



HAL
open science

Determinants of Mexico-US outwards and return migration flows: A state-level panel data analysis

Isabelle Chort, Maëlys de La Rupelle

► **To cite this version:**

Isabelle Chort, Maëlys de La Rupelle. Determinants of Mexico-US outwards and return migration flows: A state-level panel data analysis. 2017. hal-01516057

HAL Id: hal-01516057

<https://hal.science/hal-01516057>

Preprint submitted on 28 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DOCUMENT DE TRAVAIL

DT/2015-03

Determinants of Mexico-US outwards and return migration flows: A state-level panel data analysis

Isabelle CHORT
Maëlys de la RUPELLE

UMR DIAL 225

Place du Maréchal de Lattre de Tassigny 75775 • Paris Cedex 16 • Tél. (33) 01 44 05 45 42 • Fax (33) 01 44 05 45 45
• 4, rue d'Enghien • 75010 Paris • Tél. (33) 01 53 24 14 50 • Fax (33) 01 53 24 14 51

E-mail : dial@dial.prd.fr • Site : www.dial.ird.fr

DETERMINANTS OF MEXICO-US OUTWARDS AND RETURN MIGRATION FLOWS: A STATE-LEVEL PANEL DATA ANALYSIS

Isabelle Chort
PSL, Université Paris-Dauphine,
LEDa, IRD UMR DIAL, 75016 Paris, France
isabelle.chort@dauphine.fr

Maëlys de la Rupelle
THEMA,
Université de Cergy-Pontoise, 95011 Cergy-Pontoise
maelys.delarupelle@u-cergy.fr

Document de travail UMR DIAL
Février 2015

Abstract

In this paper, we investigate the determinants of the regional patterns of Mexico-US migration flows. Along with traditional economic determinants, we examine the role played by environmental factors and violence in Mexico in determining migration patterns and their evolutions. We estimate a micro-grounded gravity model of migration using a panel dataset of state-to-state emigration and return migration flows between Mexico and the US for the period 1995-2012. We exploit the time and dyadic dimension of the data to control for time-invariant and time-variant characteristics of destination states, including migration policies. Our results suggest that along with the traditional economic determinants of migration, climatic and social factors contribute to shaping regional migration patterns.

Key words: International migration, Mexico-U.S. migration, Gravity equation, Climate change, Natural disasters.

Résumé

Nous étudions dans cet article les déterminants des tendances régionales des flux migratoires entre le Mexique et les Etats-Unis. A côté des déterminants économiques traditionnels des migrations, nous explorons le rôle de facteurs environnementaux et sociaux sur les caractéristiques et l'évolution des flux migratoires entre états. Nous estimons un modèle de gravité micro-fondé, à partir de données de panel sur les flux migratoires entrants et de retour entre états mexicains et états-uniens sur la période 1995-2012. Nous exploitons la dimension temporelle et dyadique de nos données pour contrôler pour les caractéristiques des états de destination susceptibles d'affecter les flux migratoires, notamment les changements de politiques migratoires. Nos résultats suggèrent que les facteurs sociaux et climatiques contribuent à expliquer les tendances régionales observées.

Mots Clés : Migration internationale, Mexique, Etats-Unis, Modèle de gravité, Changement climatique, Catastrophes naturelles.

JEL Code : F22, J6, J68, R23.

1 Introduction

The flow of Mexican migrants to the United States is the largest bilateral migration flow worldwide. The US is by far the top destination of Mexican migrants: the percentage of Mexican international migrants going to the US is estimated to be between 94 and 99%. Yet, little is known about the regional patterns of migration flows between the two countries in the recent period. In addition, whereas Mexican migration to the US has been extensively studied since the 1980's, return migration of Mexicans is still under-documented.

This article contributes to the literature by providing an analysis of the determinants and the evolution of the regional patterns of both outward and return Mexico-US migration flows over the last 20 years. An important contribution of this paper is the creation of a panel database of Mexican state-to-US state yearly outward and return migration flows over the 1995-2012 period, using data from the individual Survey of Migration at the Northern Border of Mexico (*Encuesta sobre Migración en la Frontera Norte de México* or EMIF Norte). Based on information contained in the EMIF survey on origin and destination states of Mexicans crossing the border in either direction, we compute yearly flows for each pair of Mexican and US states. The resulting database is the first panel database of Mexico-US state-to-state outward and return migration flows covering both documented and undocumented migration. A detailed evaluation of the representativeness of the EMIF data is provided by Rendall et al. (2009). By comparing the EMIF data to US census data and data from two Mexican nationally representative demographic and employment surveys (ENE/ENOE and ENADID), they show that the EMIF data are less biased and include larger sample sizes than other annually collected US or Mexican data sources, and, which is of special interest to us, that they “represent reasonably well the geographic origins of Mexico’s migrants to and from the US”.

In order to account for the multi-dimensionality of migration decisions, we complement our state-to-state migration flow data with economic, geographic, climatic and social

data from various sources. We combine our migration data with state-level economic, demographic, and crime data on Mexican states of origin from the Mexican *Instituto Nacional de Estadística y Geografía* (INEGI). In addition, we compile different types of climatic data to account for both long-term changes in temperatures and rainfall and unexpected climatic phenomena: we construct state-level variables capturing deviations in precipitations from long-term averages using the monthly gridded time series provided by the Department of Geography of the University of Delaware, and exploit information on hurricanes having affected Mexican states from 1990 to 2012 using the Historical Hurricane Track tool of the US National Oceanic and Atmospheric Administration.

We exploit the regional variability in economic outcomes across both origin and destination states to explore the impact of traditional push and pull factors based on income differentials. We enrich the analysis by examining climatic, social, and geographic determinants of bilateral flows to cover a wide range of non-economic factors likely to explain the evolution of regional migration patterns. We also exploit the panel structure of our data to control for unobserved state-specific characteristics explaining migration flows.

This paper first contributes to the vast literature on Mexico-US migration, and more specifically relates to the few papers studying regional migration flows. As noted by Hanson and McIntosh (2010), the question of the scale of regional labor flows has given rise to surprisingly little academic research. Hanson and McIntosh (2010) partly fill this gap by exploiting the regional variation in the timing of the demographic transition across Mexican states to explain emigration flows from Mexico to the US from 1960 to 2000. They focus on the contribution of differential population growth between Mexico and the US to the observed surge in Mexico-US labor migration in the last two decades of the 20th century and use the variation in population growth across Mexican states and across time to identify their effect. However, they do not decompose flows along US states of destination, and their use of low-frequency census data is not suited to the analysis of the short-term determinants of migration.

Computing decadal emigration rates from Mexican census data, Hanson and McIntosh

(2010) find that labor supply shocks account for about one third of observed emigration from Mexico over 1977 to 2000. They recognize however that given the dramatic decline in fertility in Mexico since the 1970's, labor supply growth in Mexico is no longer a crucial push factor in the recent period.

Our paper extends and complements Hanson and McIntosh (2010) first by using information on destination states and thus considering dyadic flows, and second, by enlarging the set of potential factors affecting Mexico-US state-to-state flows based on the empirical literature on the determinants of migration flows, and finally, by focusing on the recent period. Since our data cover the 1995-2012 period, we are in particular able to assess the impact of the crisis on Mexico-US migration flows.

Villarreal (2014) also studies the determinants of migration flows from Mexico in the recent period, but with a microeconomic approach. He moreover focuses on the role played by labor demand at destination to explain the dramatic decline in Mexico-US migration rate that he observes since 2006, using data from the Mexican National Occupation and Employment Survey, and does not address the question of the evolution of regional patterns of Mexican migrants during the same period.

