Measuring Geographic Migration Patterns using *Matrículas Consulares**

Maria Esther Caballero†
Carnegie Mellon University

Brian C. Cadena‡
University of Colorado - Boulder and IZA

Brian K. Kovak§
Carnegie Mellon University NBER and IZA

October 2017

Abstract

In this paper, we show how to use administrative data from the *Matrícula Consular de Alta Seguridad* (MCAS) identification card program to measure the joint distribution of sending and receiving locations for migrants from Mexico to the U.S. While other data sources cover only a small fraction of source or destination locations or include only very coarse geographic information, the MCAS data provide complete geographic coverage of both countries, detailed information on migrants' sources and destinations, and a very large sample size. We first confirm the quality and representativeness of the MCAS data by comparing them with well-known household surveys in Mexico and the U.S., finding strong agreement on the migrant location distributions available across datasets. We then document substantial differences in the mix of destinations for migrants from different places within the same source state, demonstrating the importance of detailed sub-state geographical information. We conclude with an example of how these detailed data can be used to study the effects of destination-specific conditions on migration patterns. We find that an Arizona law reducing employment opportunities for unauthorized migrants decreased emigration from and increased return migration to Mexican source regions with strong initial ties to Arizona.

Keywords: international migration, immigration law, Mexico, United States

---

*This project was supported by a grant from the Berkman Faculty Development Fund at Carnegie Mellon University. The authors would like to thank Alexandra Chouldechova, Terra McKinnish, Fernando Riosmena, and Lowell Taylor, and participants at the 2017 Population Association of America Annual Meeting and the University of Oxford Workshop on Immigration, Health, and Well-Being for helpful comments. Benjamin Mayer provided excellent research assistance. Special thanks to Edith Soto Ramírez at the *Instituto de los Mexicanos en el Exterior* for extensive discussions regarding the *Matrícula Consular de Alta Seguridad* data. Remaining errors are our own.

†Carnegie Mellon, Heinz College, 4800 Forbes Ave., Pittsburgh, PA 15213; mcaballe@andrew.cmu.edu

‡Univ. of Colorado, Dept. of Economics, 256 UCB, Boulder, CO 80309; brian.cadena@colorado.edu; 303-492-7908

§Carnegie Mellon, Heinz College, 4800 Forbes Ave., Pittsburgh, PA 15213; bkovak@cmu.edu; 412-268-5223

1 Introduction

Research on immigration is often hampered by data limitations. For example, large scale individual-level surveys in the U.S. do not ask about immigrants' legal status, and government records on legal permanent residents are presented as aggregate tabulations with no individual-level information. As emphasized by Blau and Mackie (2016), the strategic utilization of alternative administrative datasets can help fill such gaps, facilitating innovative research questions. In this paper, we follow their suggested approach by examining administrative data providing uniquely detailed information on sources and destinations for migrants from Mexico to the U.S. Once we establish the data's appropriateness, we provide an initial example of the type of novel research that is possible with these data, documenting the international migration consequences of the Legal Arizona Workers Act.

Specifically, we evaluate data on geographic migration patterns from the *Matrícula Consular de Alta Seguridad* (MCAS) program, which issues identity cards to Mexican citizens living in the United States. Massey, Rugh and Pren (2010) introduce and describe this data source, which features complete geographic coverage of Mexico and the U.S., detailed information on migrants' source and destination regions, and very large sample sizes. The more than 7 million observations spanning 2006 to 2013 separately identify 75 U.S. destinations and all of the more than 2,000 source *municipios* in Mexico. However, the data's primary disadvantage is that they represent administrative records from a voluntary program, rather than a stratified random sample from a well-defined population, raising concerns about data quality, coverage, and representativeness (Riosmena and Massey 2012). Our first contribution is to resolve these concerns by showing strong agreement on migrant source and destination distributions between MCAS and a variety of standard data sources, including nationally representative household surveys in Mexico and in the U.S. These comparisons establish the quality and representativeness of the MCAS data and confirm its usefulness as a source of information on detailed geographic migration patterns.

For research questions requiring measures of sub-national migration patterns with broad geographic coverage and/or large sample sizes, MCAS data are uniquely well suited. Large-scale household surveys such as the Mexican Census and the American Community Surveys record sub-national geography only for their respective countries, with only national geographic information for foreign places. These surveys therefore cannot be used to measure migration flows between sub-national locations. More specialized surveys such as the *Encuesta Nacional de la Dinámica Demográfica* (ENADID) and the *Encuesta sobre Migración en la Frontera Norte* (EMIF) report migrants' sources and destinations, but suffer from small samples and quite aggregate geographic information.[1] The Mexican Migration Project (MMP) provides unparalleled detail regarding migration experiences for those surveyed, but this impressive detail comes at the cost of covering only a small number of communities in Mexico. (Massey and Zenteno 2000).[2]

Our second contribution is to demonstrate the empirical benefit of calculating geographic migration patterns for detailed source locations. While the ENADID and EMIF report migrants' sources at the relatively aggregate state level, the MCAS data provide source information at the much more detailed *municipio* level. We calculate the distribution of destinations for all migrants from a given Mexican state and compare it to the destination distributions for migrants from each *municipio* within that state. We find that, as a general rule, the state-level distribution differs substantially from the *municipio*-level measures. In fact, the typical source *municipio's* destination distribution differs from its state's distribution by as much as the typical state differs from the destination distribution of all Mexican emigrants. Thus, assigning all migrants their source state's average destination distribution introduces substantial measurement error into an analysis of the role of pre-existing local migration patterns in an individual's migration experience.

We anticipate that these data will open the door to numerous additional lines of research, especially in the literature focused on the influence of prior international migrants' destination choices on the experience of subsequent migrants. As it stands, this literature already contains a number of important findings. Larger numbers of previous migrants from the same sending community increase the probability of migrating internationally by lowering the costs of migration.[3] Previous

[1] In the case of EMIF, the data record the planned destinations of those intending to migrate.

[2] As of October 2015, the MMP had surveyed 154 communities whose combined populations accounted for 1.03 percent of the Mexican population in 2010.

[3] See, for example, Massey (1986); Massey and Espinosa (1997); Palloni, Massey, Ceballos, Espinosa and Spittel

migration has a wide variety of other effects, including altering the set of individuals choosing to migrate, affecting migrants' approach to crossing the border, and increasing investment in origin communities.[4] Similarly, migrants' destinations and eventual success in the U.S. are strongly influenced by the destinations and occupations of previous migrants from their sending community.[5] Garip (2016) provides a detailed typology of underlying mechanisms that influence these empirical relationships, including social facilitation, normative influence, and network externalities, with many examples pertaining to the context of Mexico–U.S. migration.

In addition to concerns about data quality, one reason researchers have likely avoided using the MCAS data is because the tabulations do not provide individual-level information other than place of birth and U.S. residence. The final contribution of this paper, therefore, is to demonstrate that the MCAS data can nevertheless be used in combination with traditional household survey data to address important questions related to Mexico-U.S. migration. As an initial example, we study the international migration response to the Legal Arizona Workers Act (LAWA). This law required employers throughout Arizona to submit an electronic request to confirm every prospective employee's legal authorization to work in the U.S. The law thus reduced the attractiveness of Arizona as a destination for potential migrants without legal status. Prior work has shown that restrictions like these reduced the local immigrant population (Bohn, Lofstrom and Raphael 2014) and decreased planned border crossings with Arizona as the intended destination (Hoekstra and Orozco-Aleman 2017).

We use the MCAS migration measure to analyze the effects of this policy on sending communities within Mexico. We begin by calculating the share of migrants from each source region (state or *municipio*) that selected Arizona as a destination prior to LAWA's passage. We then use Mexican

(2001); Winters, de Janvry and Sadoulet (2001); and Garip and Asad (2016). DiMaggio and Garip (2012) provide a survey of the sociology literature on networks.

[4]McKenzie and Rapoport (2010) show that the presence of previous migrants disproportionately increases migration probability for less educated individuals. Dolfin and Genicot (2010) examine the effects of family and community contacts on migration with and without the assistance of smugglers. Woodruff and Zenteno (2007) show that migration increases microenterprise development in source locations.

[5]See Bartel (1989); Jaeger (2000); Bauer, Epstein and Gang (2002); Diaz McConnell (2008); Patel and Vella (2013); and Lafortune and Tessada (2014) on destination choices and Munshi (2003) and Mundra and Rios-Avila (2016) on labor market success

Census data to determine how changes in regional migration flows depend on the importance of Arizona in a source region's initial distribution of migrant destinations. The results reveal that source areas with stronger pre-LAWA migration connections to Arizona experienced larger decreases in emigration to the U.S. and larger increases in return migration from the U.S. than sources with initially weaker connections to Arizona.

This example further validates the quality of the MCAS-based measure of migration patterns and confirms its value to researchers in a variety of ways. First, the fact that we find differential migration responses based on a source's MCAS-measured connection to Arizona directly reinforces our conclusion that these migration connection measures are informative. Second, this analysis shows how the data can be used to expand the scope of questions that can be answered. No previous analysis of LAWA or similar restrictions has examined migration responses at the source level because the policy change affected job prospects for all potential migrants throughout Mexico to some degree. The place-to-place migration measure allows us to form hypotheses about which sending communities are most affected by LAWA. This approach allows us to show that not only did migrants shift away from Arizona as a destination, but the policy also led to an overall decline in net migration to the U.S. from connected source regions. Third, we demonstrate the value of the geographic detail in the MCAS data by comparing the robustness of state-level analysis to *municipio*-level analysis. Not only does the *municipio*-level analysis yield greater precision and less influence of high-leverage outliers, but it also allows us to control for any time-varying unobservable push or pull factors common to *municipios* within the same Mexican state.

The remainder of the paper is organized as follows. The next section provides background on the MCAS program and administrative data. Section 3 provides comparisons between the MCAS and multiple U.S. and Mexican data sources, showing that MCAS-based calculations closely match available geographical migrant distributions from high quality household surveys. Section 4 demonstrates the value of measuring sub-national migration patterns with detailed information on source locations, revealing substantial differences between *municipio*-level and state-level distributions of migrant destinations. Section 5 presents the analysis of LAWA, showing that sources more connected to Arizona experienced larger decreases in net migration to the U.S. after the law's passage. Section 6 concludes.

2 *Matrícula Consular de Alta Seguridad (MCAS)*

2.1 Background

The *matrícula consular* is a document issued by the Mexican government that provides its citizens abroad with a form of identification in their country of residence. In the U.S., the *matrícula* provides proof of citizenship, identity, and residence without conferring any immigration status on the cardholder. It is used primarily for returning to Mexico, opening bank accounts, obtaining loans and home mortgages, as identification to law enforcement officials, and, in some U.S. states and cities, to obtain a driver's license and to access basic government services. The many benefits of *matrículas* may explain both the high takeup rates among unauthorized immigrants and broad representativeness of the *matrícula*-holding population that we document below.

The Mexican government has been issuing *matrículas* since 1871, but in the 1990s they were transformed into wallet-sized laminated cards resembling a driver's license. In 2002 a more secure version called the *Matrícula Consular de Alta Seguridad* (MCAS) was introduced (IME 2004), and additional security measures were added in 2006. Since the most recent security update, all MCAS issuances are recorded and verified through a centralized database, accessed by the issuing consulate (National Immigration Law Center 2015). This database of MCAS issuances forms the basis for the dataset we use to measure Mexico-U.S. migration patterns.

In order to obtain a card, the applicant must make an appointment and attend the Mexican consulate corresponding to their place of residence in the U.S. The applicant must provide proof of Mexican citizenship, identity, and residence in the relevant consular area, and they must not have a criminal record or be subject to judicial or administrative actions in the U.S. or Mexico (Secretaría de Relaciones Exteriores 2016). Cards are issued to all qualifying Mexican citizen applicants irrespective of age or immigration status, though it is generally assumed that the majority of MCAS holders are unauthorized immigrants who have limited access to other official forms of identification in the U.S. (Massey et al. 2010). The card is valid for 5 years, and it can be renewed when it expires, when the cardholder moves to another consular jurisdiction, or if it is lost or damaged. MCAS are considered valid proof of identification by a wide variety of financial institutions and police departments in the U.S., and 12 states and the District of Columbia accept them as proof of ID to obtain a driver's license (National Conference of State Legislatures 2015).[6]

2.2 Data

The recorded information from each approved MCAS application includes each card recipient's *municipio* (similar to county) of birth in Mexico and the U.S. state and consular area of current residence. Consular areas refer to the geographic area of the U.S. within the jurisdiction of each Mexican consulate. The governmental Institute for Mexicans Abroad (*Instituto para los Mexicanos en el Exterior*, IME) uses this database to produce publicly available tabulations of the numbers of cards issued in each year.[7] These tabulations include the count of cards issued for each birth *municipio* and U.S. location pair, omitting any additional individual-level information that might raise confidentiality concerns.[8] By combining published tabulations based on U.S. state of residence with separate tabulations based on consular area of residence, one can generate counts of card issuances to individuals living in the 75 mutually exclusive and exhaustive destination areas in the U.S., shown in Appendix Figure A-1. With more than 2,000 Mexican *municipios* and 75 U.S. destinations represented, the card issuance data provide very detailed information on point-to-point migration patterns between Mexico and the U.S.

A large number of MCAS are issued each year, and more than 7 million cards were issued during

[6]These 13 jurisdictions are CA, CO, CT, DC, DE, HI, IL, MD, NM, NV, UT, VT, and WA. It is likely that the take-up rate is somewhat higher in these destinations compared to the rest of the country. Our analysis focuses on differences in destinations selected by migrants from different sources. As long as any higher take-up in these destinations occurs for individuals from all sources, our central conclusions will be unaffected.

[7]As of October 2017, the tabulations are available online at `http://www.ime.gob.mx/gob/estadisticas/2016/usa/estadisticas_usa.html`.

[8]The tabulations are similar in structure to those provided by the U.S. Internal Revenue Service, reporting counts of migrants for each source-destination pair in the U.S. See Kaplan and Schulhofer-Wohl (2012) for an example.

the 2006-2013 period. Nearly all of these issuances are new applications rather than renewals.[9] To assess the scale of these numbers, we calculate the stock of valid cards outstanding in each year from 2011 to 2013. Since MCAS are valid for 5 years, we can measure the stock of valid cards in a given year by summing the numbers of cards issued during the preceding 5 years. For example, all cards issued from 2006-2010 are still valid during at least a portion of 2011. Appendix Table A-1 compares the number of valid cards in 2011-2013 to the estimated Mexican-born population of the U.S. (calculated from the ACS) and the Pew Research Center's estimates of the unauthorized Mexican-born population of the U.S. (Passell and Cohn 2014, Gonzalez-Barrera 2015). We find a quite consistent 38 percent share of Mexican-born population holding a valid MCAS in each year. This share is similar to the 46 percent share reported in Suro and Escobar (2006), and the difference may reflect either a modest decrease in takeup between 2006 and 2011 or the fact that we observe the population of cards issued rather than a sample. Massey et al. (2010) conclude that it is safe to assume that all *matrícula* holders are unauthorized immigrants, since "persons legally in the United States would have no need for such documentation" (p.132). Under this assumption, the MCAS data cover 75 to 80 percent of the unauthorized Mexican immigrants living in the U.S.

Nevertheless, applying for a MCAS is voluntary, and the cards are distributed to a self-selected population. To understand the potential selection into takeup, we analyzed data from another Pew survey, which interviewed individuals applying for *matrículas* at various Mexican consulates in 2004-2005 (Suro 2005).[10] Appendix Table A-2 shows mean demographic and educational characteristics for this sample of *matrícula* applicants in comparison to all Mexican-born U.S. residents in the 2005 American Community Survey. Men, younger adults, and those with lower educational attainment are overrepresented among *matrícula* applicants in comparison to the overall Mexican-born population. Additionally, *matrícula* applicants were more likely to have arrived in the late 1990s and early 2000s compared to the overall Mexican-born population. Together, these results are consistent with the idea that *matrícula* applicants are primarily drawn from the population of

[9] Appendix Table A-3 provides an annual breakdown of the number of card issuances. In addition, officials at IME were kind enough to provide annual summary statistics on the share of card issuances reflecting new cards vs. renewals. In every year, fewer than 3 percent are renewals.

