of Mexican immigrants from 1910 to 1990. *Explorations in Economic History*, *38*(3), 386–409.

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. (1978). Mexican illegal alien workers in the United States. *Monograph Series*, (20).

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.

González, M. (2010). Los braceros en el Porfiriato. *Estudios Agrarios*, (44), 9–26.

Gratton, B. & Merchant, E. (2015). An Immigrant's Tale: The Mexican American Southwest 1850 to 1950. *Social Sciences History*, (39), 521–50.

Greenwood, M. J. (1997). Internal migration in developed countries. *Handbook of Population and Family economics*, *1*, 647–720.

Hatton, T. J. (1995). A model of UK emigration, 1870-1913. *The Review of Economics and Statistics*, 407–415.

Hatton, T. J. (2010). The cliometrics of international migration: a survey. *Journal of Economic Surveys*, *24*(5), 941–969.

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).

Instituto Nacional de Estadística Geografía e Informática (1986). *Estadísticas Históricas de México*. Mexico City.

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.

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.

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.

López-Alonso, M. (2007). Growth with inequality: Living standards in Mexico, 1850–1950. *Journal of Latin American Studies*, *39*(1), 81–105.

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, 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.

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.

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.

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.

Ramos, R. (2016). Gravity models: a tool for migration analysis. *IZA World of Labor*, (239).

Rhode, P. W. (2002). Gallman's annual output series for the United States, 1834-1909. Technical report, National Bureau of Economic Research.

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.

Rosenzweig, F. (1965). El Desarrollo Económico de México de 1877 a 1911. *El Trimestre Económico*, *32*(127(3)), 405–454.

Samora, J. (1982). *Los mojados: The wetback story*. Notre Dame: University of Notre Dame Press.

Secretaría de Agricultura y Fomento (1910). *Tercer Censo de Población de los Estados Unidos Mexicanos*. Mexico City: Oficina Impresora de Secretaría de Hacienda.

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.

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.

US Department of Commerce and Labor (1906a). *Census of Manufactures: 1905. Arizona, Indian Territory, New Mexico and Oklahoma (Bulletin No. 30)*. Washington: Government Printing Office.

US Department of Commerce and Labor (1906b). *Census of Manufactures: 1905. Louisiana, Mississippi, and Texas (Bulletin No. 48)*. Washington: Government Printing Office.

US Department of Commerce and Labor (1907). *Bulletin of the Bureau of Labor (No. 71)*. Washington: Government Printing Office.

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.

Williamson, J. G. (1998). Globalization, labor markets and policy backlash in the past. *Journal of Economic Perspectives*, *12*(4), 51–72.

Woodruff, C. & Zenteno, R. (2007). Migration networks and microenterprises in Mexico. *Journal of Development Economics*, *82*(2), 509–528.

Appendix

A. Summary statistics

Table 7: Determinants of Mexican migration to the US (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

B. Historical context

Figure 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 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 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 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 2 for each year of the sample. Table 11 shows that results in Table 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 11).[25]

Column 2 and 5 of Table 11 shows that in 1907, municipalities sharing border with the United States had 54 percent more emigration than inland municipalities.[26] 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).[27] 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 $Contiguity$.

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

[25] 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.

[26] $exp(0.432) = 1.54$

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

Table 11: Determinants of Mexican migration to the US 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

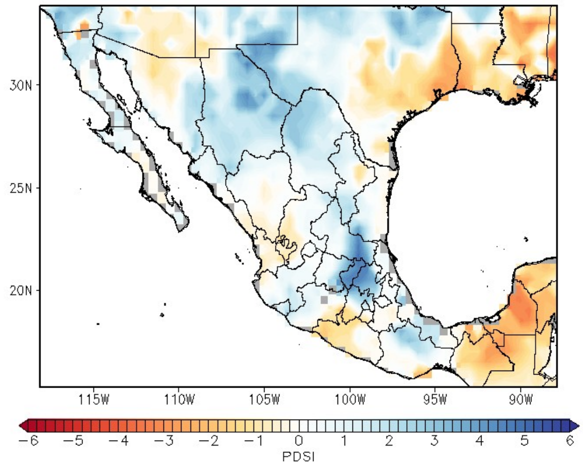
Source: Mexican Border Crossing Records. Microfilm publication number A3365.
Notes: * = Significant at 10% level; ** = Significant at 5% level; *** = Significant at 1% level.
Robust standard errors in parenthesis clustered at a state-level.
lnM_{ij} = weighted migration flow.
Contiguity = Dummy variable 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.

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 10).[28] Finally, the six-month results are very similar in significance and magnitude to complete sample results.