The geography of Mexico-US migration flows has been studied in particular by Durand, Massey and co-authors (see in particular Durand et al. (2001), Durand and Massey (2003)), with the recent contributions by Riosmena and Massey (2012) and Massey et al. (2010) who exploit new data sources providing a richer information on both the origin and destination of immigrants than census data.

The earliest statistics on both the origin and destination of Mexican migrants to the US can be found in Foerster (1925)¹, who documents the state of origin of 10,212 Mexican immigrants entering the US at three border ports² in April 1924. These data are partially cited by Woodruff and Zenteno (2007) who use them to assess the strength

¹Page 51. Made available online by Harvard University at <http://pds.lib.harvard.edu/pds/view/4905592?n=57&s=4&imagesize=1200&rotation=0>, accessed 02 January 2015.

²Two ports in Texas (San Antonio (5,205 entries) and El Paso (4,770 entries), and one in California (Los Angeles, 237 entries)

of their instrument for migration networks based on the railroad network in Mexico in the early twentieth century.

Indeed, the role of the railroad in shaping the Mexico-US migration flows has been largely documented, and continues to explain a large part of migration patterns as late as in the early 1990s (Durand et al. (2001); Borjas (2007)). Yet, in the recent period, the geography of Mexico-US migration flows has been subject to rapid changes, with both a diversification of origins and the emergence of new destinations (Riosmena and Massey (2012); Massey et al. (2010)). As rightly pointed out by Riosmena and Massey (2012), very few data sources document both the origin and destination states of Mexican immigrants in the US. Apart from the EMIF which is not mentioned by Riosmena and Massey (2012), only two data sources provide information allowing to reconstruct state-to-state migratory streams, the 2006 round of the ENADID (*Encuesta Nacional de la Dinámica Demográfica*), Mexico's National Survey of Population Dynamics, and the data collected in the *Matrícula Consular* Program. Riosmena and Massey (2012) use the 2006 round of the ENADID, which asked household members who had been to the US information about their place of destination. The data from 2006 round of the ENADID survey are representative of both legal and undocumented migrants, but they are cross-sectional and thus only provide a picture of migration flows between the Mexico and the US in 2006. Our data base constructed from the EMIF has the advantage of being a panel of 14 years covering the period 1995-2012, which allows us first to illustrate the evolution of migration flows, and second to control for state- or dyad-specific unobserved characteristics correlated with migration. Moreover, while Riosmena and Massey (2012) choose to analyse regional flows, we use more disaggregated data at the state level.

The data from the *Matrícula Consular* Program are used by Massey et al. (2010) who study the geography of undocumented migration between Mexico and the US. However, because of the data source they use, legal migration is out of the scope of their study. In addition, their data are not representative of the population of undocumented Mexican migrants and only provide a snapshot of the stock of undocumented migrants who chose

to register to the program between January 1 and October 31 of 2006.

The maps of migration flows that we constructed from the EMIF survey data provide an illustration of the rapid changes in both origins and destinations of Mexican migrants (see the maps in Appendix). Consistent with Riosmena and Massey (2012), these maps highlight in particular the end of the predominance of Mexico's West-Central states in the origin of Mexico-US flows, that accounted for around 50% of Mexican migrants from the 1920's to the early 1990s Durand et al. (2001).

The geography of return migration has been recently studied by Masferrer and Roberts (2012) who use the Mexican censuses of 1995, 2000, 2010 and the 2005 Population Count. However, since census data do not provide information on the state of residence in the US of return migrants, they cannot link information on US former states of residence of Mexicans with US experience with information on Mexican states of return.

This paper also builds on a growing literature that acknowledges the multi-dimensionality of migration decisions, and explores in particular the role of climatic factors. Our paper thus relates to the more general debate over the consequences of climate change on migration summarized in Piguet et al. (2011). The recent literature investigating the importance of climatic factors on international migration find contrasted results in general (see Beine and Parsons (2012) for a review). In the particular context of Mexico, several papers have stressed the importance of climatic factors (Munshi (2003), Pugatch and Yang (2011), Chort (2014)). Interestingly, Beine and Parsons (2012) find a significant effect of climatic factors on international migration only when conditioning upon origin countries' characteristics. In the Mexican context, Nawrotzki et al. (2013) find evidence of drought driven migration from dry states. We account for the potentially heterogeneous effects of climatic factors by considering separately rainfall shocks during the dry and the rainy season. We additionally control for violence in origin Mexican states by including a variable for the number of homicides at the state level, provided by the INEGI. Indeed, while homicides had been steadily declining through the 1990s and the mid-2000s, drug-related violence sharply increased after 2007 (Heinle et al., 2014)

and is found by Rios Contreras (2014) to explain a large part of migration flows from Mexico to the US, especially from border states.

Finally, this paper closely relates to the strand of the migration literature initiated by Pedersen et al. (2008) and Mayda (2010) that applies the gravity equation initially developed to account for trade flows (Anderson and Van Wincoop, 2003) to the estimation of the determinants of international migration flows using bilateral flow matrices. Recent papers have provided significant contributions to this field by emphasizing the microeconomic theory behind such models, drawing upon the income maximisation approach (Grogger and Hanson (2011); Bertoli and Moraga (2013); Ortega and Peri (2013) ; Beine and Parsons (2012)).

To the best of our knowledge, our paper is the first to investigate the determinants of Mexico-US migration flows using dyadic data at the federated state level for both origin and destination, and the first macroeconomic study to examine the impact of economic factors, climatic events and violence on migration simultaneously. In addition, we document the determinants of both outward and return migration flows and fully exploit the richness of our panel data to control for unobserved state and time characteristics correlated with migration. More generally, our paper contributes to the growing literature applying gravity models to the analysis of migration flows, and provides the first estimation of a gravity model of migration at the infra-country level.

Our results suggest that along with the traditional economic and geographic determinants of migration, climatic and social factors contribute to shaping regional migration patterns. We find in particular that income at origin has a non-linear impact on migration outflows, and our results confirm that the determinants of migration out- and return flows are not symmetrical: climatic and social factors in Mexico and migration networks are found to affect migration outflows, whereas they have no significant impact on return flows.

The article is structured as follows. Section 2 briefly presents the theoretical frame-

work behind our empirical specification and discusses our empirical model. The data are described in Section 3. Results are presented and discussed in section 4. Finally, Section 5 concludes.

2 Theoretical foundations and estimation

2.1 The income maximisation framework

We estimate in this paper a standard gravity model of migration which has already been used in particular by Grogger and Hanson (2011), Beine et al. (2011), Bertoli and Moraga (2013), and Beine et al. (2014). In this section, we first recall the main theoretical assumptions of such models, borrowing from the presentation framework proposed by Beine et al. (2014). A micro-foundation of gravity models applied to migration can be found in the canonical random utility model (RUM) of migration. We adapt the general theoretical framework to model migration between Mexican and US states. We assume that agents decide whether or not to migrate and where to move based on the maximisation of their utility across the full set of destinations including their home state. For the sake of simplicity, and since we do not have data on internal migration nor on migration to countries other than the US, we assume that the set of possible destinations for Mexican migrants is only made of all US states³.

The utility of an individual i located in origin (Mexican) state j at time t and migrating to US state k is written as⁴:

$$U_{ijkt} = a_{jkt} - c_{jkt} + \varepsilon_{ijkt} \tag{1}$$

³Note that unlike the above cited papers applying similar models to dyadic data, in our application the sets of origins and destinations do not overlap. The implications of such data structure are discussed below.

⁴The theoretical framework presented here is adapted to the modelling of Mexico-US migration. In the following empirical application, we additionally apply the same model to represent return migration. The only change as regards the initial theoretical framework consists in considering that Mexican immigrants in the US choose whether to stay in host US state j or return to Mexican state k .

with $a_{jk,t}$ the attractiveness of destination k at time t to individuals from origin j , is a deterministic component of utility based on observable characteristics of destination k ; $c_{jk,t}$ is the cost of migrating from j to k at time t , and $\varepsilon_{ijk,t}$ is an individual stochastic component of the utility of migrating from j to k at time t .