[10] The survey covered consulates in Los Angeles, New York, Chicago, Atlanta, Dallas, Raleigh, and Fresno.

recently arrived unauthorized immigrants who are most likely to benefit from having access to an official identification card under the MCAS program.

3 MCAS Data Quality

The preceding results suggest that the MCAS data provide very good coverage of the recently-arrived unauthorized immigrant population. In this section, we examine the quality of the data in measuring the strength of migration ties between sources in Mexico and destinations in the U.S., while contrasting the MCAS data with other nationally representative data sources that one might use as alternatives.[11] We use a variety of data sources, and Table 1 provides a reference guide for which data sources, time periods, and migration measures are used in each set of results.

3.1 MCAS Data Match High Quality Survey Data

We begin by comparing the migration patterns measured using MCAS to those in the largest and highest quality household surveys in the U.S. and Mexico. With these datasets, we are able to compare estimates of the marginal distributions, i.e. source locations in Mexico and destination locations in the U.S.[12]

Figure 1 provides the first such comparison, showing strong agreement between MCAS data and American Community Survey (ACS) data regarding the distribution of Mexicans across U.S. destination states. To construct this figure, we use the MCAS cards issued from 2006-2010, all of which were still valid on January 1, 2011. We then calculate the share of these cards reporting a migrant's residence in each of the 50 U.S. states plus the District of Columbia at the time of

[11]We do not make comparisons between MCAS and the Mexican Migration Project (MMP) data because MMP communities are generally far smaller than and not necessarily representative of their *municipios*, so there is no comparable geographic unit across the two data sets.

[12]Throughout this analysis, we use the U.S. state as the destination-level geography. Much of the analysis relies on the public-use microdata version of the ACS (Ruggles, Alexander, Genadek, Goeken, Schroeder and Sobek 2010), and the sub-state geographical definitions in that data source (PUMAs) do not align well with the consular areas. For consistency, therefore, we adopt the US state as the definition of a destination. The consular areas, however, are composed of US counties, and future work can match US datasets with county-level coverage to the destination geography shown in Figure A-1.

application. We construct a similar measure of the share of Mexican-born residents living in each state using the 2010 and 2011 American Community Surveys. Because the ACS is conducted throughout the year, by combining the 2010 and 2011 samples, we obtain a measure centered on January 1, 2011. We then compare the two distributions using a scatter plot, with the MCAS-based shares on the y-axis and the ACS-based shares on the x-axis. Since Mexican population is distributed unevenly across U.S. states, we plot the natural log of the state shares, allowing one to visually compare the two data sources for large and small states on the same figure.[13] We include a 45-degree line, which shows how the two sets of shares would relate if the datasets agreed perfectly. It is readily apparent that the two datasets strongly agree, with only minor deviations from the 45-degree line.[14] Moreover, the largest differences appear in states with very small numbers of Mexican immigrants. We have created similar figures for each of the time periods available in the data, centered on the beginning of 2012, 2013, and 2014, and they show similar agreement between the two data sources.

Figure 2 Panel (a) examines the distribution of Mexican source states for migrants to the U.S. We use the 2010 Mexican Census for this comparison, taking advantage of a question that asks respondents whether anyone currently or formerly living in the household migrated internationally between June 2005 and June 2010. Since we know the household's location, we can calculate each Mexican state's share of individuals observed leaving for the U.S. during this time period. We compare the log of this share with similar source-state shares from the MCAS data covering 2006-2010. Again, the close agreement across datasets is visually apparent from the figure, with only minor deviations from the 45-degree line.

Figure 2 Panel (b) repeats the analysis of Panel (a) but changes the geographic definition to the Mexican *municipio*. Although these measures are somewhat noisier, especially in municipios with smaller populations, the agreement between MCAS and the Mexican Census at this fine level of geographic detail is remarkable and reflects the very large number of migrants present in the MCAS

[13]We report the R^2 for both the logged and non-logged version of these comparisons in each figure.

[14]Because the shares sum to one across states in each dataset, states whose shares are larger in MCAS than in the ACS are offset by states whose MCAS shares are smaller than in the ACS. Because the log function is concave, states with larger shares in the MCAS appear closer to the 45-degree line than states with smaller shares in the MCAS.

data, which facilitates high quality measures of migration patterns even for small geographic areas. These results show that the MCAS data closely coincide with the best available measures of source and destination information for Mexican migrants to the U.S.

3.2 ENADID and EMIF as alternatives

We next consider whether either of the other two datasets with complete geographic coverage of sending and receiving locations compares as well with these Census distributions. We begin with the *Encuesta Nacional de la Dinámica Demográfica* (ENADID), which reports each migrant's state of residence in the U.S. in addition to their source location. The question identifying migrants is similar to the one in the Mexican Census, and it allows us to observe migrants who left for the U.S. between August 2009 and August 2014. Figure 3 Panel (a) is constructed analogously to Figure 1, and compares the distribution of destination states among migrants in the ENADID who left during 2009-2014 to the 2014 ACS. The ENADID destination measure performs reasonably well, but there are important ways in which the MCAS data in Figure 1 align more closely with the baseline ACS distribution. First, Figure 3 Panel (a) includes only 41 U.S. states, as there are ten U.S. states that were not reported as the destination for any migrants observed in the ENADID. Second, the figure shows that the distribution aligns closely for large-population states such as Texas and California, but there is noticeably more disagreement between the two data sources for mid-range population states (those with log shares between -4 and -6). The MCAS data matched very closely in this range; larger differences from the ACS were observed only for very small states that the ENADID omits entirely. These differences highlight the primary shortcoming of the ENADID: a much smaller sample of migrants than in the MCAS database.[15]

Figure 3 Panel (b) provides a comparison of source states analogous to Figure 2 Panel (a), using the ENADID in place of the MCAS. Again, the ENADID performs reasonably well, although the share comparisons are not as tightly clustered around the 45-degree line as are the observations

[15]The different number of covered migrants is not surprising as the sample size for the ENADID is calibrated to ensure accurate reporting of domestic fertility rates rather than migration rates. We thank Fernando Riosmena for helpful discussion on the design of the ENADID.

using the MCAS data, likely resulting from the ENADID's smaller sample size. Appendix Figures A-2 and A-3 present analogous comparisons using the EMIF. The EMIF performs no better than the ENADID data, and the EMIF destination distribution is much less consistent with the ACS.[16] Because the ENADID appears to be the best survey option, we focus our remaining comparisons on the MCAS and ENADID data.

3.3 Comparison of Joint Distributions in MCAS and ENADID

We next turn to a comparison of the joint distributions available in both the MCAS and ENADID. Our analysis focuses on each dataset's measure of where migrants from a given Mexican state are likely to locate within the United States. Using MCAS applications from 2009-2013 and ENADID data from 2014, we calculate for each source state the share of migrants selecting each U.S. state as their place of residence. Figure 4 graphs the relationship between the natural log of these shares for each source-destination combination, labeling each observation with the U.S. destination state. Although there is some agreement between these two data sources, the influence of the ENADID's smaller sample size is readily apparent. Notably, there are relatively few observations in the lower left quadrant of the figure. These "missing" observations reflect destinations that are relatively uncommon in the ENADID or that fail to appear at all due to its small sample size.

In fact, the sample sizes in the ENADID are sufficiently small that empty cells are guaranteed to occur. Mexican states with fewer than 51 observed out-migrants must have at least some empty source-destination cells. As shown in detail in Appendix Figure A-4, nearly 80 percent of Mexican source states have fewer than 51 observed migrants in the ENADID. In contrast, every Mexican state has at least 51 individual MCAS issuances observed over the five-year period, and the vast majority of states have more than 100,000 observed migrants. Even at the relatively aggregate state level, the ENADID simply does not observe a sufficient number of migrants to credibly estimate source-specific destination distributions for the majority of Mexican sources.

Taken together, the results in this section imply that the MCAS data provide an excellent

[16] This discrepancy likely occurs because the EMIF asks about a migrant's *intended* destination, which is subject to change.

way of measuring place-based migration patterns between Mexico and the U.S. The distributions of both sources and destinations closely match the highest quality available survey datasets, and the joint distribution of sources and destinations corresponds reasonably well with the distribution observed in the ENADID, despite the latter data source's small sample size. Finally, the number of observations per cell is orders of magnitude larger in the MCAS data, which both increases the precision of the estimated share of migrants from a source choosing a particular destination and also greatly reduces the potential for entirely missing source-destination pairs with small numbers of migrants.

4 The Value of Detailed Source Data

Having established that MCAS data are a superior resource for examining Mexico-U.S. migration patterns, we next examine the importance of using fine rather than coarse geography. Our primary analytical tool is a measure of the dissimilarity of two discrete distributions, as described by Duncan (1957). This index quantifies the difference in the destination distribution between individuals from source s and individuals from a reference group.

$$\Delta_s \equiv \frac{1}{2} \sum_{d=1}^{D} \left| \pi_{sd} - \pi_d^{ref} \right|, \tag{1}$$

where $\pi_{sd} \equiv \frac{N_{sd}}{N_s}$ is the share of emigrants from source s residing in destination d, and π_d^{ref} is the share of the reference population in destination d. This index is bounded between zero and one, with zero representing identical distributions and one representing distributions with no overlap whatsoever. The magnitude of the measure can be interpreted as the share of migrants from source s who would need to be reallocated across destinations in order to exactly match the destination distribution of the reference group. Note that Figures 1 through 3 report this dissimilarity measure for the comparisons reflected in each figure.

We begin by using the MCAS data from 2006-2010 to document the variation in chosen U.S. locations based on migrants' state of birth. Specifically, we examine how each Mexican state's distribution of migrant destinations in the U.S. compares to the destination distribution of all Mexican emigrants. In this case, s in (1) refers to Mexican states, d refers to U.S. destination states, and ref refers to all Mexican migrants in the U.S. In order to clarify how this measure is calculated, Appendix A.6 steps through the calculation of $\pi_{sd} - \pi_d^{ref}$ for migrants from Michoacán, relative to all Mexican migrants. Compared to the average Mexican migrant, migrants from Michoacán are more likely to live on the West Coast and in Illinois, and they are much less likely to live in Texas. Adding up the absolute value of these differences yields a value of $\Delta_s = 0.21$, indicating that 21 percent of the migrants from Michoacán would need to relocate within the U.S. in order to match the overall distribution of Mexican migrants' chosen destinations. We repeat this analysis for each sending state, and Figure 5 Panel (a) provides a histogram showing the distribution of Δ_s for all Mexican source states. The values of the index range from 0.09 to 0.63, with a mean of 0.31. Thus, most states' measures are reasonably high, indicating that the various Mexican source states send migrants to quite different sets of locations in the U.S.

The histogram with solid bars in Figure 5 Panel (b) provides the distribution of the same dissimilarity measure calculated at the *municipio* rather than state level, and it shows that the destination distributions for source municipios are even more different than are the distributions for source states. Although a few municipios have destination distributions close to the national average (the minimum dissimilarity measure is 0.08), most are quite different (the mean dissimilarity measure is 0.46 and the maximum is 0.99). Many source municipios would require more than half of their migrants to choose different destinations in order to match the destination distribution of all Mexican migrants.

We next examine variation in migrant destinations among *municipios* in the same sending state. For each *municipio*, we calculate a new version of Δ_s using the Mexican state containing the source *municipio* as the reference group. The histogram with hollow bars in Figure 5 Panel (b) provides the distribution of these within-state dissimilarity measures.[17] This distribution is noticeably shifted to left compared to the solid bars, which confirms that the destinations selected by migrants from

[17] For more detail on the data points underlying this histogram, see Appendix A.6, which includes example *municipio* distributions and the relevant comparisons with those *municipios'* state.

a given *municipio* are, in general, more similar to their state's distribution than they are to the national average. Yet the histogram reveals that it is very common for *municipios* within the same state to have very different destination distributions – the values of this version of the dissimilarity index range from 0.03 to 0.99 with a mean of 0.34. Thus, for many *municipios*, the state distribution is a poor proxy for the true *municipio*-level distribution.[18]

As an example, consider two *municipios* in the state of Michoacán: Ciudad Hidalgo and Tiquicheo. These source locations are only a 3 hour drive apart, but their destination distributions differ sharply. Migrants from Hidalgo settle primarily in Illinois (likely Chicago), while more than two thirds of emigrants from Tiquicheo reside in Texas.[19] This difference in destinations occurs *within* the same source state, ruling out the possibility that it arises due to other factors affecting destination choice such as distance, climate similarity, etc. It is also noteworthy that neither distribution is particularly close to the state level distribution, which is more concentrated in California.

The results in this section have important implications for researchers using previous settlement patterns as a source of identifying variation. State of birth destination measures are, in general, a poor proxy for settlement patterns that operate at a finer level of geography. In fact, the typical *municipio's* destination distribution is as different from its state's as the typical state's distribution is different from the overall distribution. The ability to construct a measure of the destinations chosen by previous migrants from an individual's *municipio* of birth, therefore, makes the MCAS data a particularly valuable resource for researchers studying Mexico-U.S. migration.

[18]Random variation could account for some observed differences between *municipio*-level and state-level destination distributions, particularly since some *municipios* are quite small. In Appendix A.7, we implement a permutation test to demonstrate that the observed differences are far larger than those that could plausibly be explained by small *municipio* samples and random variation.

[19]For the full distribution of destinations chosen by migrants from these two sources, see Appendix Figures A-8 and A-9.

5 Effects of LAWA on Migration in Connected Mexican Sending Regions

As an example of the value of this level of geographic detail, this section examines the effect of the Legal Arizona Workers Act (LAWA) on migration rates into and out of sending regions in Mexico. This act, passed by the Arizona state legislature in 2007 with an effective date of January 2008, led to a decline in the "likely unauthorized" population living in Arizona (Bohn, Lofstrom and Raphael 2011). Data limitations, however, prevented an analysis of whether this decline reflected changes in international migration or simply the movement of unauthorized immigrants into different U.S. states. Our analysis leverages the MCAS data to show that emigration fell and return migration rose in source locations that were more exposed to LAWA through migration connections. Thus, the decline in destination labor market opportunity driven by LAWA reduced the number of Mexican immigrants living in the United States from regions with ties to Arizona.

5.1 The Legal Arizona Workers Act

LAWA mandates the use of E-verify, an online system administered by the federal government, to verify prospective employees' identity and authorization to work in the United States. The E-verify system compares Social Security numbers and names of new workers against a centralized database from the Social Security Administration (SSA) and the Department of Homeland Security (DHS). When there is no match between the employee's name or social security number and the official records, the system sends a report of non-confirmation to the employer. The law imposes sanctions on employers who hire unauthorized workers, ranging from business license suspensions for the first offense to license revocation (Bohn et al. 2011).

LAWA's main purpose was to increase the costs for employers hiring unauthorized migrants and for unauthorized employees looking for jobs. As a result, Arizona became a less desirable destination to live and work for immigrants without legal status. We examine the effect of LAWA on *international* migration by assessing whether Mexican sending regions initially more connected to Arizona experienced larger increases in return migration and larger decreases in emigration rates after LAWA was passed. Importantly, these analyses use Mexican Census data to measure migration rates in and out of Mexico; the MCAS data are used only to characterize the degree to which a sending location was initially connected to Arizona.

5.2 Effects of LAWA on Migration Rates

Our analysis treats the implementation of LAWA as a quasi-experiment that manipulates Mexican-born individuals' U.S. job prospects. We expect that LAWA negatively affects the job prospects for all Mexican-born individuals without U.S. work authorization, but that these effects will be larger for individuals with strong network connections to Arizona. Because we cannot observe an individual's social network contacts directly, we rely on the geography-based networks observable in MCAS data. Specifically, we assume that LAWA's effect on the labor market prospects in the U.S. for the average migrant from region s is proportional to the share of MCAS card recipients from the same sending region who lived in Arizona in 2006.