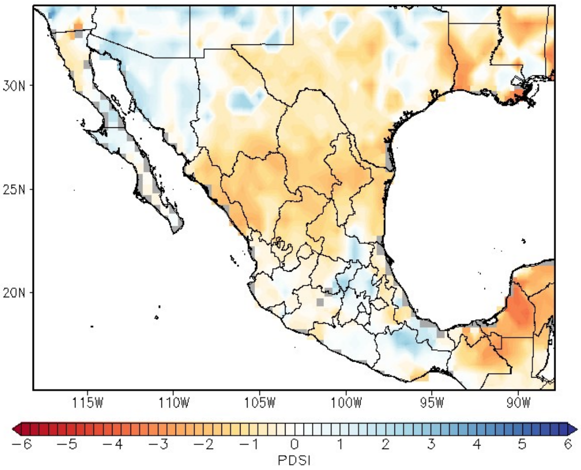
[28]The average expected wage differential in Arizona was 1.82 dollars per day, while in Texas was about 1 dollar per day.

E. Droughts

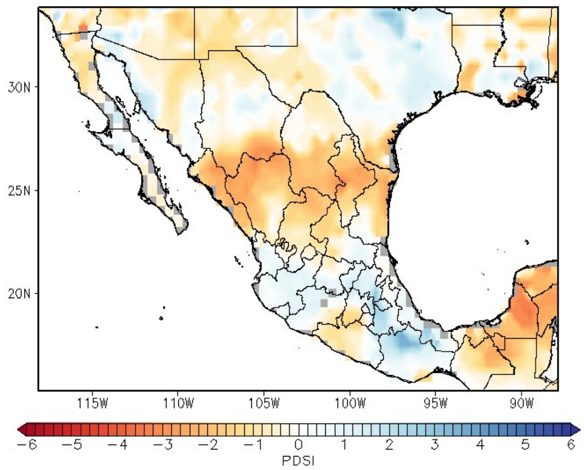
Figure 9: Droughts in Mexico 1906-1909.
Palmer Drought Severity Index (PDSI)