If we assume that $\varepsilon_{ij,t}$ is an iid extreme value distributed random term (McFadden, 1974) and denote p_{jkt} the proportion of individuals from country j deciding to migrate to country k at time t , we obtain:

$$E(p_{jkt}) = \frac{\exp(a_{jkt} - c_{jkt})}{\sum_{l \in D} \exp(a_{jlt} - c_{jlt})} \quad (2)$$

with D the set of available destinations.

The expected gross migration flow of migrants moving from state j to state k at time t , $E(m_{jkt})$, is obtained by multiplying the previous expression by the stock of population living in the state j at time t :

$$E(m_{jkt}) = \frac{\exp(a_{jkt} - c_{jkt})}{\sum_{l \in D} \exp(a_{jlt} - c_{jlt})} s_{jt} \quad (3)$$

As noted in Beine et al. (2014), if we assume that the attractiveness of state k does not depend on the state of origin j , and if we note $\exp(a_{kt}) = y_{kt}$, $\exp(-c_{jkt}) = \phi_{jkt}$ and $\sum_{l \neq j,k} \exp(a_{lt} - c_{jlt}) = \Omega_{jt}$, we obtain the following expression :

$$E(m_{jkt}) = \phi_{jkt} \frac{y_{kt}}{\Omega_{jt}} s_{jt} \quad (4)$$

In this equation, Ω_{jt} , which is referred to as a multilateral resistance term by Bertoli and Moraga (2013), in analogy with the trade literature (Rose and Van Wincoop, 2001) captures the fact that bilateral migration flows between j and k do not only depend on the relative attractiveness of origin state j and destination state k , but also on the attractiveness of all other destinations.

2.2 Estimation issues

Based on the above theoretical framework, the standard approach in the migration literature relies on the property of the independence from irrelevant alternatives (IIA) derived from the assumptions on the distribution of the stochastic component of utility. Indeed, when taking the ratio between $E(m_{jkt})$, the expected migration flow from j to k , and $E(m_{jjt})$ the expected number of stayers in j , under the above assumptions, the expression in 4 simplifies and Ω_{jt} and s_{jt} cancel out. Many papers exploit the IIA property and estimate the empirical counterpart of the ratio between $E(m_{jkt})$ and $E(m_{jjt})$ by taking the gross migration rate from j to k and log-linearizing the formula to express this migration rate as a function of the differential attractiveness of origin j and destination k (Clark et al. (2007),). However, as emphasized by Beine et al. (2014), failing to account for multilateral resistance leads to biased estimates of the impact of the different factors of interest on bilateral migration flows.

In order to circumvent the first of these two issues, we choose to estimate equation 4 in levels, using as a dependent variable alternately the gross migration flows from origin state j to destination state k , or the gross return migration flow from US state k to Mexican state j , with Poisson Pseudo-Maximum Likelihood (PPML). Using the PPML estimator, we avoid the problem of log-linearization which is found by Santos Silva and Tenreyro (2006) to lead to inconsistent estimates in the presence of heteroskedasticity. Moreover, as argued by Santos Silva and Tenreyro (2011), the PPML estimator performs well even with a large share of zeros in the the data. We are faced with the absence of flows on a number of bilateral corridors. Although the share of zero flows tends to decrease over time, reflecting the diversification of both origins and destinations of Mexican migrants, at the end of the period it is still close to 50% (see table 1 for Mexico-US flows, the percentages of zero cells being very similar for return flows). As Beine and Parsons (2012) who are faced with the same problem, we thus choose to rely on PPML methods.

Second, following Beine and Parsons (2012), and exploiting the longitudinal dimension

Table 1: Proportion of zero migration outflows in our state-level bilateral matrix

Year	Pct	
	Total flows	Male flows
1995	84.9	85.0
1999	75.7	76.3
2000	71.1	72.2
2001	70.8	71.5
2002	73.1	73.6
2004	73.9	74.1
2005	71.1	71.6
2006	66.7	67.5
2007	62.6	63.5
2008	61.8	62.5
2009	62.6	63.7
2010	54.7	55.9
2011	49.3	50.3
2012	48.4	49.2

Source: EMIF data, authors' calculations

of our data, we include a rich structure of fixed effects to partially capture the unobserved factors affecting bilateral migration flows, and more specifically, to account for the multilateral resistance factors to migration. Indeed, recent works in the trade literature (Fally, 2014) show that gravity equations estimated using the Poisson Pseudo-Maximum Likelihood estimator with origin and destination fixed effects are fully consistent with the structural constraints imposed to account for multilateral resistance factors (Anderson and Van Wincoop, 2003)⁵. We include in all regressions Mexican state dummies, that account for all time-invariant origin specific unobserved factors that may affect migration outflows (or return flows), and US state-year dummies, that capture both time-invariant and time-variant host specific characteristics susceptible to explain migration inflows. US state-year dummies allow us to control for economic factors (income, either GDP per capita or wages) and state-specific changes in immigration policies.

We detail in the next paragraphs the various factors entering the deterministic com-

⁵The alternative proposed by Bertoli and Moraga (2013) based on the CCE estimator developed by Pesaran (2006) is not applicable here given the relatively short longitudinal dimension of our panel (14 years) and the large proportion of zero flows.

ponent of utility.

Following the literature, we assume that migration costs depend on dyadic factors such as distance, contiguity, cultural proximity, networks, and on changes in immigration policies. First note that in our application, since we consider only two countries, Mexico and the US, we need to adapt our set of explanatory variables to account for the relative homogeneity of origins on the one hand, and destinations on the other. For instance, linguistic proximity is one of the traditional dyadic factors in the international migration literature which cannot be included in the set of dyadic variables in our study. Indeed, while the share of Hispanic population varies across destination states, and may capture a mixture of cultural and linguistic proximity and networks, the absence of heterogeneity with respect to these dimensions in the set of origin states justifies that we consider such a variable to be destination-specific instead of dyadic. Similarly, in our application, immigration policies implemented by US state legislatures apply to all Mexicans whatever their state of origin and are not included in the set of dyadic variables. In summary, all destination-state specific variables that account for cultural proximity, which may be hard to disentangle from migration networks, and immigration policies, are captured by the destination-year fixed effects.

Strictly speaking, the only dyadic factors that we include in our different econometric specifications are the geographical distance between states j and k , defined as the long-circle distance between their capital city, a dummy which equals one for pairs of Mexican and US states sharing a border, and a measure of migration networks between states j and k . As regards dyadic network variables, the classical approach in the international migration literature consists in proxying migration networks by historical bilateral migration stocks (Beine and Parsons (2012), see also Beine et al. (2014) for a review). However, as emphasized in the introduction of this paper, very few data sources document the regional dimension of the Mexico-US migration. In particular, exhaustive and representative statistics on the state of origin of the stock of Mexican migrants who reside in the different US states do not exist. We thus choose to capture bilateral migration

networks by constructing a dyadic variable made of the historical emigration rate from Mexican state j to the US in 1987 weighted by the distance between Mexican state j and US state k .

According to the above theoretical framework, outward and return migration flows from j to k depend on $y_{j,t}$ and $y_{k,t}$, the attractiveness of states j and k . Note that we focus our analysis on the determinants of migrations at origin (in Mexican states), and that all time-variant and time-invariant factors affecting the attractiveness of destination k are absorbed in the destination-year dummies. As for the attractiveness of origin state j , we include economic, climatic and social characteristics. The first economic factor to be considered should account for the level of income per capita. We follow the literature in using the GDP per capita.