We use the following specification to relate changes in migration rates from 2005 to 2010 to the importance of Arizona as a destination:

$$\Delta Y_s = \beta_0 + \beta_1 \pi_{s,2006} + \epsilon_s \tag{2}$$

where $\Delta Y_s = \ln(y_{s,2010}) - \ln(y_{s,2005})$ is the change in the log of the return migration or emigration rate from 2005 to 2010 in Mexican sending region s. We calculate migration rates using the 2010 Mexican Census and the 2005 *Conteo de Población y Vivienda*. $\pi_{s,2006}$ is the share of emigrants from source s selecting Arizona as a residence in 2006, before LAWA was passed, using the MCAS data. ϵ_s is an error term. $\hat{\beta}_1$ therefore captures the differential change in migration rates for Mexican source locations that were more connected to Arizona prior to LAWA. We conduct this analysis using states or *municipios* as source regions, treating states as independent observations and computing standard errors clustered at the state level when using *municipios* as the unit of analysis.[20]

5.2.1 Return Migration Rates

The first specification uses return migration rates as the outcome variable. Return migrants are defined as individuals living in Mexico during the 2010 Census or 2005 *Conteo* reference period, but who lived in the U.S. five years before. The return migration rate is then the number of return migrants divided by the source's population at the beginning of each period (2000 or 2005).[21]

Figure 6 Panel (a) presents the underlying data and the fitted values for equation (2) using the change in the natural log of this measure from 2005 to 2010 as the dependent variable. Consistent with expectations, return migration rates rose more in sending states with stronger network connections to Arizona. The first column of Panel A of Table 2 provides the coefficient estimates from this specification. To understand the magnitude of this effect, it is important to bear in mind that this time period saw dramatic declines in net migration to the United States overall, both through increased return migration and through decreased emigration. The average Mexican state's return migration rate nearly quadrupled over this time period (rates were 0.3% on average in 2005 and 1.13 % on average in 2010). Yet these results imply that Mexican states with strong connections to Arizona saw larger increases in return migration than in states with weaker connections. A state like Sonora, where roughly 50 percent of migrants had historically settled in Arizona, experienced 30 percent larger growth in its return migration rate compared to a state with no connection to Arizona.[22]

Columns (2) and (3) of Table 2 examine the robustness of this result. The second column in

[20] In Appendix A.8, we corroborate the regression analysis following (2) with summary statistics on the time-series evolution of return migration and emigration rates for *municipios* with initially higher and lower rates of connection to Arizona.

[21] Return migration flows are identified in the 2005 *Conteo* and in the 2010 Mexican Census through a question that records country of residency five years prior to the date when the survey was administered. The count of return migrants does not include any individuals who were living in Mexico five years previously, but who moved to the U.S. and back within the five year window.

[22] The percentage difference in the growth of return migration rates would be $100 * e^{.5*.558} - 1 = 32.1\%$. Compared to growth rates of roughly 300 percent in untreated states, this treated state would see a growth rate of roughly 400 percent.

Panel A shows results when the observations are weighted by the 2000 Mexican population.[23] The third column provides results from a robust regression technique that reduces the impact of high leverage outliers.[24] The point estimates are positive and quite similar across all three specifications, but the large standard error in column (3) makes clear that these state-level results are highly dependent on the inclusion of a single high-leverage observation (Sonora).

We now take advantage of the finer geographical detail available in the MCAS to conduct similar analysis using the Mexican *municipio* as the unit of analysis. Panel B of Table 2 presents similar estimates for the change in log return migration rates from 2005 to 2010 at the Mexican *municipio* level. The first column provides the baseline estimates using *municipios* as observations, and Figure 6 Panel (b) provides the raw data and fitted line from this regression. The point estimate is comparable in magnitude with the first column of Panel A. Notably, the standard error is substantially smaller, and the scatter plot makes clear that there is no particular high leverage outlier driving the results. In fact, implementing the same robust regression technique used in the state-level analysis (column 3) leads to standard errors that are not much different from those in the baseline results in column (1).

An additional advantage of conducting a *municipio*-level analysis is the ability to add controls for other changes over time that affect return migration rates. Changing conditions in sending regions will alter individuals' incentives to return or to leave for the U.S. To the extent that these changes in conditions are correlated with a location's ties to Arizona, they represent a threat to the causal interpretation of these regression results. In columns (4)-(6) we therefore add Mexican state fixed effects as additional controls. Because the dependent variable is already expressed as a difference within *municipios*, these fixed effects remove the influence of any *changes* in the sending areas that affected migration rates similarly for all *municipios* within a Mexican state. For example, this specification controls for changes in state-level labor market conditions and changes in state-level crime. This specification leverages the within-state variation in destinations shown in Figure

[23]This weighting addresses the fact that population growth rates are heteroskedastic, with smaller populations experiencing more variable percentage growth in migration rates.

[24]Specifically, we use the `rreg` command in Stata, which implements the robust regression procedure described by Li (1985).

5b and compares municipios that are geographically close to each other and yet are differentially connected to Arizona's labor market. In each case, the results in columns (4)-(6) are similar to the corresponding results without fixed effects in columns (1)-(3). The ability to conduct analysis at the *municipio* level thus strengthens the results by increasing precision, by decreasing the importance of outliers, and by allowing for flexible controls for unobserved changes that could be correlated with the strength of a location's ties to Arizona.

5.2.2 Emigration Rates

Having shown that LAWA increased the rate at which individuals return to Mexico, we now turn to the other component of net migration – emigration to the United States. Emigration is measured using the 2010 Census, which reports the year in which household members traveled to the U.S. We calculate the emigration rate as the number of people who reported emigrating in a given year divided by the source-area population in that year, for 2005 and 2010.[25]

Table 3 is analogous to Table 2 and examines changes in the log of the emigration rate from 2005 to 2010 at the Mexican state and *municipio* levels. The scatter plot and fitted line for the regression in the first column of Panel A is provided in Figure 7 Panel (a); Figure 7 Panel (b) provides a similar graph for the specification in the first column of Panel B. On the whole, the results for emigration mirror the results for return migration, with sources more connected to Arizona seeing larger decreases in emigration from 2005 to 2010. Again, the importance of a single observation (Sonora) in the state-level results is apparent both visually in the scatter plot and in the large standard error in the third column of Panel A. In contrast, the *municipio*-level results are robust to mitigating the influence of high leverage outliers (columns 3 and 6 of Panel B). Further, the results are robust to the addition of state fixed effects (columns 4-6 of Panel B), and are in fact stronger when including these controls.

Together, these results demonstrate that the LAWA-induced declines in Arizona's likely-unauthorized

[25]A small number of emigrants who left in 2005 are not observable because the survey asks only about the start date of the most recent trip, so individuals who emigrated first in 2005, then returned to Mexico and then emigrated again are treated as having emigrated in the year of their most recent trip to the U.S.

population documented by Bohn et al. (2014) occurred partly through substantial changes in international migration patterns. Return migration rose more sharply in Mexican source locations where migrants had historically settled in Arizona, and emigration fell more in those same locations. These effects of LAWA are, to our knowledge, a novel result, and they imply that policies affecting migrants' job opportunities can be effective at deterring or reversing the flow of unauthorized migrants into the United States.[26] Moreover, these results further strengthen our conclusion that source-specific destination distributions based on MCAS data are informative. Had the data been sharply non-representative or highly noisy, we would not observe these important differences across Mexican source regions with strong initial connections to Arizona.

6 Conclusion

In this paper, we evaluate the use of administrative data from the *Matrícula Consular de Alta Seguridad* (MCAS) program to measure geographic migration patterns between Mexico and the U.S. Unlike other available datasets that one could use to characterize these patterns, the MCAS data provide very large sample sizes, detailed geographic identifiers, and complete geographic coverage of both Mexico and the U.S. We find strong agreement between the MCAS and each country's Census of Population when measuring the distributions of migrants' Mexican source states and U.S. destination states. We then demonstrate that different sending regions within the same Mexican state regularly send migrants to very different sets of locations in the U.S. In fact, the typical *municipio's* distribution matches its state's distribution only as well as the typical state's matches the overall destination distribution. Using more aggregate state-level migration information would therefore obscure the differences between detailed sending regions, likely reducing the apparent influence of previous migration choices on a variety of outcomes.

We demonstrate the practical usefulness of these data by using information on source-specific

[26] Hoekstra and Orozco-Aleman (2017) examine a related question using a later Arizona law, SB 1070, which imposed unprecedented immigration enforcement measures. They use EMIF data to document decreased intended migration to Arizona among unauthorized Mexican migrants after the law's passage. Our novel contribution is to document changes in realized return migration and emigration at the Mexican source level. Given the shortcomings of EMIF described in Section 3, such an analysis would not be feasible using that data source.

migrant destinations to study the effects of the 2008 Legal Arizona Workers Act (LAWA) on international migration between Mexico and the U.S. We find increased return migration to and decreased emigration from Mexican regions with stronger pre-existing migration ties to Arizona. These findings indicate that labor market interventions can strongly affect international migration patterns and confirm the value of the MCAS data in measuring meaningful links between source and destination regions.

The MCAS data therefore represent a valuable resource for researchers seeking to understand the influence of previous migration patterns on a variety of subsequent outcomes. As in our analysis of LAWA, one could use these data to examine the effects of local enforcement measures on international migration to and from the most affected source areas. These data can also be used to track the evolution of migration patterns to examine new migrant sources and destinations and the relationships among them. Alternatively, source-specific or source-destination-specific measures of previous migration could be included as second-level measures in multilevel models. For example, these data provide an excellent measure of the intensity of previous migration out of a source community, which has been shown to affect both the likelihood that an individual migrates and the composition of migrants (McKenzie and Rapoport 2010). Because these data do not contain information at the individual level, however, they have important limitations. For instance, they could not be used to examine the relative labor market success of migrants based on whether they follow previous migrants from their home town (Munshi 2003). For these types of questions, the Mexican Migration Project continues to be the best available source of individual level data.

References

Bartel, Ann P., "Where Do the New U.S. Immigrants Live?," *Journal of Labor Economics*, 1989, *7* (4), 371–391.

Bauer, Thomas K, Gil S Epstein, and Ira N Gang, "Herd effects or migration networks? The location choice of Mexican immigrants in the US," *Manuscript, CEPR Discussion Paper No. 3505*, 2002.

Blau, Francine D and Christopher Mackie, "The Economic and Fiscal Consequences of Immigration," *A Report of the National Academies, Washington DC*, 2016.

Bohn, Sarah, Magnus Lofstrom, and Steven Raphael, "Lessons from the 2007 Legal Arizona Workers Act," *Public Policy Institute of California*, 2011.

___ , ___ , and ___ , "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?," *The Review of Economics and Statistics*, 2014.

DiMaggio, Paul and Filiz Garip, "Network Effects and Social Inequality," *Annual Review of Sociology*, 2012, *38*, 93–118.

Dolfin, Sarah and Garance Genicot, "What do networks do? The role of networks on migration and "coyote" use," *Review of Development Economics*, 2010, *14* (2), 343–359.

Duncan, Otis Dudley, "The Measurement of Population Distribution," *Population Studies*, 1957, *11* (1), 27–45.

Garip, Filiz, *On the move: Changing mechanisms of Mexico-US migration*, Princeton University Press, 2016.

___ and Asad L Asad, "Network Effects in Mexico–US Migration: Disentangling the Underlying Social Mechanisms," *American Behavioral Scientist*, 2016, *60* (10), 1168–1193.

Gonzalez-Barrera, Ana, "More Mexicans Leaving Than Coming to the U.S.," Technical Report, Pew Research Center, Washington, D.C. November 2015.

Hoekstra, Mark and Sandra Orozco-Aleman, "Illegal Immigration, State Law, and Deterrence," *American Economic Journal: Economic Policy*, May 2017, *9* (2), 228–252.

IME, "Matricula Consular de Alta Seguridad (MCAS)," *Mexicanos en el Exterior*, September 2004, *1* (10).

Jaeger, David A, "Local labor markets, admission categories, and immigrant location choice," *Manuscript, College of William and Mary*, 2000.

Kaplan, Greg and Sam Schulhofer-Wohl, "Interstate Migration Has Fallen Less Than You Think: Consequences of Hot Deck Imputation in the Current Population Survey," *Demography*, 2012, *49* (3), 1061–1074.

Lafortune, Jeanne and José Tessada, "Smooth (er) landing? the dynamic role of networks in the location and occupational choice of immigrants," *Mimeo, May*, 2014.

Li, Guoying, "Robust Regression," in David C. Hoaglin, Frederick Mosteller, and John W. Tukey, eds., *Exploring Data Tables, Trends, and Shapes*, New York: Wiley, 1985, chapter 8, pp. 281–343.

Massey, Douglas S, "The social organization of Mexican migration to the United States," *The Annals of the American Academy of Political and Social Science*, 1986, *487* (1), 102–113.

——— **and Kristin E Espinosa**, "What's driving Mexico-US migration? A theoretical, empirical, and policy analysis," *American journal of sociology*, 1997, *102* (4), 939–999.

Massey, Douglas S. and Rene Zenteno, "A Validation of the Ethnosurvey: The Case of Mexico-U.S. Migration," *International Migration Review*, 2000, *34* (3), 766–793.

———**, Jacob S. Rugh, and Karen A. Pren**, "The Geography of Undocumented Mexican Migration," *Mexican Studies/*Estudios Mexicanos, Winter 2010, *26* (1), 129–152.

McConnell, Eileen Diaz, "The US Destinations of Contemporary Mexican Immigrants1," *International Migration Review*, 2008, *42* (4), 767–802.

McKenzie, David and Hillel Rapoport, "Self-selection patterns in Mexico-US migration: The role of migration networks," *The Review of Economics and Statistics*, 2010, *92* (4), 811–821.

Mundra, Kusum and Fernando Rios-Avila, "Immigrant Birthcountry Networks and Unemployment Duration: Evidence around the Great Recession," *Manuscript, IZA Discussion Papers No. 10233*, 2016.

Munshi, Kaivan, "Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market," *The Quarterly Journal of Economics*, 2003, *118* (2), 549–599.

National Conference of State Legislatures, "States Offering Driver's Licenses to Immigrants," `http://www.ncsl.org/documents/immig/DL_EnactedTable_July2015.pdf` July 2015.

National Immigration Law Center, "Basic Facts About the *Matrícula Consular*," `https://www.nilc.org/wp-content/uploads/2015/11/Basic-Facts-about-the-Matricula-Consular.pdf` December 2015.

Palloni, Alberto, Douglas S Massey, Miguel Ceballos, Kristin Espinosa, and Michael Spittel, "Social Capital and International Migration: A Test Using Information on Family Networks 1," *American Journal of Sociology*, 2001, *106* (5), 1262–1298.

Passell, Jeffrey S. and D'Vera Cohn, "Unauthorized Immigrant Totals Rise in 7 States, Fall in 14," Technical Report, Pew Research Center, Washington, D.C. November 2014.

Patel, Krishna and Francis Vella, "Immigrant networks and their implications for occupational choice and wages," *Review of Economics and Statistics*, 2013, *95* (4), 1249–1277.

Riosmena, Fernando and Douglas S. Massey, "Pathways to El Norte: Origins, Destinations, and Characteristics of Mexican Migrants to the United States," *International Migration Review*, 2012, *46* (1), 3–36.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek, "Integrated Public Use Microdata Series: Version 5.0," 2010. Machine-readable database. Minneapolis: University of Minnesota.

Secretaría de Relaciones Exteriores, "*Matrícula Consular*," `https://consulmex2.sre.gob.mx/nuevayork/index.php/espanol/servicios-consulares/matricula-consular` July 2016.