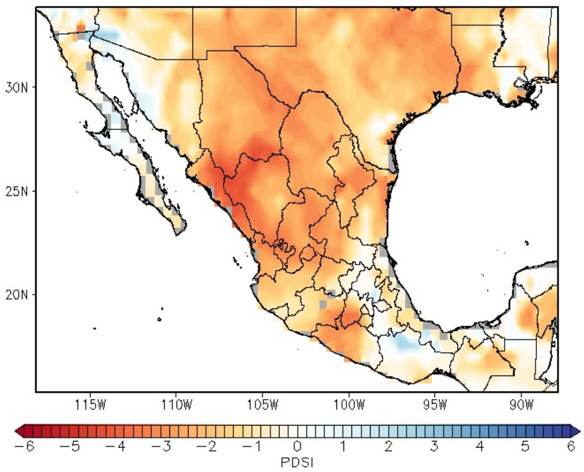
(a) Jan - Dec 1906



(b) Jan - Dec 1907



(c) Jan - Dec 1908

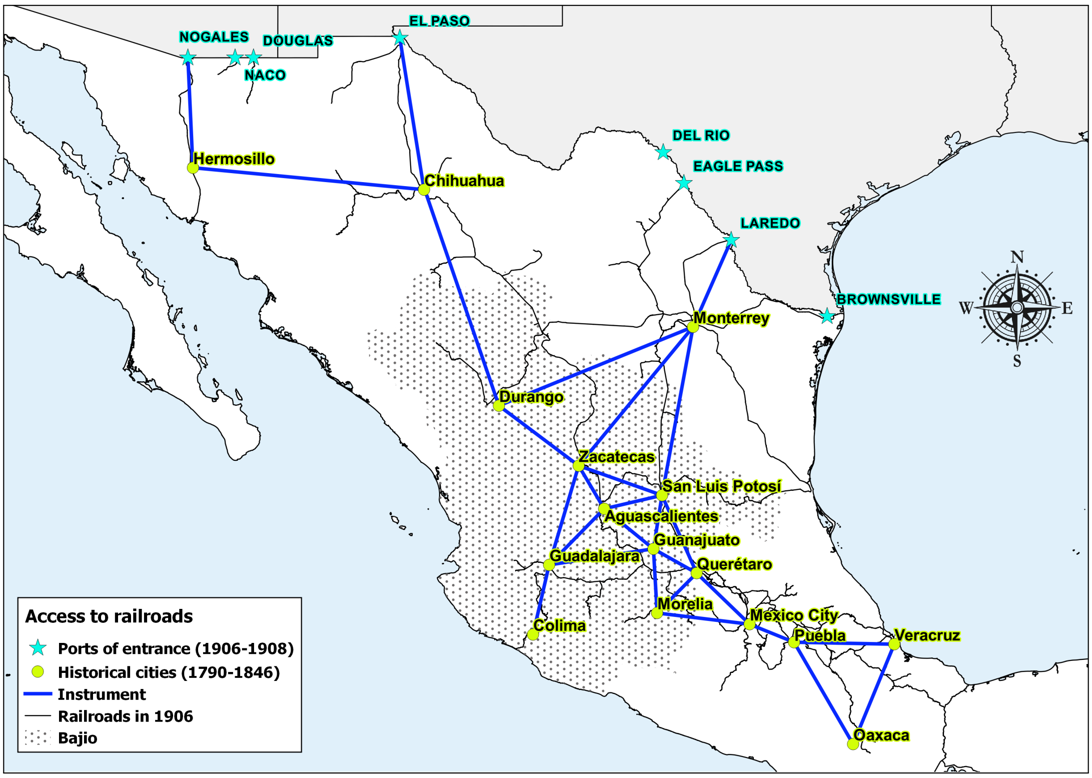


(d) Jan - Dec 1909

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.

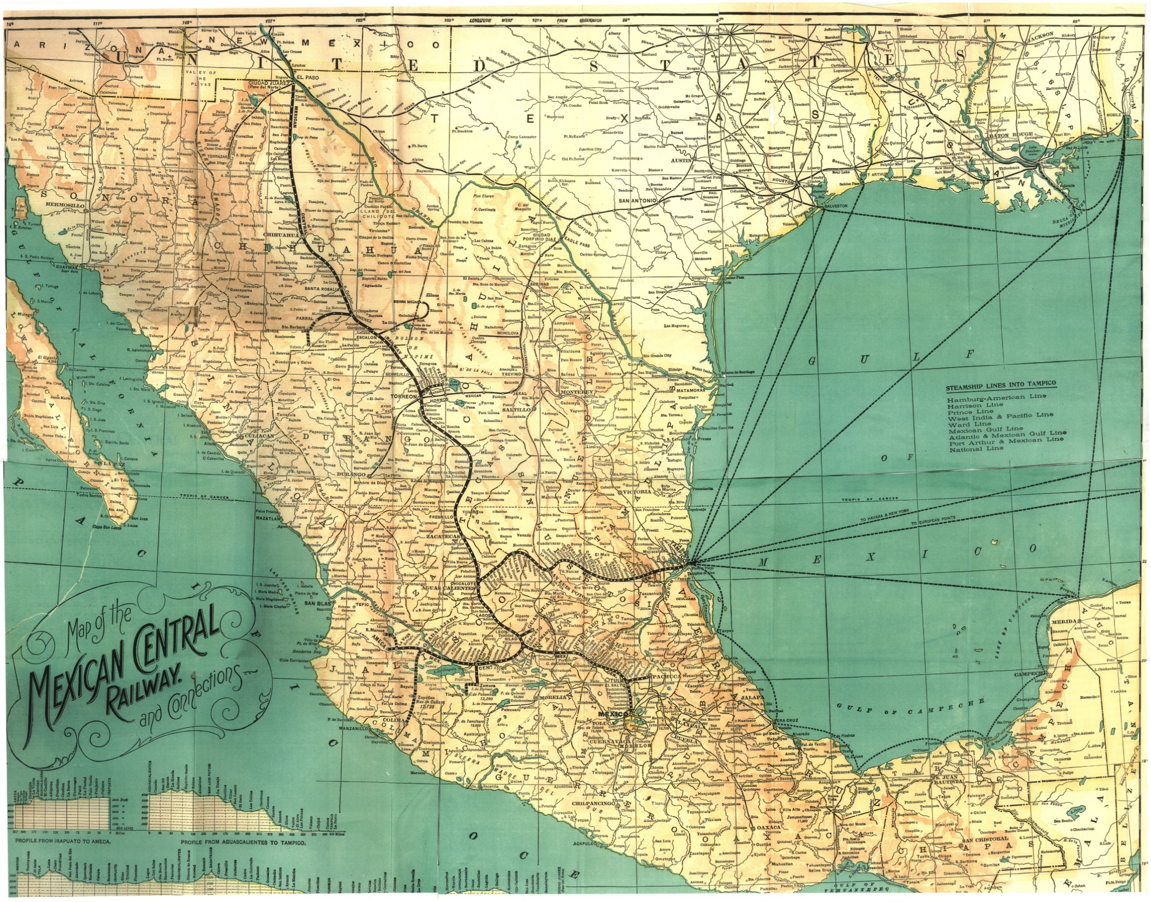
F. Instrumental variable strategy: straight lines

Figure 10: Historical cities (1790–1846) and straight lines as an 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 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).