Our environmental factors and climatic shock variables are of two types. Piguet et al. (2011) and Beine and Parsons (2012) differentiate short-run unexpected natural disasters, and deviation in climatic factors around long run average. We consider hurricanes and deviations in long-term averages of yearly precipitations.

Because of obvious endogeneity concerns when dealing with migration flows and economic variables simultaneously, migration flows at time t are related to lagged values of explanatory variables.

Finally, our estimation results may be biased if the error terms are serially and spatially correlated. The literature estimating gravity models of migration, reviewed by Beine et al. (2014) only focuses on country-to-country migration flows. As our observation unit are state-to-state migration flows, we are all the more concerned with spatial correlation issues. Indeed, since we consider Mexican-to-US flows, the structure of our database is such that all origin states on the one hand, and all destination states on the other hand are part of a single country. Moreover we have no data on internal migration flows in Mexico. As a result, the coefficient on the state level Mexican GDP variable would be most probably upward biased if we were not controlling for the correlation between GDP levels of all Mexican states: a decrease in the GDP per capita in Mexican

state j is very likely correlated with a decrease in the GDP per capita of other Mexican states, thus reducing the attractiveness of other Mexican states to individuals living in state j (which we cannot directly control for since we do not observe internal migration in our data) and altogether making US destinations more attractive for Mexicans from all states. In order to partially capture the impact of a change in the attractiveness of other potential destinations *which are not in our dataset*, ie other Mexican states, we include in the set of regressors a variable which is equal for each origin state to the log of the mean population weighted value of the GDP per capita in all other Mexican states.

3 Empirical specification

The basic equation of gross migration flows from Mexico to the US, that we estimate with PPML, is the following :

$$m_{jk,t} = \beta_0 + \beta_1 \ln(GDP_{j,t-1}) + \beta_2 CLIM_{j,t-1} + \beta_3 HOM_{j,t-1} + \beta_4 \ln(MEANGDPMEX_{j,t-1}) \\ + \beta_5 \ln(DIST_{jk}) + \beta_6 NETW_{jk} + \beta_7 BORDER_{jk} + D_j + D_{k,t} + \epsilon_{jk,t}$$

$m_{jk,t}$: the gross migration flow from the Mexican origin state j to the US destination state k at time t

$\ln(GDP_{j,t-1})$: the log of the real GDP per capita in state j at time $t - 1$

$CLIM_{j,t-1}$: different climatic variables for origin state j including the z-score of yearly precipitations at the state level and a dummy variable for hurricanes in Mexico at time $t - 1$ (see the Data section for more details on the construction of the climatic variables). We further control for unexpected weather shocks by including a variable for the number of hurricanes and storms affecting state j at time $t - 1$ and a variable for the maximum intensity of hurricanes affecting state j at time $t - 1$.

$\ln(HOM_{j,t-1})$: the log of the number of homicides in Mexican state j in year $t - 1$
divided by the total population of the state

$\ln(MEANGDPMEX_{j,t-1})$: the log of the mean value of the GDP per capita in all
Mexican states other than j at time $t - 1$.

$\ln(DIST_{jk})$: the log of the great-circle distance between the capitals of origin state j
and destination state k

$NETW_{jk}$: a measure of migration networks between origin state j and destination
state k . We use the migration rate to the US in Mexican state j in 1987 (provided
by the INEGI) multiplied by the inverse of the distance between the capitals of
origin state j and destination state k .

$BORDER_{jk}$: a dummy variable that equals 1 if Mexican and US states j and k have a
common border.

$D_j, D_{k,t}$ origin and destination-year fixed effects

The same equation is adapted to the modelling of gross return migration flows from
the US to Mexico :

$$m_{kj,t} = \beta_0 + \beta_1 \ln(GDP_{j,t-1}) + \beta_2 CLIM_{j,t-1} + \beta_3 HOM_{j,t-1} + \beta_4 \ln(MEANGDPMEX_{j,t-1}) \\ + \beta_5 \ln(DIST_{kj}) + \beta_6 NETW_{jk} + \beta_7 BORDER_{kj} + D_j + D_{k,t} + \epsilon_{kj,t}$$

The dependent variable is now $m_{kj,t}$, the gross return migration flow from US host
state k to Mexican state j at time t . We choose to keep the focus on the determinants of
return migration flows measured in Mexico: consistent with this choice, all independent
variables and the structure of the fixed effects included in the model are the same as
in the previous equation: ie Mexican states j fixed effects, which in the case of return

migration are destination states, and US states -year fixed effects, which are equivalent to the destination - year fixed effects in the first equation for outward flows.

Both our specifications thus includes variables for push-factors of migration (or equivalently pull factors of return migration), measured in Mexican states (income per capita, climatic factors, and homicides), dyadic variables specific to each migration corridor (distance, border and networks), and Mexican states and US states-year dummies capturing the impact of time-invariant origin-specific push-factors (or pull factors in the case of return migration) and both time-variant and time-invariant characteristics related to the attractiveness of each US state.

We also estimate several complementary specifications including as additional regressors the squared log of the GDP per capita, following Mayda (2010), and interaction terms between the log of the GDP per capita and income quartiles in order to control for potential credit constraints affecting migration costs. We additionally control for the unemployment rate in the Mexican state of origin (or destination in the case of return migration), to better capture economic opportunities in the home state.

In some specifications, we further investigate the impact of climate shocks by differentiating rain deviations from long term average during the rainy (from May to October) and the dry seasons, since they may have different implications.

4 Data

The data used in this paper come from different sources. Immigration and return migration flows are constructed using data from the EMIF surveys (Encuesta sobre Migración en la Frontera Norte de México)⁶.

⁶<http://www.colef.net/emif/>

4.1 The EMIF data

The EMIF data have been collected since 1993 at different points of the Mexico-U.S. border and aim at providing a representative picture of the flows of Mexicans crossing the border in either direction. The survey design relies on a multistage probability spatiotemporal sampling frame, where geographical and time units are chosen interactively. The sampling of geographical units (cities, zones and crossing points) is based on prior studies on the characteristics of Mexican migration flows, and more specifically on the probability proportionate to the size of migratory flows in the area. Further details on the survey design and the computation of the sampling weights are provided on the EMIF's website⁷. Each individual questionnaire is then assigned a sampling weight that accounts for this multistage sampling frame. We use sampling weights to construct aggregate flow variables.

4.2 Construction of the state-to-state flow matrix

15 waves of the EMIF survey are currently available: 1995 and 1999 to 2012⁸. We use information on the state of origin and destination of all surveyed individuals to construct a matrix of aggregate yearly migration flows from Mexican states to US states and return flows of Mexican immigrants from US states to Mexican states. The origin of individual i is here defined as the state of last residence in Mexico (respectively the US for return flows), and the destination is the self-declared state of destination of individual i , either in the US for outward flows, or in Mexico for return migrants. Observations with missing information on either origin or destination are dropped in the following analysis. Each individual observation is then weighted using the EMIF sampling weights in order to ensure the representativeness of our constructed flows. We also exploit individual data on gender to construct specific sub-flows. In the following empirical analysis, we choose to

⁷<http://www.colef.net/emif/diseniometodologico.php>

⁸We also drop the year 2003 because information on the US state of destination for Mexican migrants is missing.

focus on male flows, since the evaluation of the survey coverage by Rendall et al. (2009) suggests that migrant women are under-represented in the EMIF⁹.

In order to investigate the determinants of state-to-state flows, we merge additional data to the migration flows matrix. As regards climatic factors, we construct a state-level data set of hurricanes affecting Mexico between 1990 and 2012, from the Historical Hurricane Track tool developed by the U.S. National Oceanic and Atmospheric Administration (NOAA)¹⁰. We gather information on the number and intensity of hurricanes and storms affecting each Mexican State. We construct three yearly state-level variables synthesizing this information: the number of hurricanes and storms, the maximal storm intensity registered in the year, and the cumulated intensity of all registered storms.