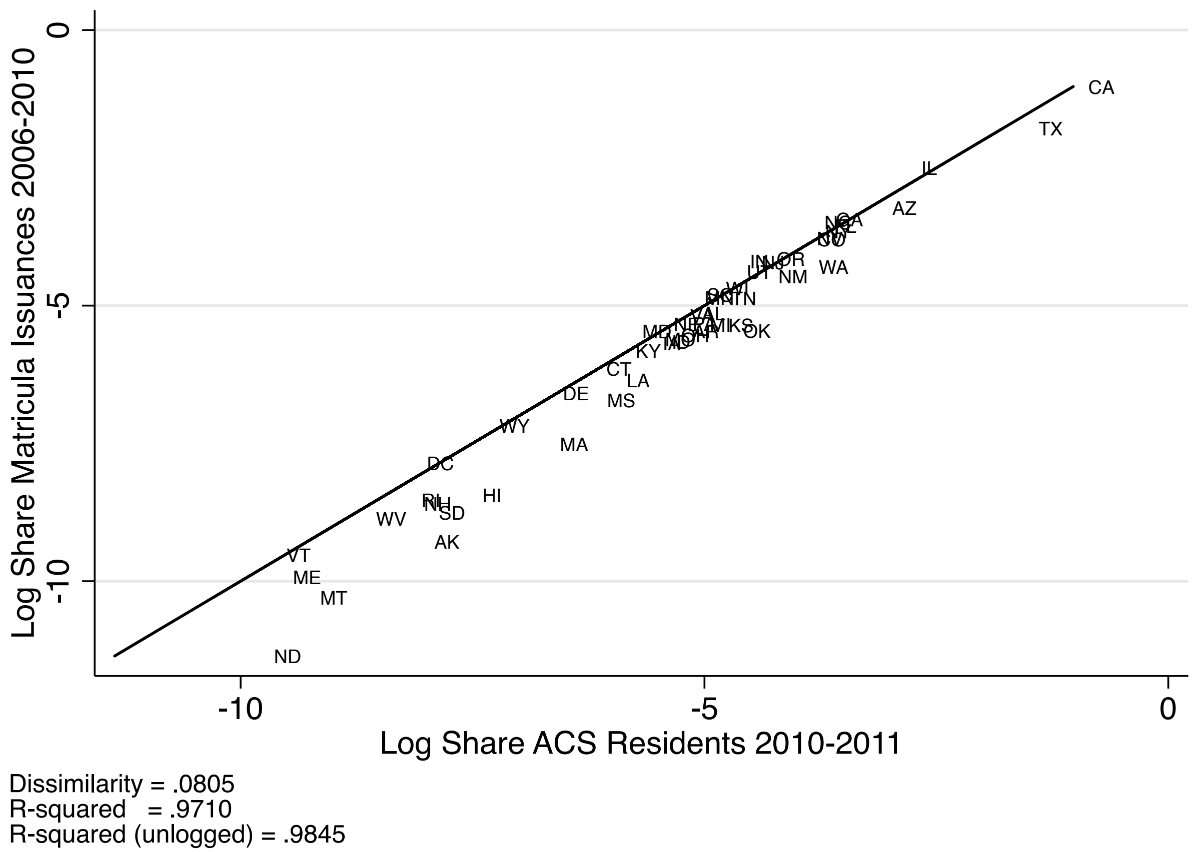
Suro, Robert, "Survey of Mexican Migrants," Technical Report, Pew Hispanic Center March 2005.

___ **and Gabriel Escobar**, "Survey of Mexicans Living in the U.S. on Absentee Voting in Mexican Elections," Technical Report, Pew Hispanic Center Februray 2006.

Winters, Paul, Alain de Janvry, and Elisabeth Sadoulet, "Family and Community Networks in Mexico-U.S. Migration," *The Journal of Human Resources*, 2001, *36* (1), 159–184.

Woodruff, Christopher and Rene Zenteno, "Migration networks and microenterprises in Mexico," *Journal of development economics*, 2007, *82* (2), 509–528.

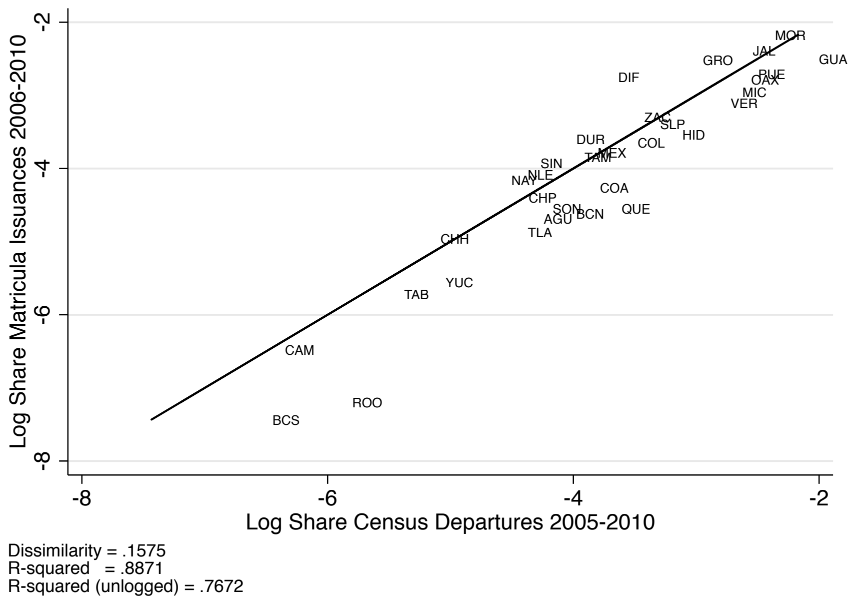
Figure 1: Comparison of U.S. Destination Distribution: MCAS vs. ACS



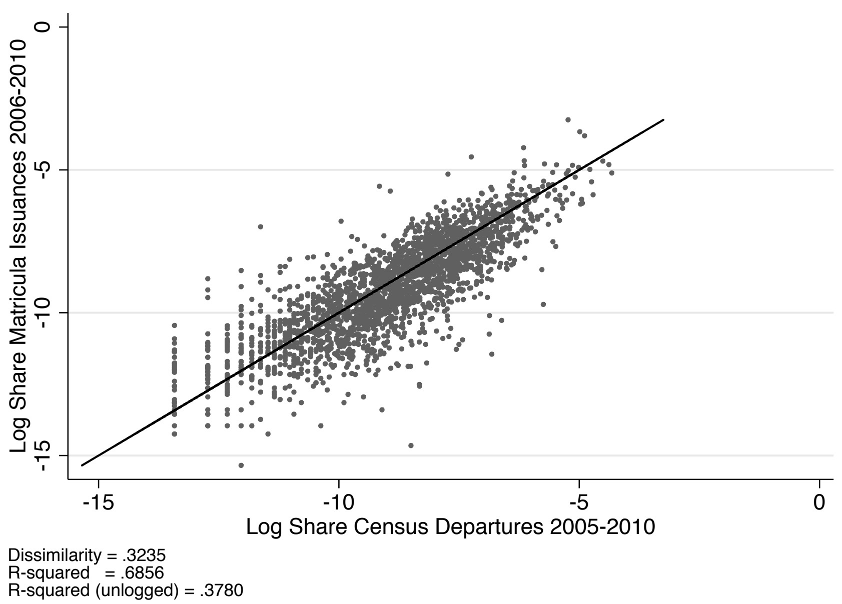
The figure plots the distribution of Mexican-born individuals across U.S. destinations. Each point represents the natural log of the share of individuals in each dataset living in each U.S. state. The ACS sample includes Mexican-born individuals sampled in 2010 or 2011. The MCAS data include the universe of identity cards issued during 2006-2010. These cards were valid through the 2010-2011 time frame covered by the ACS sample. The 45-degree line, which would indicate perfect agreement between the two data sources, is shown for reference. The dissimilarity index, shown in the lower left corner, is defined in Equation (1) in the text and is interpreted as the share of individuals that would need to be reallocated to make the two datasets' distributions match exactly. The R-squared value corresponds to the specification shown in the figure, while the "unlogged" version applies to a comparison of raw unlogged shares.

Figure 2: Comparison of Mexican Source State and Source *Municipio* Distributions: MCAS vs. Mexican Census

(a) Source States



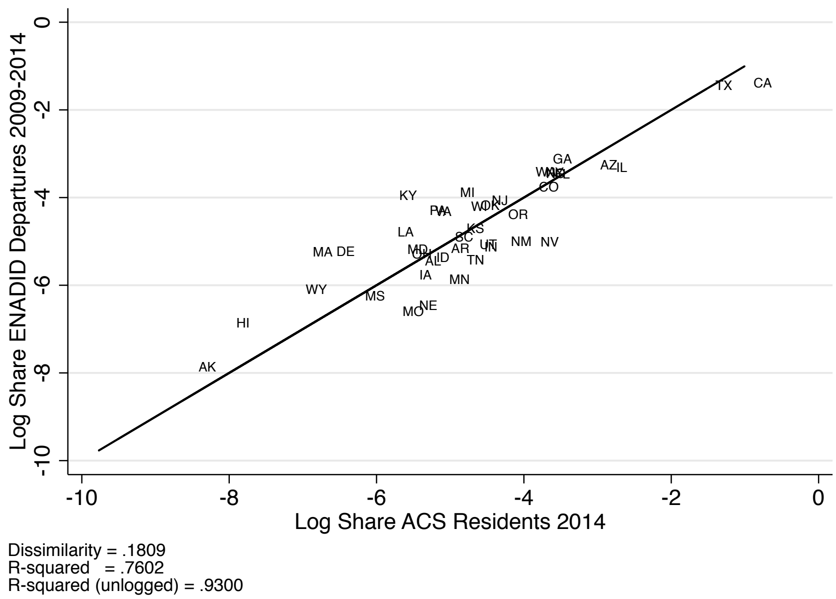
(b) Source *Municipios*



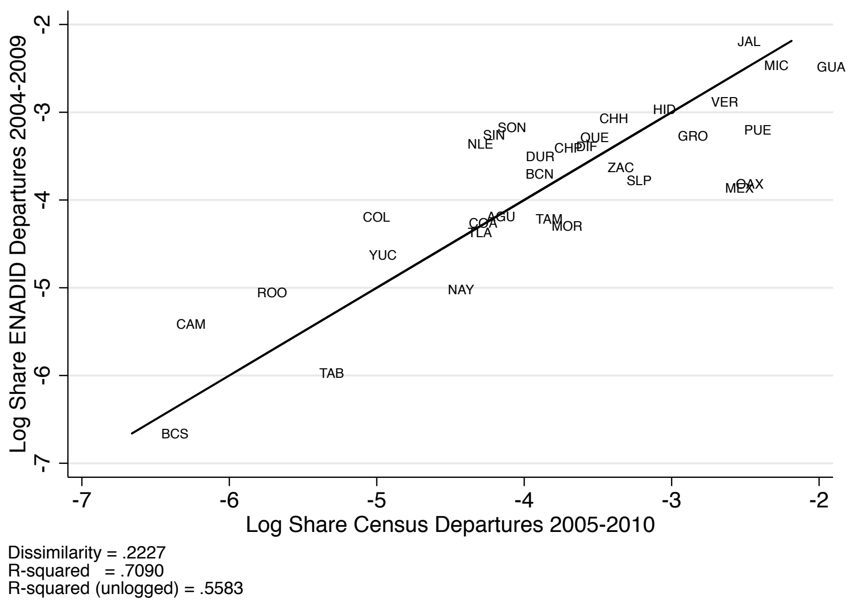
The figures plot the distribution of Mexican source states and source *municipios* for migrants to the U.S. Each point represents the natural log of the share of individuals in each dataset from each Mexican state or *municipio*. The Mexican Census sample includes individuals who moved to the U.S. during the five year period from June 2005 to June 2010. The MCAS sample includes the universe of identity cards issued during 2006-2010. Vertical striping in Panel (b) reflects municipios with very small numbers of observations. The 45-degree line, which would indicate perfect agreement between the two data sources, is shown for reference. The dissimilarity index, shown in the lower left corner of each panel, is defined in Equation (1) in the text and is interpreted as the share of individuals that would need to be reallocated to make the two datasets' distributions match exactly. The R-squared value corresponds to the specifications shown in the figure, while the "unlogged" version applies to comparisons of raw unlogged shares.

Figure 3: Comparisons of ENADID Source and Destination Distributions

(a) Distribution of U.S. Destination States: ENADID vs. ACS

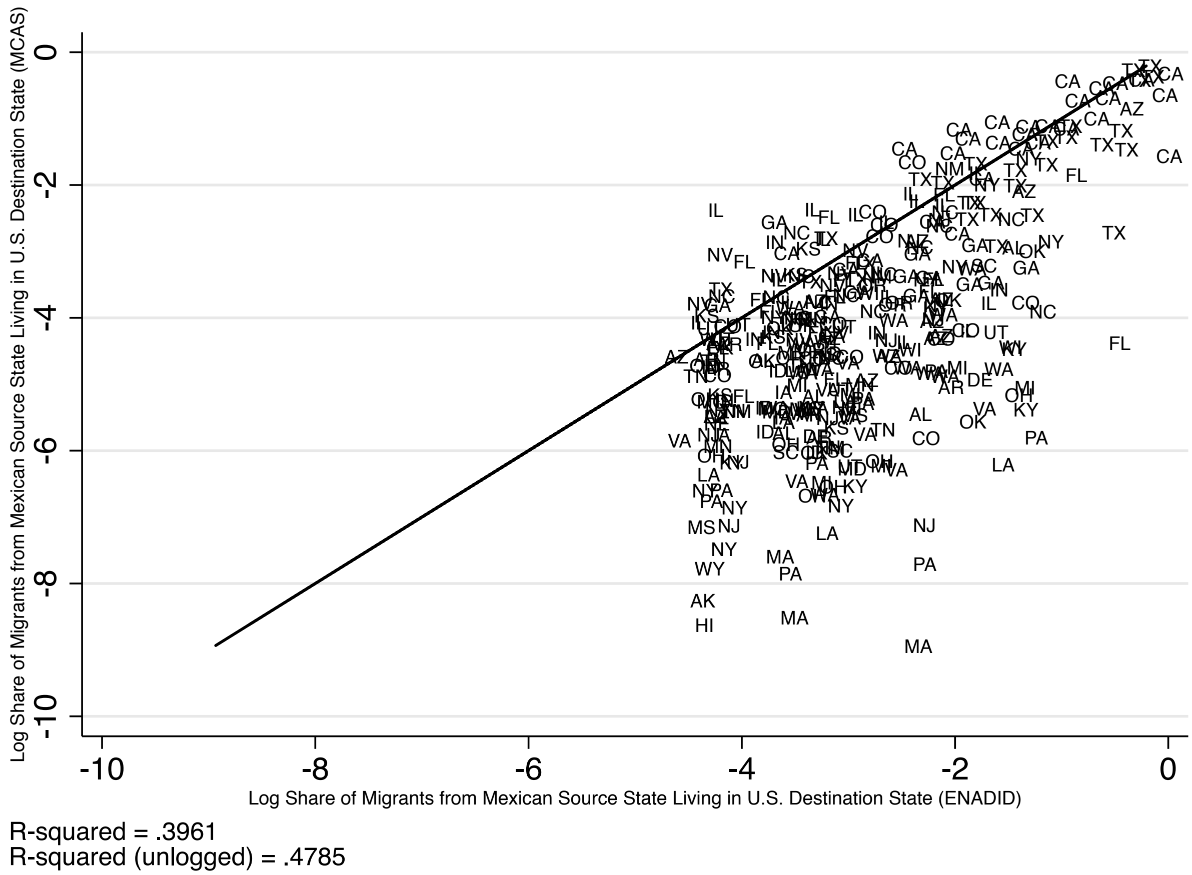


(b) Distribution of Mexican Source States: ENADID vs. Mexican Census



Panel (a) plots the distribution of Mexicans across U.S. destination states. Each point represents the natural log of the share of individuals in each dataset living in each U.S. state. The ACS sample includes Mexican-born individuals living in the U.S. by 2014. In both panels, The ENADID sample includes those who moved to the U.S. during the five year period from May 2009 through May 2014. Panel (b) plots the distribution of Mexican source states for migrants to the U.S. The Mexican Census sample includes individuals who moved to the U.S. during the five year period from June 2005 through June 2010. The 45-degree line, which would indicate perfect agreement between the two data sources, is shown for reference. The dissimilarity index, shown in the lower left corner of each panel, is defined in Equation (1) in the text and is interpreted as the share of individuals that would need to be reallocated to make the two datasets' distributions match exactly. The R-squared value corresponds to the specifications shown in the figure, while the "unlogged" version applies to comparisons of raw unlogged shares.