In addition, we use satellite data from the “Tropical Rainfall Measuring Mission” (TRMM)¹¹ to construct state-level variables capturing deviations in precipitations from long-term averages. The TRMM is a joint project between the NASA and the Japanese Aerospace Exploration Agency which has been launched in 1997 to study tropical rainfalls, and is therefore well adapted to the Mexican context. Moreover, various technological innovations (including a precipitation radar, flying for the first time on an earth orbiting satellite) and the low flying altitude of the satellite increase the accuracy of the climatic measures. Interestingly enough, the TRMM products combine satellite measures with monthly terrestrial rain gauge data. Last, the measures are provided for 0.25 x 0.25 degree grid squares (around 25 km X 25 km), which allows us to construct very precise climatic variables. We construct rainfall variables for the two main meteorological seasons in Mexico, the rainy season and the dry season (the rainy season spanning from May to October). Measures are computed at the Mexican state level using the boundaries of Mexican states. We apply the same strategy as Chort (2014) and Pugatch and Yang (2011) and create state-level yearly normalized rainfall variables (rainfall z-scores)¹².

⁹The results presented here are robust to using total (both male and female) flow data

¹⁰<http://www.csc.noaa.gov/hurricanes/>

¹¹A survey published in 1998 in the *American Journal of Agricultural Economics* stresses the progress expected in improved climate measure and forecast from the TRMM mission.

¹²To construct these rainfall z-scores, we first assign grid points to states based on latitude and longitude coordinates, then sum up monthly data to obtain yearly rainfall variables and compute state-

Geographic controls include, for each observation, the log distance between the two states, calculated using the great circle formula, and a dummy variable for the pairs of bordering states.

State-level data on population, income, agriculture and crime for Mexico come from the Mexican Instituto Nacional de Estadística y Geografía (INEGI)¹³¹⁴.

4.3 Summary statistics

Our matrix is made of 1,632 cells corresponding to all possible pairs of the 51 US and 32 Mexican states. Since we exploit 14 waves of the EMIF survey, we have a total of 22,848 observations, including zero cells. In the following analysis we drop from our sample all observations corresponding to corridors exhibiting zero migration flows *throughout* the 1995-2012 period. Note that according to Santos Silva and Tenreyro (2006) estimates obtained with the PPML method are rather insensitive to the restriction of the sample to non-zero flows. As shown in table 1, even after dropping pairs with zero flows throughout the period, the proportion of zeros remains high, which justifies our choice of the PPML estimator.

Based on the bilateral flow matrix obtained using information contained in the EMIF, we constructed maps showing for each survey year the origins and destinations of Mexican migrants to the US. These maps, shown in Appendix, provide a clear illustration of the changes in Mexico-US migration patterns since the mid-2000s, consistent with recent studies using other data sources (see in particular Riosmena and Massey (2012)). Indeed,

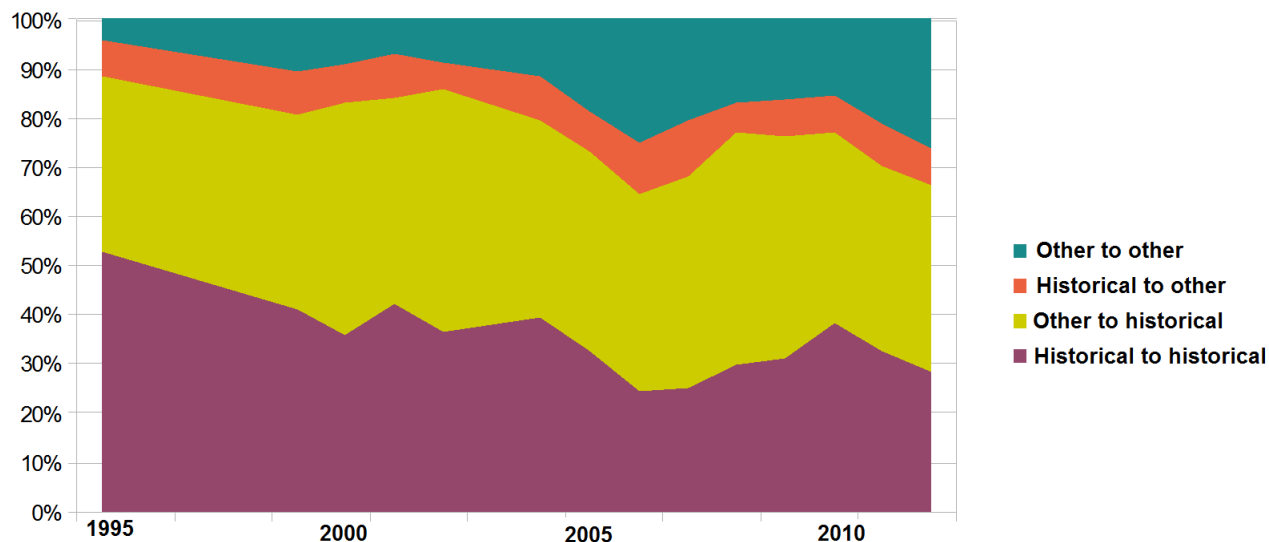
level averages for each year, state-level long term averages and state-level standard deviations. Long term averages are obtained by combining the above described satellite data for the 1998-2012 period, to the monthly gridded time series provided by the Department of Geography of the University of Delaware (http://climate.geog.udel.edu/~climate/html_pages/download.html#P2011rev) for the 1949-1997 period. The normalized variable is the state-level rainfall value minus the state-level long-run mean, divided by the state-level standard deviation over the observation period. For example, a positive value for year t in state j means that t has been an especially rainy year in state j . Conversely, a negative value means that precipitations have been lower than (long-term) average in that state in year t .

¹³<http://www.inegi.org.mx/>

¹⁴Some of our variables, and in particular Mexican population at the state level, are linearly extrapolated for the years in which they are not available, based on the values for other years for the same state.

we observe a double shift in migration patterns with both the diversification of origin states and the emergence of new destinations in the US.

Figure 1: Regional decomposition of Mexico-US state-to-state flows from 1995 to 2015



Source: authors' calculation based on the database constructed using EMIF data. Historic origin Mexican states are Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Luis Potosí and Zacatecas (Durand et al., 2001). Historic destination US states are Arizona, California, New Mexico and Texas (Woodruff and Zenteno (2007) citing Foerster (1925))

The relationship between the two geographic shifts is further explored in Figure 1 which represents the evolution of the share of Mexico-US migrants from and to historical migration states. Following (Durand et al., 2001), historical migrant sending Mexican states are Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Luis Potosí and Zacatecas. Historical destination US states are Arizona, California, New Mexico and Texas (Woodruff and Zenteno (2007) citing Foerster (1925)). Figure 1 illustrates the decline in the proportion of migrants from the Mexican Historical Region to traditional US destination states. While the proportion of this migration stream was still above 50% of total Mexico-US flows in 1995, it has fallen to 30% at the end of the period. At the same time, the proportion of migrants from new origins to non traditional

destinations has risen from less than 5% to 25%.

Table 2: Evolution of Mexico-US migration flows

Year	Male outflow size	Male return flow size	Male return flow size (airports)
1995	399,173	435,610	
1999	429,695	361,542	
2000	362,326	314,425	
2001	303,204	607,411	
2002	618,515	490,863	
2004	429,587	326,326	
2005	601,318	346,442	
2006	730,496	421,606	
2007	750,919	471,217	
2008	608,272	511,555	
2009	517,693	711,576	206,084
2010	367,489	370,943	183,700
2011	230,742	315,242	140,700
2012	228,040	259,689	113,072

Source: EMIF data, authors' calculations.

Finally, table 2 shows the sharp decline in raw Mexico-US male migration flows, in absolute numbers, after the 2008 crisis, following the jump of the first half of the 2000s. Interestingly, relying here on data collected at the border, and focusing on flows in level, we find results that slightly contrast with those of Villarreal (2014) regarding the responsiveness of migration to the economic downturn of 2008. Indeed, Villarreal (2014), using Mexican Employment Survey data (ENOE), finds that the migration *rate* to the US (the ratio of Mexican immigrants to the Mexican population) began to decrease as early as 2006, while we find that flows in absolute terms only began to decline in 2008. Table 2 also suggest that return flows outnumber immigration flows after 2009. However, a change in the construction of flows series has occurred in 2009, since the EMIF has been extended to airports, in order to include all migrants flying back to Mexico from the US (shown in column 3). In the following regression analysis, we consider total return flows, including airport passengers, but year dummies capture the potential effect of the change in the definition of our dependent variable in 2009.

Summary statistics for all explanatory variables are presented in tables 5 and 6 in

Appendix.

5 Results

5.1 Mexico-US flows

Estimation results for Mexico-US flows are shown in table 3. The dependent variable is the size of the flow (in levels), and all specifications are estimated using the Poisson Pseudo-Maximum Likelihood estimator. Column (1) presents the basic specification, column (2) explores the potential heterogeneity in the impact of climatic variables depending on the meteorological season, column (3) and (4) investigate the existence of credit constraints, or more generally the potential non-linear effect of income on migration, and column (5) additionally controls for unemployment in origin states to better proxy for labor market conditions at home. All specifications include origin and destination-year fixed effects, which capture the impact of time invariant characteristics of origin Mexican states, and the effect of both time-variant and time-invariant characteristics of US states of destination, including economic performances, or state-specific immigration policies. Results on dyadic variables, such as the geographical distance, sharing a common border, and migration networks all have the expected sign: the former leads to a decrease in migration flows while the latter two increase bilateral flows. The average distance between origin and destination is 2505 km. An increase by one standard deviation represents an increase of 32%, and decreases migration flows by 68%. Sharing a common border has a strong impact on migration flows : flows between border states are 133% larger than other flows. Migration networks are found to have a strong and significant impact. The migration network variable is a dyadic variable combining the emigration rate to the US for Mexican state j in 1987 with the distance between Mexican state j and US state k . The variable has been constructed so that it depends positively on both the distance between Mexican state j and US state k decreases, and the historical migration rate. When considering the average value of the network variable, one standard deviation more means

an increase by 89%, and translates into an increase in migrants flows by 180%. This is a sizeable effect. If one considers the average Mexico-US migrants flow size, an increase of 180% corresponds approximately to an increase by one third of a standard deviation, and thus explains an important part of migration flows.

More surprising at first is the sign of the coefficient on the log of the GDP per capita at origin. Whereas the literature exploring the determinants of bilateral flows of international migrants in general finds a negative impact of income at origin on migration outflows, our results suggest that income at origin increases migration. The elasticity of migration flows to the GDP per capita at origin is found to be close to 1.5. The structure of fixed effects included in all specifications capture all origin time invariant characteristics, and destination time variant and time invariant characteristics likely to affect migration. After controlling for other time variant factors at the Mexican state level (violence, climatic shocks) and dyadic factors, we find that the GDP per capita at origin has a positive impact on the size of outward migration flows. This finding specifically means that after controlling for the average level of the GDP per capita at origin with origin fixed effects, and after controlling for changes at destination which would affect the income differential between origin and destination through destination - year fixed effects, economic growth has a positive impact on migration flows. The average growth rate of the GDP per capita in Mexican states over the period of interest is 1.97%, and implies thus a 3% increase in migration outflows.

This results suggest that Mexico-US migrants are credit-constrained, which is confirmed by our estimates of specification (3). We follow Mayda (2010) and include in column (3) a quadratic term for income per capita. Specification in column (3) suggests that income at origin, proxied by the GDP per capita, has a positive impact on migration at low levels of income, and has a negative effect at high levels. Column (4) further explores this issue by interacting the GDP per capita variable with the different state-level income quartiles in 1995. Our results suggest that income at origin has a non-linear impact on outward migration flows: all else equal, a one percent increase in the GDP per

capita leads to an increase in emigration by 2% for origin Mexican states which were in the first income quartile in 1995, while the increase is only 1.2% for states in the second income quartile, and is not significant for richer states.

The coefficients on the population-weighted mean of the GDP per capita in all other Mexican states, included in all specifications to capture the potential impact of economic conditions in alternative (Mexican) destinations, are not significant in any of our specifications. According to our results, an increase in the GDP per capita in other Mexican states does not significantly divert flows away from international routes.

We find a negative impact of violence, measured by the log share of homicides in the total population of the state, on emigration. The effect is however small - while the share of homicides has an average value of 0.14 per thousand and a standard deviation of 0.169 per thousand, an elasticity of -0.3 implies limited consequences on migration flow size. This result may be reconciled with those of Rios Contreras (2014) who finds a significant impact of violence on Mexican emigration to the US. Indeed, while the analysis of Rios Contreras (2014) is focused on border towns, our bilateral flow matrix includes all Mexican states.

As for climatic variables, we find no significant impact of any of our hurricanes variables on migration flows. Specification (1) in Table 3 suggests that the deviations in yearly precipitations have a negative impact on migration flows, however this effect is small and non significant. The relationship between rain shocks and emigration is further explored in the following specifications: we create two variables to separately investigate the impact of rain shocks during the rainy and the dry season. We then find no significant effect of rain anomalies on migration during the rainy season, and a negative effect of abnormally high precipitations on migration during the dry season. The latter effect is consistent with drought driven migration: lower than average precipitations in the dry season increase emigration flows.

Finally, the unemployment rate at home is not found to have a significant impact on migration outflows.

Table 3: Determinants of state-to-state Mexico-US migration flows - PPML estimation

Dependent variable:	Migration flows (in level)				
	(1)	(2)	(3)	(4)	(5)
Ln distance between capitals	-1.999*** (0.70)	-2.025*** (0.69)	-2.110*** (0.67)	-3.833 (4.23)	-1.377* (0.75)
Common border	0.846*** (0.28)	0.857*** (0.28)	0.861*** (0.28)	0.856*** (0.28)	0.832*** (0.28)
Ln migration network	1.983*** (0.40)	1.940*** (0.39)	1.833*** (0.39)	0.122 (4.28)	2.453*** (0.49)
Ln GDP per capita (lag 1)	1.529** (0.64)	1.497** (0.63)	10.000** (4.95)		2.451*** (0.79)
Ln mean GDP per cap other Mexican states (lag 1)	6.758 (6.31)	5.449 (6.28)	5.559 (5.67)	5.801 (6.16)	8.047 (8.53)
Ln share of homicides (lag 1)	-0.312*** (0.06)	-0.301*** (0.06)	-0.275*** (0.06)	-0.296*** (0.06)	-0.269*** (0.08)
Nb hurricanes and tropical storms (lag 1)	0.103 (0.07)	0.064 (0.07)	0.049 (0.07)	0.056 (0.07)	0.107 (0.08)
Hurricane maximum intensity (lag 1)	-0.044 (0.04)	-0.040 (0.04)	-0.034 (0.04)	-0.038 (0.04)	-0.064 (0.04)
Yearly rain anomalies (z-score) (lag 1)	-0.008 (0.03)				
Rain anomalies during rainy season (z-score) (lag 1)		0.013 (0.03)	0.013 (0.03)	0.013 (0.03)	-0.030 (0.03)
Rain anomalies during dry season (z-score) (lag 1)		-0.110*** (0.02)	-0.111*** (0.02)	-0.108*** (0.02)	-0.092*** (0.03)
Ln state GDP per capita squared (lag 1)			-1.070* (0.60)		
Ln GDP per cap - quartile 1 (lag 1)				2.011** (0.99)	
Ln GDP per cap - quartile 2 (lag 1)				1.246* (0.65)	
Ln GDP per cap - quartile 3 (lag 1)				1.488 (1.03)	
Ln GDP per cap - quartile 4 (lag 1)				0.504 (0.85)	
Unemployment rate (lag 1)					0.006 (0.03)
Constant	-21.496 (30.20)	-15.497 (30.21)	-31.674 (30.39)	-2.606 (51.38)	-36.473 (42.17)
Origin FE	Yes	Yes	Yes	Yes	Yes
Destination - year FE	Yes	Yes	Yes	Yes	Yes
R^2	0.749	0.753	0.754	0.754	0.757
N	15667	15667	15667	15667	13435