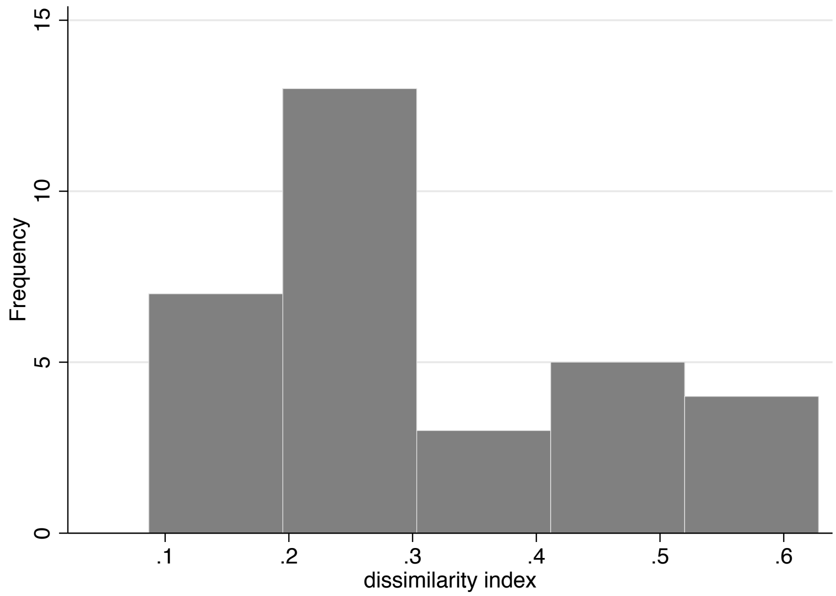
Figure 4: Comparison of Joint Source and Destination State Distribution: MCAS vs. ENADID



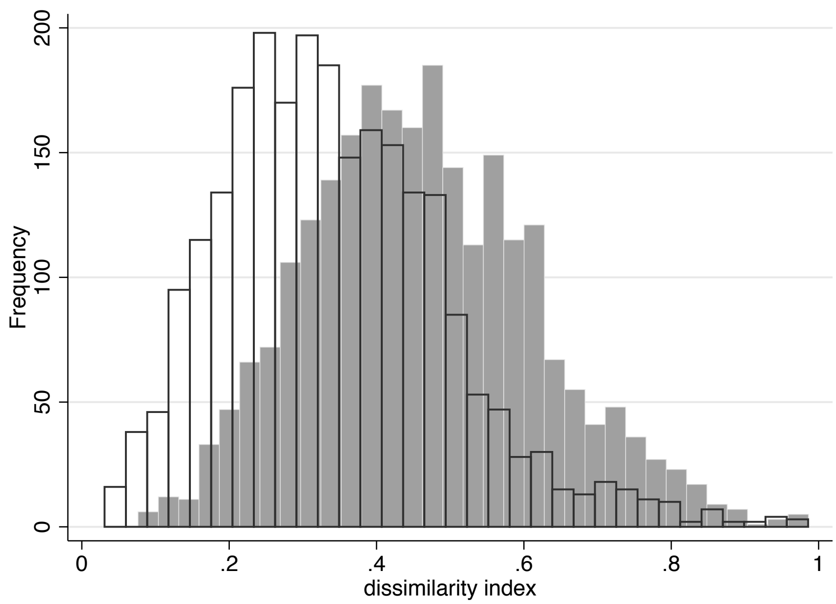
The figure plots the distribution U.S. destination states for migrants from each Mexican source state. Each point represents the natural log of the share of migrants living in the labeled U.S. state for each Mexican source state, from each dataset. MCAS data include the universe of identity cards issued during 2009-2013. The ENADID sample includes individuals who moved to the U.S. in the five year period from May 2009 through May 2014. The 45-degree line, which would indicate perfect agreement between the two data sources, is shown for reference. Source-destination pairs with no observations in the ENADID are omitted. The R-squared value corresponds to the specification shown in the figure, while the "unlogged" version applies to a comparison of raw unlogged shares.

Figure 5: Distribution of Dissimilarity Index

(a) Mexican States vs. Overall Distribution



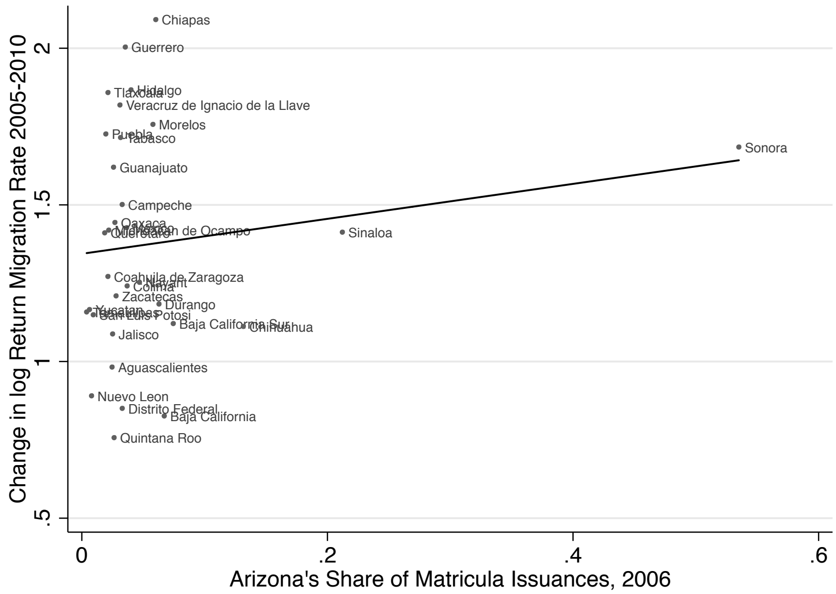
(b) Mexican Municipios vs. Overall or State-Specific Distribution



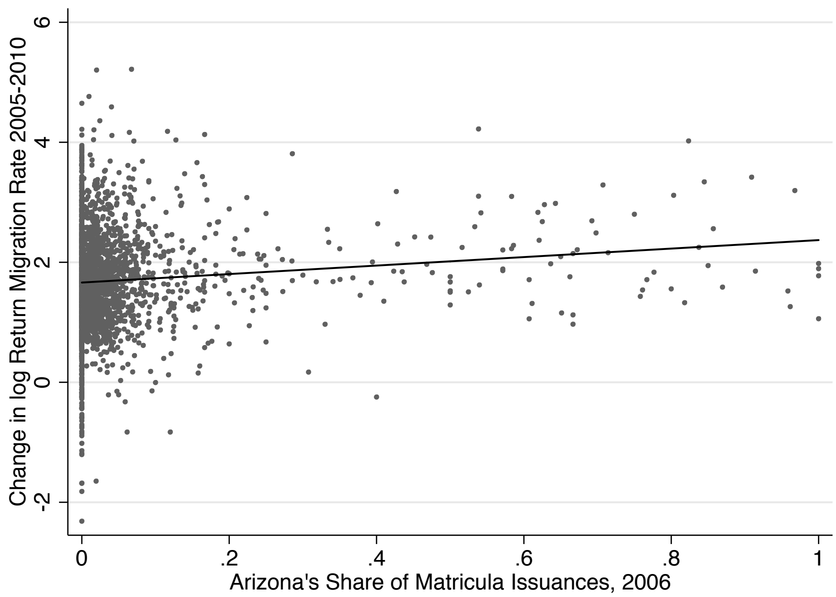
Panel (a) shows the distribution of dissimilarity indexes for the 32 Mexican states, using the national distribution as the reference. Panel (b) shows distributions of dissimilarity indexes for the 2,442 Mexican *municipios*, compared to either the national reference distribution (filled bars) or the reference distribution for the state containing each *municipio* (hollow bars). All analyses use the universe of MCAS identity cards issued from 2006-2010. In both panels, the x-axis represents the dissimilarity index, indicating the share of individuals that would need to be reallocated across destinations for a given source's destination distribution to match the reference distribution.

Figure 6: Change in Log Return Migration Rates vs. Initial Share of Migrants Living in Arizona

(a) Mexican State Level



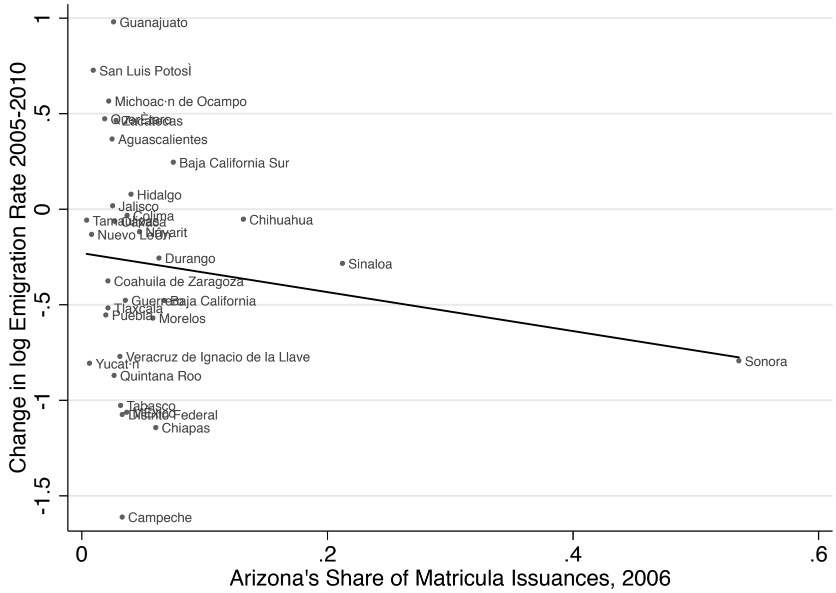
(b) Mexican *Municipio* Level



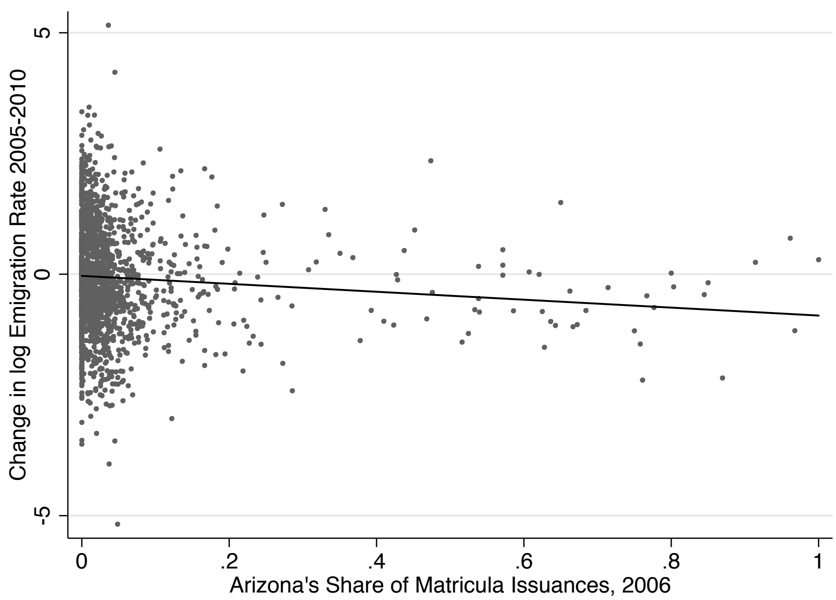
The figures plot the relationship between the change in log return migration rate from 2005 to 2010 in Mexican sending regions (ΔY_s) vs. Arizona's initial share of emigrants from each Mexican sending region (π_s) in 2006, before LAWA went into effect. Migration rates were calculated using the 2005 Mexican *Conteo de Población y Vivienda* and the 2010 Mexican Census, whereas π_s was calculated using the 2006 MCAS. Panel (a) shows the relationship at the Mexican state level, and Panel (b) shows the relationship at the *municipio* level. The regression lines are based on the regression model in Equation (2), with no controls. See Table 2 for alternative specifications.

Figure 7: Change in Log Emigration Rates vs. Initial Share of Migrants Living in Arizona

(a) Mexican State Level



(b) Mexican *Municipio* Level



The figures plot the relationship between the change in log emigration rate from 2005 to 2010 in Mexican sending regions (ΔY_s) vs. Arizona's initial share of emigrants from each Mexican sending region (π_s) in 2006, before LAWA went into effect. Migration rates were calculated using the 2005 Mexican *Conteo de Población y Vivienda* and the 2010 Mexican Census, whereas π_s was calculated using the 2006 MCAS. Panel (a) shows the relationship at the Mexican state level, and Panel (b) shows the relationship at the *municipio* level. The regression lines are based on the regression model in Equation (2), with no controls. See Table 3 for alternative specifications.

Table 1: Data Sources and Measures Used

Data Source	Description	Time Period	Geography	Measure	Figure/Table
Matrícula Consular de Alta Seguridad (MACS)	Administrative data from the MCAS program	2006-2010	U.S. state	Share of unexpired cards issued to residents of each US state as of 2011.	Figure 1
			Mexican state/county	Share of unexpired cards issued to emigrants from each state/county as of 2011.	Figure 2a, Figure 2b
			U.S. state by Mexican state/county	Share of unexpired cards issued to emigrants from each state/county as of 2011(for each source).	Figure 5a, Figure 5b
		2009-2013	U.S. state by Mexican state/county	Share of unexpired cards issued to residents of each US state as of 2014 (for each source).	Figure 4
		2006	Mexican state/county	Share of cards issued to residents of AZ	Figure 6a, Figure 6b, Figure 7a, Figure 7b
American Community Survey (ACS)	IPUMS version of annual 1% survey of U.S. residents.	2010-2011	U.S. state	Share of Mexican-born individuals living in each US state.	Figure 1
		2014	U.S. state	Share of Mexican-born individuals living in each US state.	Figure 3a
Mexican Census	Decennial Population Census microdata conducted by the Mexican Statistics Office (INEGI).	2010	Mexican state/county	Share of 2005-2010 emigrants from each Mexican state or county.	Figure 2a, Figure 2b, Figure 3b
			Mexican county	Return migration rate - number of return migrants from 2005 to 2010 divided by county population in 2005.	Figure 6a, Figure 6b
			Mexican county	Emigration rate in 2005 and 2010 - number who emigrated divided by county population.	Figure 7a, Figure 7b
Encuesta Nacional de la Dinámica Demográfica (ENADID)	National household survey from the Mexican Statistics Office (INEGI) - representative at the Mexican state level.	2014	U.S. state	Share of 2009-2014 emigrants living in each U.S. state.	Figure 3a
			U.S. state by Mexican state	Share of 2009-2014 emigrants living in each U.S. state (for each source).	Figure 4
		2009	Mexican state	Share of 2004-2009 emigrants from each Mexican state.	Figure 3b
Mexican Conteo de Población y Vivienda	Population Census conducted by the Mexican Statistics Office INEGI.	2005	U.S. state by Mexican state	Return migration rate - number of return migrants from 2000 to 2005 divided by county population in 2000.	Figure 6a, Figure 6b

Table 2: Estimated Effect of LAWA on Return Migration Rates

	Change in log return migration rate					
	unweighted	weighted by 2000 population	control for outliers	unweighted	weighted by 2000 population	control for outliers
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Mexican State level Estimates						
Arizona's share of migrants in 2006	0.56**	0.54*	0.58			
	(0.22)	(0.29)	(0.78)			
Constant	1.34***	1.37***	1.33***			
	(0.07)	(0.09)	(0.09)			
Observations	32	32	32			
R-squared	0.022	0.02	0.02			
Panel B. Mexican Municipio level Estimates						
Arizona's share of migrants in 2006	0.71***	0.86***	0.65***	0.76***	1.22***	0.53**
	(0.12)	(0.21)	(0.14)	(0.23)	(0.40)	(0.22)
Constant	1.66***	1.35***	1.65***	1.09***	0.89***	1.09***
	(0.07)	(0.09)	(0.02)	(0.00)	(0.01)	(0.2)
State FE	No	No	No	Yes	Yes	Yes
Observations	2,188	2,188	2,188	2,188	2,188	2,188
R-squared	0.01	0.02	0.01	0.13	0.36	0.16