Notes: Standard errors clustered at the Mexican state-US state pair level in parenthesis

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Determinants of state-to-state return migration flows - PPML estimation

Dependent variable:	Migration flows (in level)				
	(1)	(2)	(3)	(4)	(5)
Ln distance between capitals	-4.467*** (0.68)	-4.461*** (0.68)	-4.451*** (0.68)	-5.016** (2.49)	-4.595*** (0.72)
Common border	1.195*** (0.27)	1.196*** (0.27)	1.194*** (0.27)	1.194*** (0.27)	1.202*** (0.27)
Ln network	-0.193 (0.46)	-0.184 (0.46)	-0.170 (0.45)	-0.729 (2.52)	-0.109 (0.51)
Ln GDP per capita (lag 1)	-1.866*** (0.68)	-1.860*** (0.67)	0.117 (3.26)		-1.894*** (0.71)
Ln mean GDP per capita other Mex. states (lag 1)	4.338 (5.71)	4.184 (5.61)	4.632 (5.57)	6.681 (5.41)	0.070 (7.50)
Ln share of homicides (lag 1)	-0.174 (0.11)	-0.170 (0.11)	-0.158 (0.12)	-0.147 (0.11)	-0.130 (0.13)
Nb hurricanes and tropical storms (lag 1)	0.037 (0.08)	0.049 (0.08)	0.040 (0.09)	0.046 (0.08)	0.090 (0.10)
Hurricane maximum intensity (lag 1)	0.009 (0.04)	0.008 (0.04)	0.011 (0.04)	0.006 (0.04)	-0.049 (0.06)
Yearly rain anomalies by state (z-score) (lag 1)	-0.005 (0.04)				
Rain anomalies during rainy season (z-score) (lag 1)		-0.024 (0.04)	-0.023 (0.04)	-0.022 (0.04)	-0.072 (0.05)
Rain anomalies during dry season (lag 1)		0.031 (0.03)	0.031 (0.03)	0.039 (0.03)	0.059* (0.03)
Ln GDP per capita squared (lag 1)			-0.240 (0.39)		
Ln of GDP per cap - quartile 1 (lag 1)				-0.918 (0.72)	
Ln of GDP per cap - quartile 2 (lag 1)				-2.142*** (0.69)	
Ln of GDP per cap - quartile 3 (lag 1)				-1.188 (0.78)	
Ln of GDP per cap - quartile 4 (lag 1)				-2.387*** (0.85)	
Unemployment rate (lag 1)					0.043 (0.04)
Constant	24.858 (28.29)	25.404 (27.89)	19.552 (28.96)	16.819 (34.40)	46.873 (37.18)
Mexican state FE	Yes	Yes	Yes	Yes	Yes
US state - year FE	Yes	Yes	Yes	Yes	Yes
R^2	0.877	0.876	0.876	0.878	0.887
N	16312	16312	16312	16312	13984

Notes: Standard errors clustered at the Mexican state-US state pair level in parenthesis

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 Return migration

Table 4 presents the estimation results of the determinants of state-level return migration flows. The specifications are the same as in table 3. All specifications include Mexican state fixed effects and US state - year fixed effects capturing the effect of all time invariant factors at the Mexican state level that correlate with return migration, and the impact of time variant and time invariant characteristics of US states on return flows, including the economic downturn of 2008 of the restrictive immigration policies implemented in several states in the second part of the 2000s. Unsurprisingly, the signs of the coefficients on our first two dyadic variables proxying for migration costs, the distance and contiguity variables, are the same as for outward flows: distance between US state k and Mexican state j is negatively correlated with return migration from k to j . Indeed, return migration costs related to distance are expected to be symmetrical for Mexican emigrants and return migrants. Conversely, the coefficient on the network variable is no longer significant for return migration flows: a larger network in US state k is not found to have any impact on return migration from this state. Again, this finding is not surprising: the cost-decreasing impact of migrant networks are specific to outward migration, since networks may provide help to find a job or a shelter, or more generally provide potential migrants with information on economic opportunities. By contrast, a stronger network at destination is not expected to decrease the cost of potential returns.

We find no significant impact of the share of homicides, suggesting that the recent drug-related increase in violence observed in some Mexican states does not deter Mexican migrants in the US from coming back home.

Contrary to what we found for migration outflows, climatic variables are not found to have any significant impact on return flows, suggesting that return migration is less sensitive to climatic shocks in Mexico than emigration.

The impact of the GDP per capita is symmetrical to what was observed in the case of Mexican outmigration: an increase in the GDP per capita of Mexican state j is found

to negatively affect return flows from the US to this state. One possible interpretation could be that settling or investing in a Mexican state could be more profitable when growth is low. Specifications in columns (3) and (4) suggest that the effect of income is non-linear: the negative coefficient is only significant for states which were in the second and fourth income quartile in 1995. Finally, unemployment rate in Mexico is not found to significantly affect return migration.

6 Conclusion

This paper studies the economic, climatic, and social determinants of state-to-state Mexico-US outward and return migration flows over the 1995-2012 period. Using data from a survey conducted at the Mexico-US border we constructed the first panel of bilateral state-to-state flows. Based on this bilateral flow matrix, we first constructed maps that illustrate the recent changes in the regional migration patterns between Mexico and the US: the diversification of origin Mexican states coincide with the emergence of new destinations in the US. Bilateral flow data emphasize the connection between the two evolutions.

Second, we estimate a gravity model of migration to assess the impact of economic, climatic and social variables on regional migration flows. We focus our analysis on out migration “push-factors” (or, symmetrically, return migration “pull factors”), measured in Mexican states of origin, whilst controlling for both time-variant and time-invariant characteristics of US states by including a rich structure of fixed-effects, fully exploiting the longitudinal dimension of our data. Moreover, we control for dyadic factors likely to affect bilateral flows, such as distance, contiguity, and networks. Our results suggest that along with the traditional economic and geographic determinants of migration, climatic and social factors contribute to shaping regional migration patterns. We find in particular that income at origin has a non-linear impact on migration outflows, and our results confirm that the determinants of migration out- and return flows are not symmetrical:

climatic and social factors in Mexico and migration networks are found to affect migration outflows, whereas they have no significant impact on return flows.