The reported coefficients come from a regression of the change in log return migration rate from 2005 to 2010 on Arizona's initial share of migrants from the Mexican source region. Positive regression coefficients indicate that regions with a stronger initial connection to Arizona experience a larger increase in return migration following the implementation of LAWA. The return migration rate is calculated using data from the 2000 and 2010 Mexican Census and the 2005 *Conteo*, as the number of return migrants during 2000-2005 or 2005-2010, divided by the source region's initial population. Arizona's initial migrant share is calculated using the 2006 MCAS. Panel A examines Mexican state-level source regions, while Panel B examines *municipio*-level sources. Columns (1) and (4) estimate unweighted (equally weighted) regressions across source regions. Columns (2) and (5) weight sources by their 2000 population, and columns (3) and (6) use robust regression (`rreg` in Stata) to reduce the influence of high leverage outliers. In Panel B, columns (4)-(6) control for Mexican state fixed effects. Heteroskedasticity robust standard errors are shown in parentheses. In Panel B, the standard errors are clustered by Mexican state. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Estimates of the Effect of LAWA on Emigration Rates

	Change in log emigration rate					
	unweighted	weighted by 2005 population	control for outliers	unweighted	weighted by 2005 population	controls for outliers
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Mexican State level Estimates						
Arizona's share of migrants in 2006	-1.02**	-0.89*	-0.98			
	(0.38)	(0.51)	(1.25)			
Constant	-0.23*	-0.36**	-0.23			
	(0.13)	(0.17)	(0.14)			
Observations	32	32	32			
R-squared	0.03	0.01	0.02			
Panel B. Mexican Municipio level Estimates						
Arizona's share of migrants in 2006	-0.82**	-0.60	-0.83***	-1.16**	-1.74**	-1.26***
	(0.31)	(0.46)	(0.24)	(0.55)	(0.79)	(0.41)
Constant	-0.04	-0.57***	-0.05	0.41***	0.25***	0.47
	(0.11)	(0.17)	(0.03)	(0.01)	(0.021)	(0.31)
State FE	No	No	No	Yes	Yes	Yes
Observations	1,752	1,752	1,752	1,752	1,752	1,752
R-squared	0.01	0.00	0.01	0.18	0.25	0.17

The reported coefficients come from a regression of the change in log emigration rate from 2005 to 2010 on Arizona's initial share of migrants from the Mexican source region. Negative regression coefficients indicate that regions with stronger initial connection to Arizona experience a larger decline in emigration following the implementation of LAWA. The emigration rate is calculated using data from the 2010 Mexican Census, as the number reporting emigration in 2005 or 2010, divided by the source region's population in that year. Arizona's initial migrant share is calculated using the 2006 MCAS. Panel A examines Mexican state-level source regions, while Panel B examines *municipio*-level sources. Columns (1) and (4) estimate unweighted (equally weighted) regressions across source regions. Columns (2) and (5) weight sources by their 2000 population, and columns (3) and (6) use robust regression (`rreg` in Stata) to reduce the influence of high leverage outliers. In Panel B, columns (4)-(6) control for Mexican state fixed effects. Heteroskedasticity robust standard errors are shown in parentheses. In Panel B, the standard errors are clustered by Mexican state. *** p<0.01, ** p<0.05, * p<0.1.

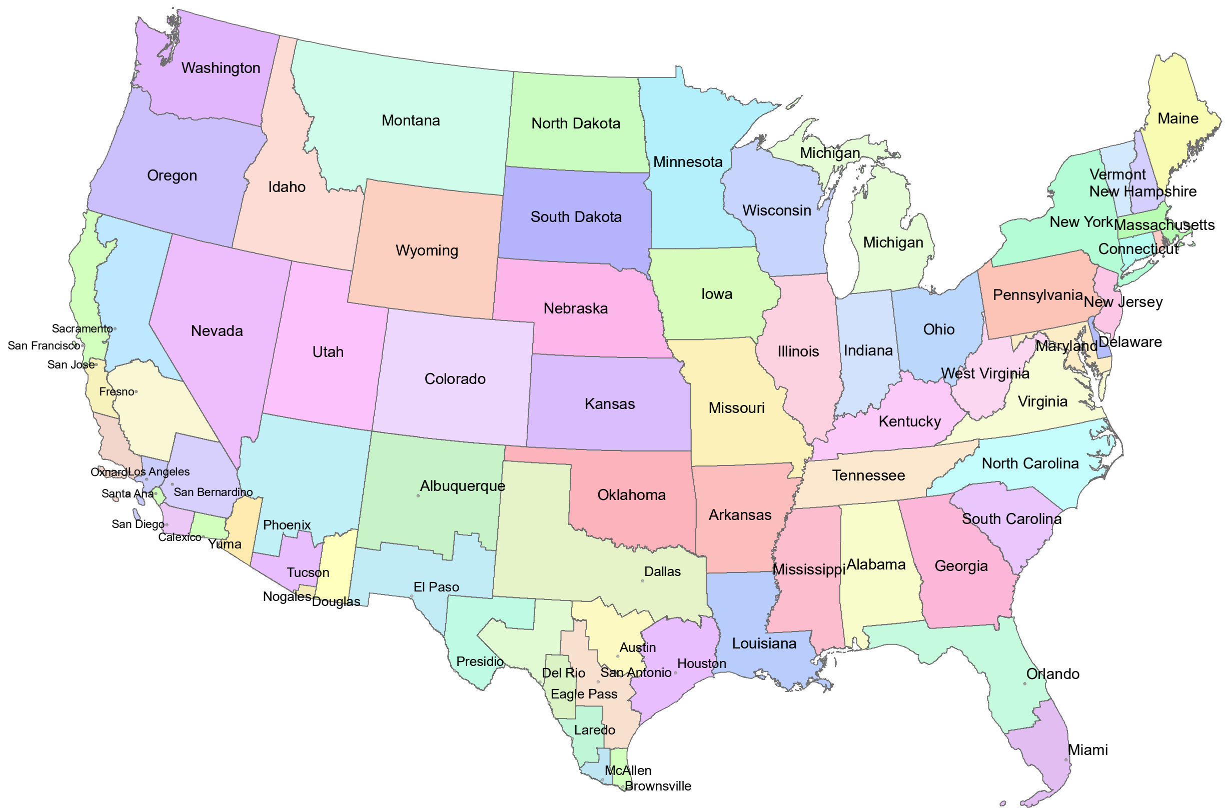
For Online Publication

A Appendix

A.1 MCAS Destination Geography

Figure A-1 shows the 75 destination areas that can be separately observed in the MCAS tabulations by combining information at the U.S. state level and the consular area levels. Note that areas with larger populations tend to have finer geographic detail.

Figure A-1: Map of U.S. Destinations Identifiable in MCAS Tabulations



This map shows the smallest geographic partition of the U.S. that can be constructed using MCAS tabulations by combining tabulations at the U.S. state and consular area levels.

A.2 MCAS Coverage

To assess the takeup rate of MCAS cards, Table A-1 compares the number of valid cards in 2011-2013 to the estimated Mexican-born population of the U.S. (calculated from the ACS) and the Pew Research Center's estimates of the unauthorized Mexican-born population of the U.S. (Passell and Cohn 2014, Gonzalez-Barrera 2015). Each column reports the total count of individuals of all ages.

To understand the potential selection into takeup, we analyzed data from a Pew survey, Suro (2005), which interviewed individuals applying for *matrículas* at various Mexican consulates in Los Angeles, New York, Chicago, Atlanta, Dallas, Raleigh, and Fresno in 2004-2005. Table A-2 shows mean demographic and educational characteristics for this sample of *matrícula* applicants in comparison to all Mexican-born U.S. residents in the 2005 American Community Survey.

Table A-1: Stock of Valid MCAS, Mexican-born Population, and Unauthorized Mexican-born Population

Year	Valid MCAS	Mexican-born Population	Unauthorized Mexican-born Population
2011	4,623,531	11,940,336	6,150,000
2012	4,506,012	11,813,907	5,850,000
2013	4,521,100	11,762,136	5,600,000

The "Valid MCAS" column is based on MCAS tabulations provided by Institute for Mexicans Abroad (*Instituto de Los Mexicanos en el Exterior*, IME). Each row reports the number of cards issued in the preceding 5 years, which are still valid as of the year listed. The "Mexican-born Population" column is estimated using American Community Survey data. We average population estimates for the year listed and the prior year to center the estimates around the beginning of the listed year, corresponding to the timing of MCAS card validity. The "Unauthorized Mexican-born Population" column reports figures from Pew Research Center Reports (Passell and Cohn 2014, Gonzalez-Barrera 2015).

Table A-2: Characteristics of *Matrícula* Applicants vs. Overall Mexican-Born Population

Characteristics	*Matrícula* Applicants	Overall Mexican-born Population
Gender		
Male	0.57	0.55
Female	0.40	0.45
Age		
18-29	0.48	0.30
30-39	0.29	0.31
40-49	0.13	0.20
50-54	0.03	0.06
55+	0.05	0.12
Education		
Primary or less	0.34	0.23
Lower secondary / vocational	0.36	0.23
High school	0.23	0.32
College+	0.07	0.21
Arrival Year		
Before 1990	0.22	0.41
1990-1994	0.16	0.15
1995-1999	0.28	0.20
2000-2004	0.38	0.24

Reports the individuals with each of the characteristics listed in the first column. The "*Matrícula* Applicants" column is based on the Survey of Mexican Migrants (Suro 2005). The "Overall Mexican-born Population" column is based on the 2005 American Community Survey (ACS).

A.3 MCAS Renewals and Issuances

Table A-3 reports the number of MCAS cards issued each year and the share that were new issuances vs. renewals. On average, slightly fewer than 1,000,000 of cards were issued each year from 2006 to 2013, with more than 7,000,000 millions cards issued over the entire period. The vast majority of cards issued were new issuances rather than renewals, which mitigates concerns about double-counting individuals receiving renewed cards within 5 years of the original card's issuance.

Table A-3: MCAS Cards Issued

Year	Cards Issued	Share New Issuances	Share Renewals
2006	945,065	99.81%	0.19%
2007	906,802	99.78%	0.22%
2008	1,007,778	99.73%	0.27%
2009	921,565	99.63%	0.37%
2010	842,321	98.98%	1.02%
2011	827,546	98.05%	1.95%
2012	921,890	97.67%	2.33%
2013	951,570	97.04%	2.96%
Total	7,324,537	98.84%	1.16%

Authors' calculations using data on MCAS cards issuances and the share of new issuances and renewals, provided by the Institute for Mexicans Abroad (*Instituto de los Mexicanos en el Exterior*, IME).

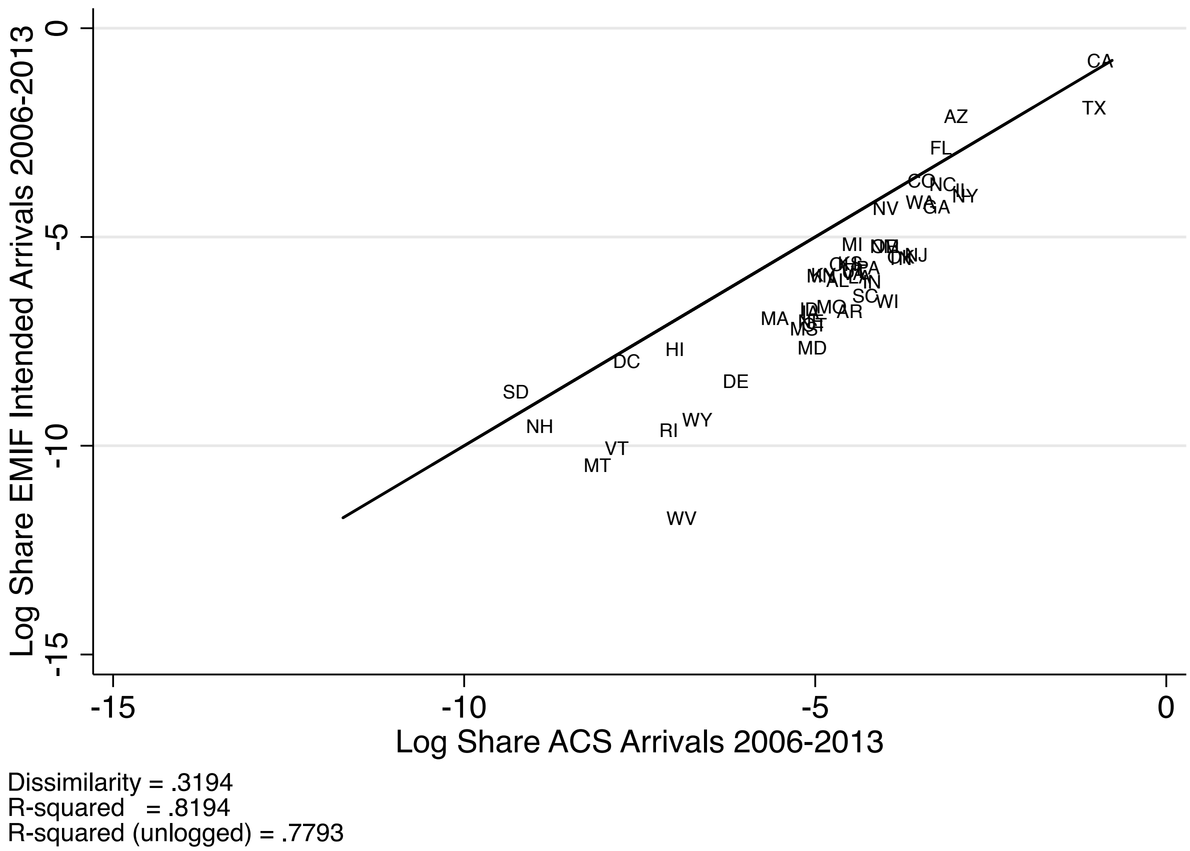
A.4 Comparisons between EMIF and Census data sources

This section examines the source and destination distributions for potential migrants, reported in the *Encuesta sobre Migración en la Frontera Norte* (EMIF).

Figure A-2 compares the share of Mexican migrants planning to migrate to each U.S. state, as reported in the EMIF, to the distribution of Mexican arrivals in each U.S. state, as reported in the ACS, during 2006-2013. There are substantial differences for many U.S. destination states, reflected in the dissimilarity index indicating that 32 percent of the planned migrants from the EMIF would need to be reallocated across U.S. states to exactly match the realized destination distribution in the ACS. Note that the EMIF performs far worse than the MCAS (Figure 1, dissimilarity = 0.08) and the ENADID (Figure 3 Panel (a), dissimilarity = 0.18). Because of this poor performance in matching the distribution of U.S. destinations, we do not examine the EMIF in detail in the main text.