Appendix

Table 5: Summary statistics Mexico-US flows

Variable	Mean	Std. Dev.	Min.	Max.	N
Weighted size male MEX-US flow	337.488	1934.396	0	71518	15667
Distance	2505.15	793.657	327.703	6775.89	15667
Common border	0.007	0.084	0	1	15667
Network	1.195	1.075	0.031	9.582	15667
GDP per capita (lag 1)	68.137	30.257	27.548	175.109	15667
GDP per cap (lag 1) by quartiles :					
quartile 1	40.39	6.944	27.548	53.868	4214
quartile 2	56.574	14.32	39.06	109.291	3933
quartile 3	72.952	10.734	51.059	93.356	3956
quartile 4	108.359	30.2	74.680	175.109	3564
Mean GDP per cap other Mex. states (lag 1)	70.566	4.65	52.105	79.375	15667
Share of homicides (lag 1)	0	0	0	0.002	15667
Share of homicides per 1000 inhabitants (lag 1)	0.14	0.169	0.017	1.885	15667
Nb hurricanes and tropical storms (lag 1)	0.268	0.611	0	4	15667
Hurricane maximum intensity (lag 1)	0.439	1.037	0	7	15667
Yearly rain anomalies (z-score) (lag 1)	0.265	0.972	-2.661	3.248	15667
Rain anomalies during rainy season (z-score) (lag 1)	0.331	1.03	-2.666	3.424	15667
Rain anomalies during dry season (lag 1)	-0.067	1.068	-2.958	4.379	15667
Unemployment rate (lag 1)	3.048	1.608	0.2	9	13435

Table 6: Summary statistics US-Mexico flows

Variable	Mean	Std. Dev.	Min.	Max.	N
Weighted size male US-MEX flow	354.51	3389.316	0	210020	16312
Distance	2512.286	777.062	327.703	6775.89	16312
Common border	0.007	0.083	0	1	16312
Network	1.189	1.065	0.031	9.582	16312
GDP per capita (lag 1)	68.001	30.291	27.548	175.109	16312
GDP per cap (lag 1) by quartiles :					
quartile 1 (lag 1)	40.259	6.92	27.548	53.868	4400
quartile 2 (lag 1)	56.364	14.196	39.06	109.291	4094
quartile 3 (lag 1)	72.843	10.709	51.059	93.356	4097
quartile 4 (lag 1)	108.277	30.178	74.680	175.109	3721
Mean GDP per cap other Mex. states (lag 1)	70.412	4.667	52.105	79.375	16312
Share of homicides (lag 1)	0	0	0	0.002	16312
Share of homicides per 1000 inhabitants (lag 1)	0.141	0.17	0.017	1.885	16312
Nb hurricanes and tropical storms (lag 1)	0.264	0.607	0	4	16312
Hurricane maximum intensity (lag 1)	0.431	1.024	0	7	16312
Yearly rain anomalies (z-score) (lag 1)	0.263	0.974	-2.661	3.248	16312
Rain anomalies during rainy season (z-score) (lag 1)	0.326	1.033	-2.666	3.424	16312
Rain anomalies during dry season (lag 1)	-0.06	1.078	-2.958	4.379	16312
Unemployment rate (lag 1)	3.041	1.612	0.2	9	13984

References

- Anderson, J. E. and Van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *The American Economic Review*, 93(1):170–192.
- Beine, M., Bertoli, S., and Fernández-Huertas Moraga, J. (2014). A practitioners' guide to gravity models of international migration. Working Papers 2014-03, FEDEA.
- Beine, M., Docquier, F., and Özden, a. (2011). Diasporas. *Journal of Development Economics*, 95(1):30–41.
- Beine, M. and Parsons, C. (2012). Climatic factors as determinants of International Migration. Discussion Papers (IRES - Institut de Recherches Economiques et Sociales) 2012002, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).
- Bertoli, S. and Moraga, J. F.-H. (2013). Multilateral resistance to migration . *Journal of Development Economics*, 102(0):79 – 100.
- Borjas, G. (2007). *Mexican Immigration to the United States*. National Bureau of Economic Research Conference Report. University of Chicago Press.
- Chort, I. (2014). Mexican Migrants to the US: What Do Unrealized Migration Intentions Tell Us About Gender Inequalities? . *World Development*, 59:535 – 552.
- Clark, X., Hatton, T. J., and Williamson, J. G. (2007). Explaining us immigration, 1971-1998. *The Review of Economics and Statistics*, 89(2):359–373.
- Durand, J. and Massey, D. (2003). *Clandestinos: migración México-Estados Unidos en los albores del siglo XXI*. América Latina y el nuevo orden mundial. Universidad Autónoma de Zacatecas.

- Durand, J., Massey, D. S., and Zenteno, R. M. (2001). Mexican Immigration to the United States: Continuities and Changes. *Latin American Research Review*, 36(1):pp. 107–127.
- Fally, T. (2014). Structural gravity and fixed effects.
- Foerster, R. F. (1925). *The racial problems involved in immigration from Latin America and the West Indies to the United States: a report submitted to the secretary of labor/ by Robert F. Foerster.*
- Grogger, J. and Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, 95(1):42–57.
- Hanson, G. H. and McIntosh, C. (2010). The great Mexican emigration. *The Review of Economics and Statistics*, 92(4):798–810.
- Heinle, K., Rodríguez Ferreira, O., and Shirk, D. A. (2014). Drug violence in Mexico: Data and analysis through 2013. *Justice in Mexico Project, University of San Diego.*
- Masferrer, C. and Roberts, B. (2012). Going Back Home? Changing Demography and Geography of Mexican Return Migration. *Population Research and Policy Review*, 31(4):465–496.
- Massey, D. S., Rugh, J. S., and Pren, K. A. (2010). The geography of undocumented Mexican migration. *Mexican studies= Estudios mexicanos*, 26(1):129.
- Mayda, A. (2010). International migration: a panel data analysis of the determinants of bilateral flows. *Journal of Population Economics*, 23(4):1249–1274.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press: New York.
- Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market. *The Quarterly Journal of Economics*, 118(2):549–599.

- Nawrotzki, R., Riosmena, F., and Hunter, L. (2013). Do rainfall deficits predict u.s.-bound migration from rural mexico? evidence from the mexican census. *Population Research and Policy Review*, 32(1):129–158.
- Ortega, F. and Peri, G. (2013). The effect of income and immigration policies on international migration. *Migration Studies*, 1(1):47–74.
- Pedersen, P. J., Pytlikova, M., and Smith, N. (2008). Selection and network effects—migration flows into oecd countries 1990–2000. *European Economic Review*, 52(7):1160–1186.
- Pesaran, M. H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4):967–1012.
- Piguet, E., Pécoud, A., and De Guchteneire, P. (2011). Migration and climate change: an overview. *Refugee Survey Quarterly*, page hdr006.
- Pugatch, T. and Yang, D. (2011). The Impact of Mexican Immigration on U.S. Labor Markets: Evidence from Migrant Flows Driven by Rainfall Shocks. Technical report.
- Rendall, M. S., Aguila, E., Basurto-Dávila, R., and Handcock, M. S. (2009). Migration between Mexico and the US estimated from a border survey. In *annual meeting of the Population Association of America*.
- Rios Contreras, V. (2014). The Role of Drug-Related Violence and Extortion in Promoting Mexican Migration: Unexpected Consequences of a Drug War. *Latin American Research Review*, 49(3):199–217.
- Riosmena, F. and Massey, D. S. (2012). Pathways to El Norte: Origins, Destinations, and Characteristics of Mexican Migrants to the United States. *International Migration Review*, 46(1):3–36.
- Rose, A. K. and Van Wincoop, E. (2001). National money as a barrier to international trade: The real case for currency union. *American Economic Review*, pages 386–390.

- Santos Silva, J. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics*, 88(4):641–658.
- Santos Silva, J. M. and Tenreyro, S. (2011). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters*, 112(2):220–222.
- Villarreal, A. (2014). Explaining the Decline in Mexico-U.S. Migration: The Effect of the Great Recession. *Demography*, 51(6):2203–2228.
- Woodruff, C. and Zenteno, R. (2007). Migration networks and microenterprises in Mexico. *Journal of Development Economics*, 82(2):509 – 528.