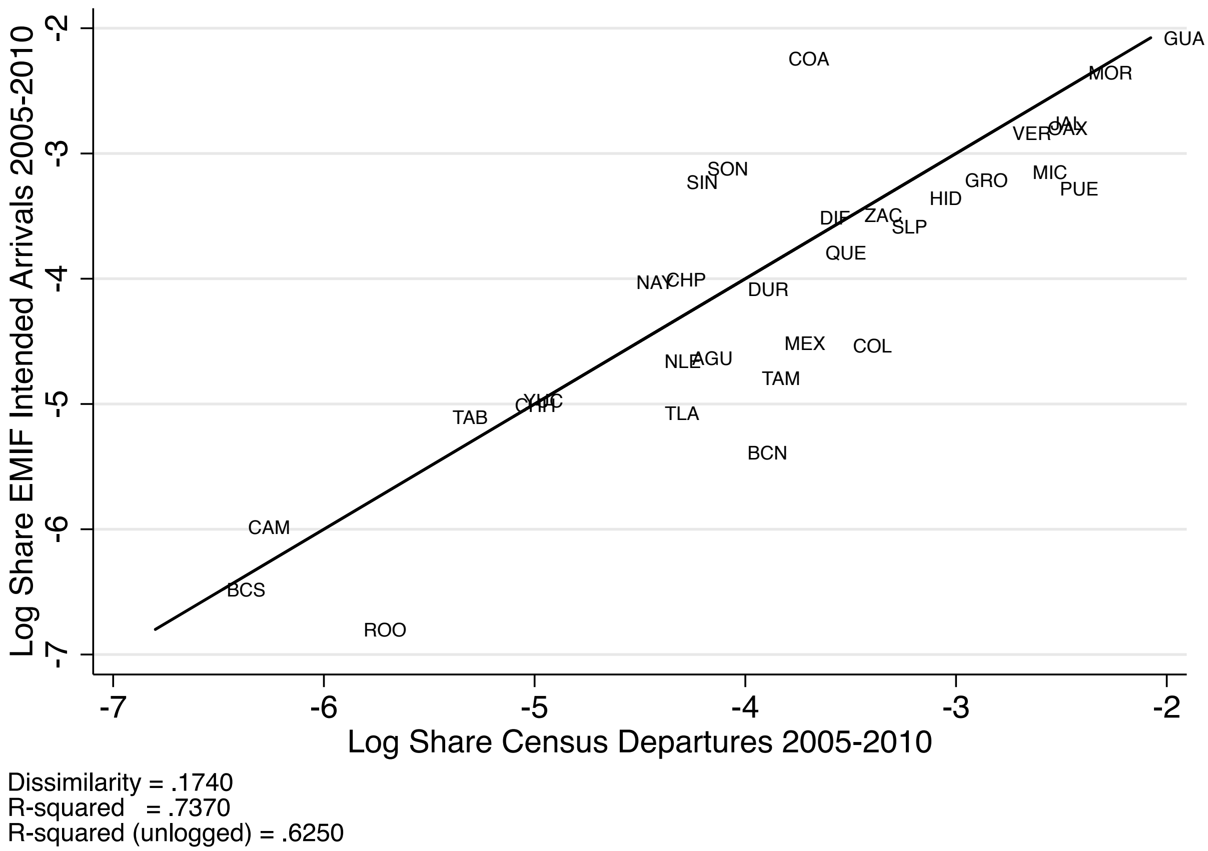
Figure A-3 compares the share of migrants from each Mexican state planning to migrate to the U.S., as reported in the EMIF, to the distribution of emigrants across Mexican State, as reported in the Mexican Census. In this case, the EMIF has a dissimilarity index of 0.17, performing similarly to the MCAS (Figure 2 Panel (a), dissimilarity = 0.16) and the ENADID (Figure 3 Panel (a), dissimilarity = 0.22).

Figure A-2: Comparison of Distribution of US Destinations: ACS vs. EMIF



The figure plots the distribution of Mexican-born individuals across U.S. destinations. Each point represents the natural log of the share of individuals in each dataset selecting the relevant U.S. state as a destination. The ACS sample includes Mexican-born individuals who in the 2013 ACS report arriving in the U.S. during 2006-2013. The EMIF sample includes individuals who report intending to cross into the U.S. during the same time period. The 45-degree line, which would indicate perfect agreement between the two data sources, is shown for reference. The dissimilarity index, shown in the lower left corner, is defined in (1) and is interpreted as the share of individuals that would need to be reallocated to make the two datasets' distributions match exactly.

Figure A-3: Comparison of Distribution of Mexican Source States: Mexican Census vs. EMIF



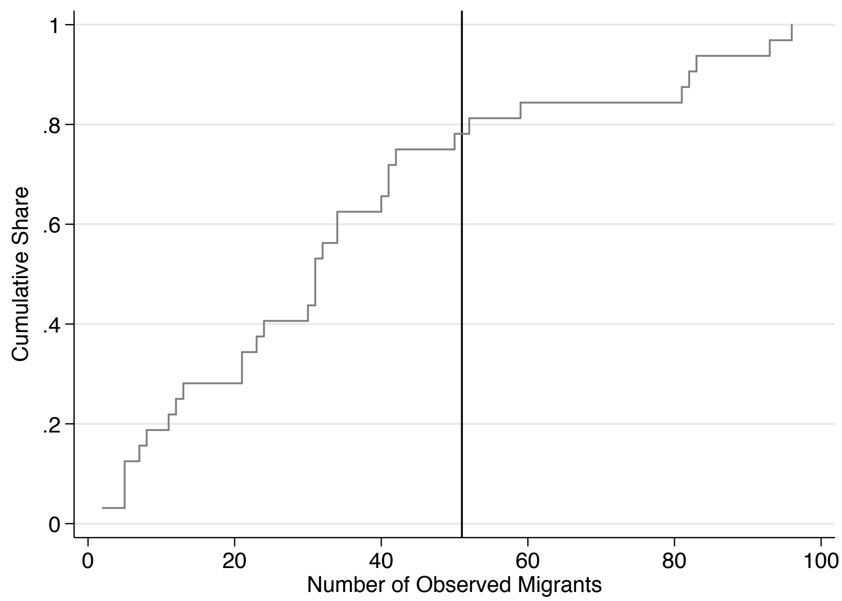
The figure plots the distribution of Mexican source states for migrants to the U.S. Each point represents the natural log of the share of U.S. migrants in each dataset from each Mexican state. The Mexican Census sample includes individuals who moved to the U.S. during the five year year period from June 2005 through June 2010. The EMIF sample includes individuals from Mexican sending states who intended to cross into the U.S. during 2005-2010. The 45-degree line, which would indicate perfect agreement between the two data sources, is shown for reference. The dissimilarity index, shown in the lower left corner, is defined in (1) and is interpreted as the share of individuals that would need to be reallocated to make the two datasets' distributions match exactly.

A.5 ENADID Sample Size Limitation

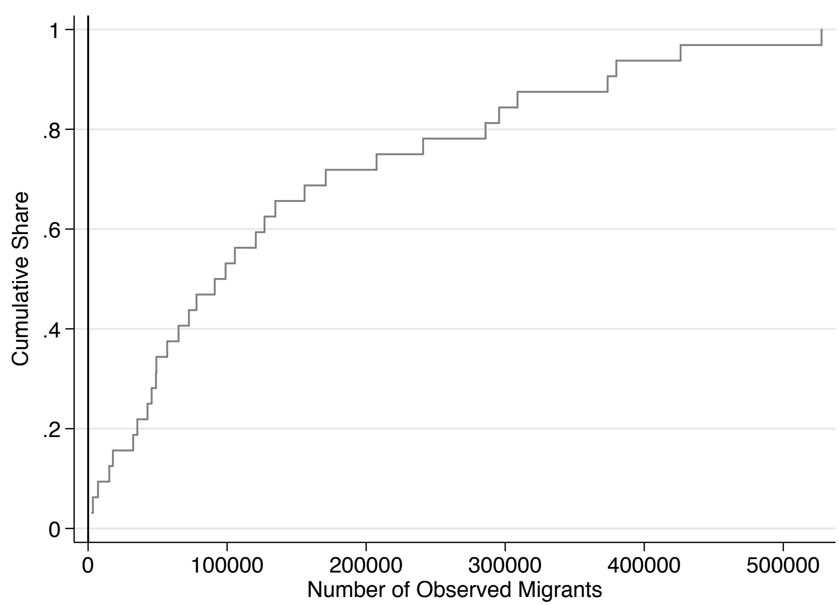
Figure A-4 provides further evidence for the small sample size issue in the ENADID. We plot the empirical cumulative distribution function (CDF) for the number of (unweighted) individual migrant observations for each Mexican source state for the ENADID in Panel (a) and the MCAS in Panel (b). Both figures show a reference line at 51 observations. The intersection of this reference line with the empirical CDF reports the share of Mexican source states for which we observe fewer than 51 migrants, which is the minimum number needed to have no empty source-destination cells when using U.S. states as destinations. Panel (a) reveals that about 80 percent of Mexican source states in ENADID have fewer than 51 observations while Panel (b) reveals that all source states in the MCAS have more than 51 observations. In fact, the majority of source states have more than than 100,000 observed migrants in the MCAS data.

Figure A-4: Comparison of the Distribution of Number of Observations for each Mexican State: ENADID vs. MCAS data

(a) ENADID



(b) MCAS



Each figure plots the empirical cumulative distribution function for the number of (unweighted) individual migrant observations for each Mexican source state. Panel (a) uses the ENADID, and Panel (b) uses the MCAS. The reference line in each figure is 51, which is the minimum number of observations needed to have no empty source-destination cells when using U.S. states as destinations.

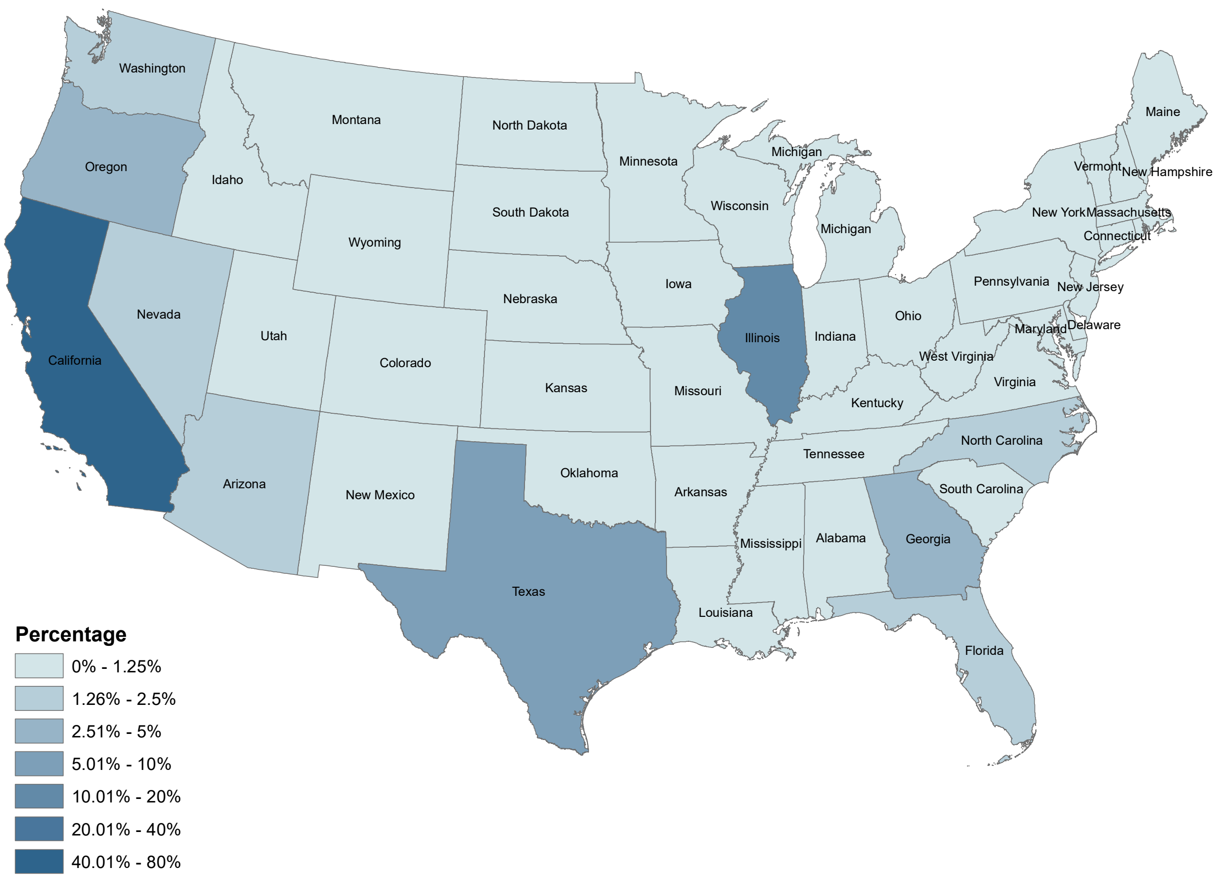
A.6 Detailed Destination Distributions for Different Sources

Figure A-5 shows the distribution of U.S. residences for MCAS cardholders born in Michoacán, and Figure A-6 provides the distribution for all cardholders born anywhere in Mexico. These two distributions are used to calculate the differences shown in Figure A-7.

Figures A-8 and A-9 provide an example of the variation in destinations selected by migrants from two *municipios* within the same state. Migrants from Tiquicheo (Figure A-8) and from Ciudad Hidalgo (Figure A-9) choose substantially different destinations, even though both *municipios* are in Michoacán.

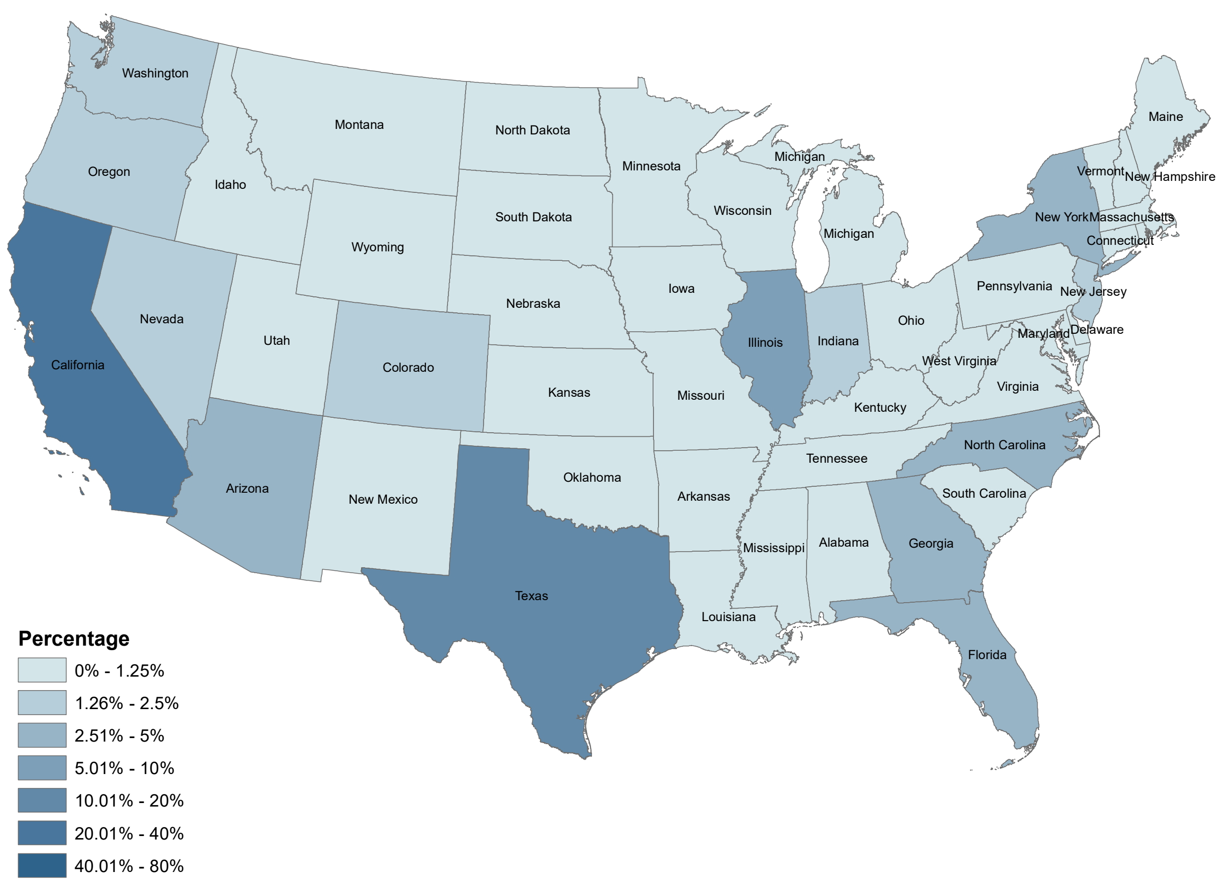
These examples demonstrate that for many *municipios* the destination distribution of their state is not be a good proxy. To reinforce this conclusion, Figure A-10 provides state-by-state differences in shares between Tichicheo and all migrants from Michoacán, revealing a substantial divergence.

Figure A-5: Distribution of MCAS Card Issuances for Migrants Born in Michoacán



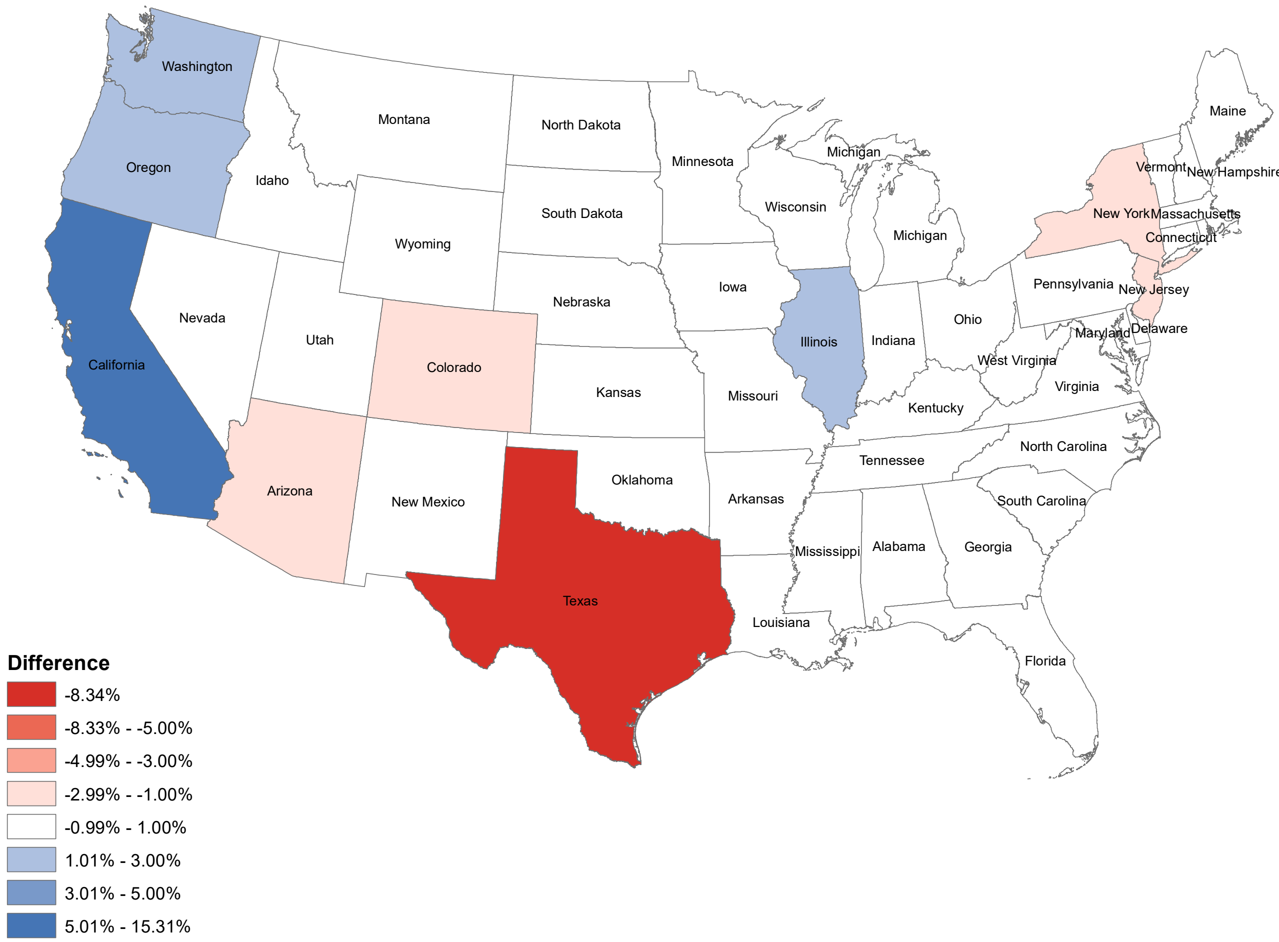
This map shows the share of MCAS identity cards issued to migrants born in Michoacán who had current addresses in each US state. The data source is the universe of MCAS identity cards issued during the 2006-2010 time period.

Figure A-6: National Distribution of MCAS Card Issuances



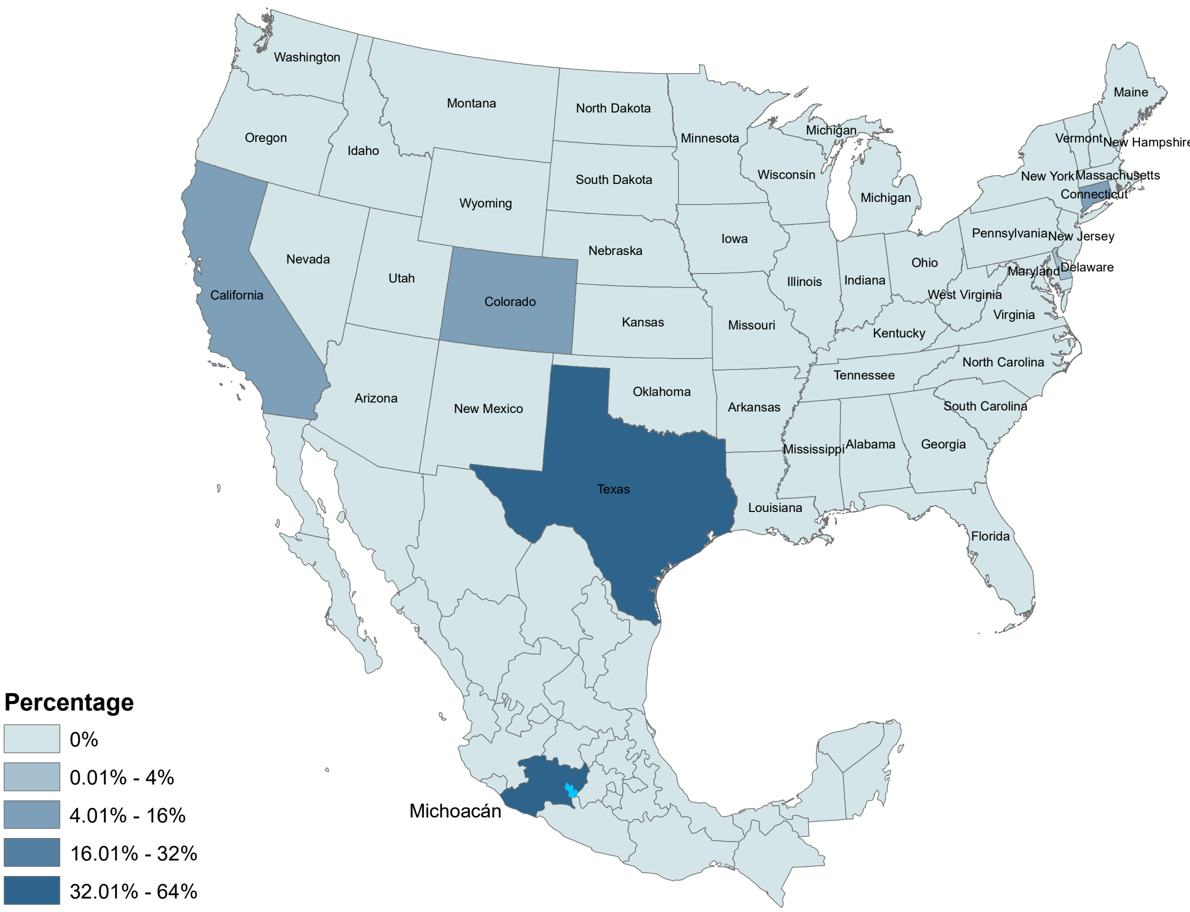
This map shows the share of MCAS identity cards issued to migrants with addresses in each US state. The data source is the universe of MCAS identity cards issued during the 2006-2010 time period.

Figure A-7: Difference in Destination Distribution for Migrants Born in Michoacán vs. All Mexican-Born



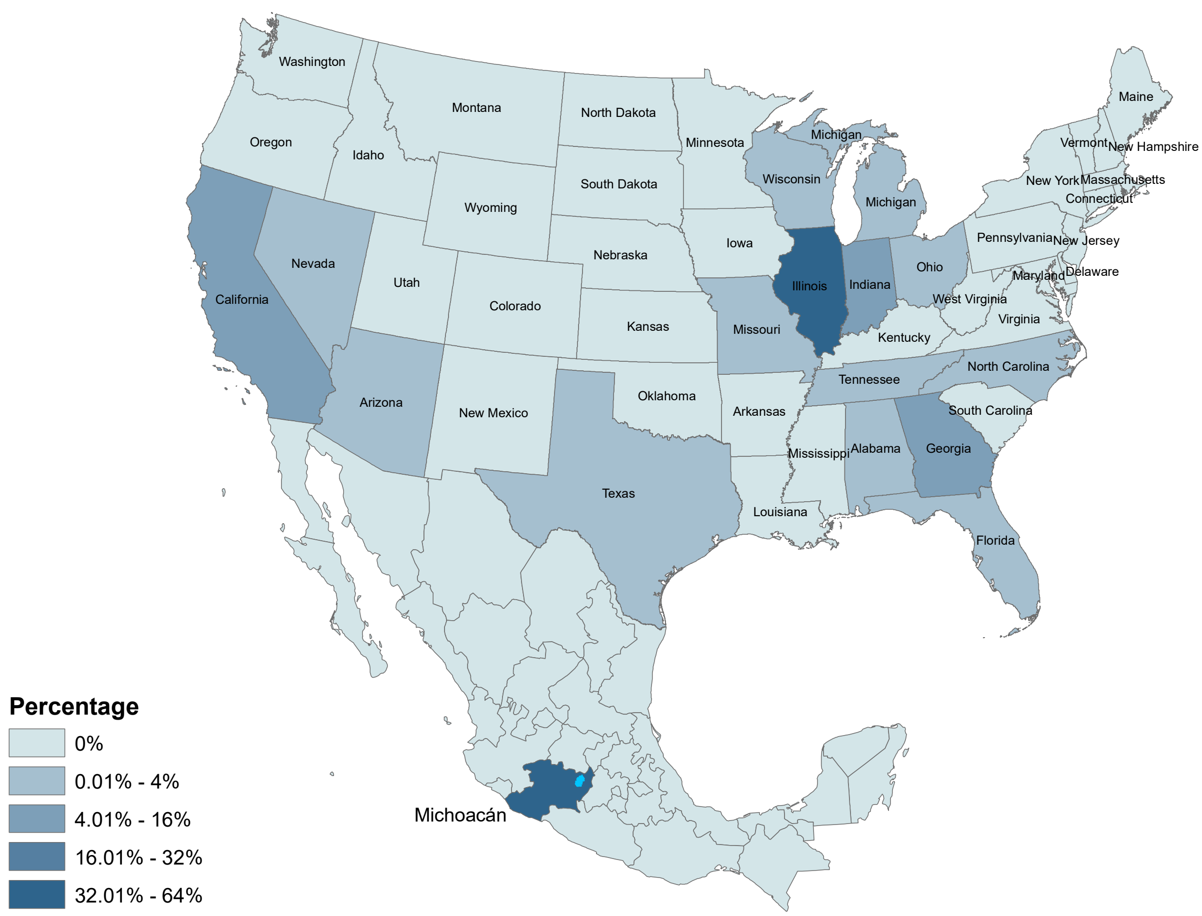
This figure shows the difference in the distributions of U.S. destination states for migrants born in Michoacán vs. all Mexican-born migrants, calculated using the universe of MCAS identity cards issued from 2006-2010. The distributions for each group are mapped separately in Figures A-6 and A-5. Darker blue (red) indicates that the state accounts for a larger (smaller) share of migrants from Michoacán than for all Mexican-born migrants. Migrants born in Michoacán are more likely to live in the West Coast and Illinois and less likely to live in Texas, relative to the average Mexican migrant.

Figure A-8: Distribution of MCAS Card Issuances for Migrants Born in Tiquicheo



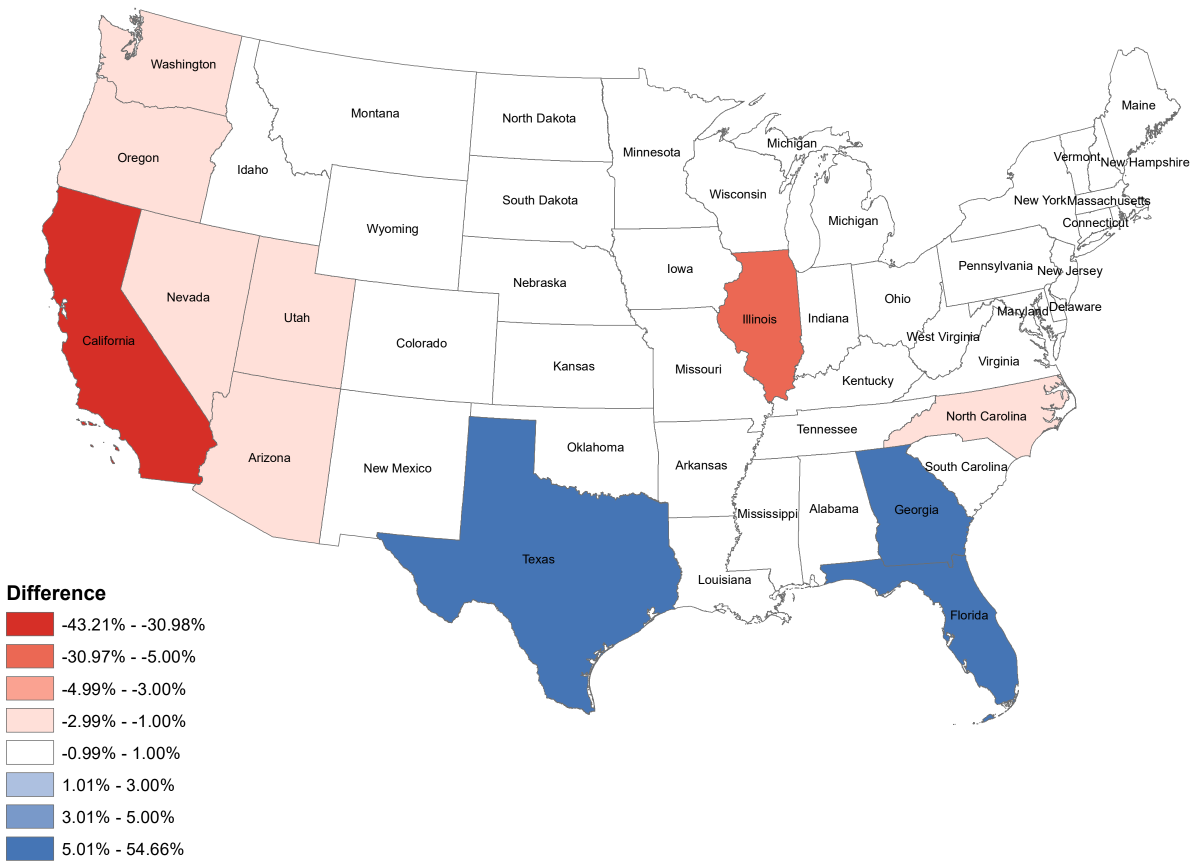
This map shows the destination distribution of the share of MCAS identity cards issued to migrants born in the *municipio* of Tiquicheo (shown in bright blue) who had current addresses in each US state. The data source is the universe of MCAS identity cards issued during the 2006-2010 time period.

Figure A-9: Distribution of MCAS Card Issuances for Migrants Born in Hidalgo



This map shows the destination distribution of the share of MCAS identity cards issued to migrants born in the *municipio* of Hidalgo (shown in bright blue) who had current addresses in each US state. The data source is the universe of MCAS identity cards issued during the 2006-2010 time period.

Figure A-10: Difference in Distribution of MCAS Card Issuances for Migrants Born in Tiquicheo Compared to Michoacán



This figure shows the destination distribution of the state-by-state difference between the share of individuals born in Tiquicheo (a Mexican *municipio* within Michoacn) shown in Figure A-8 and the share of individuals born in the Mexican state of Michoacn shown in Figure A-5.The data source is the universe of MCAS identity cards issued during the 2006-2010 time period.

A.7 Permutation Test of *Municipio* Destination Distribution Variation

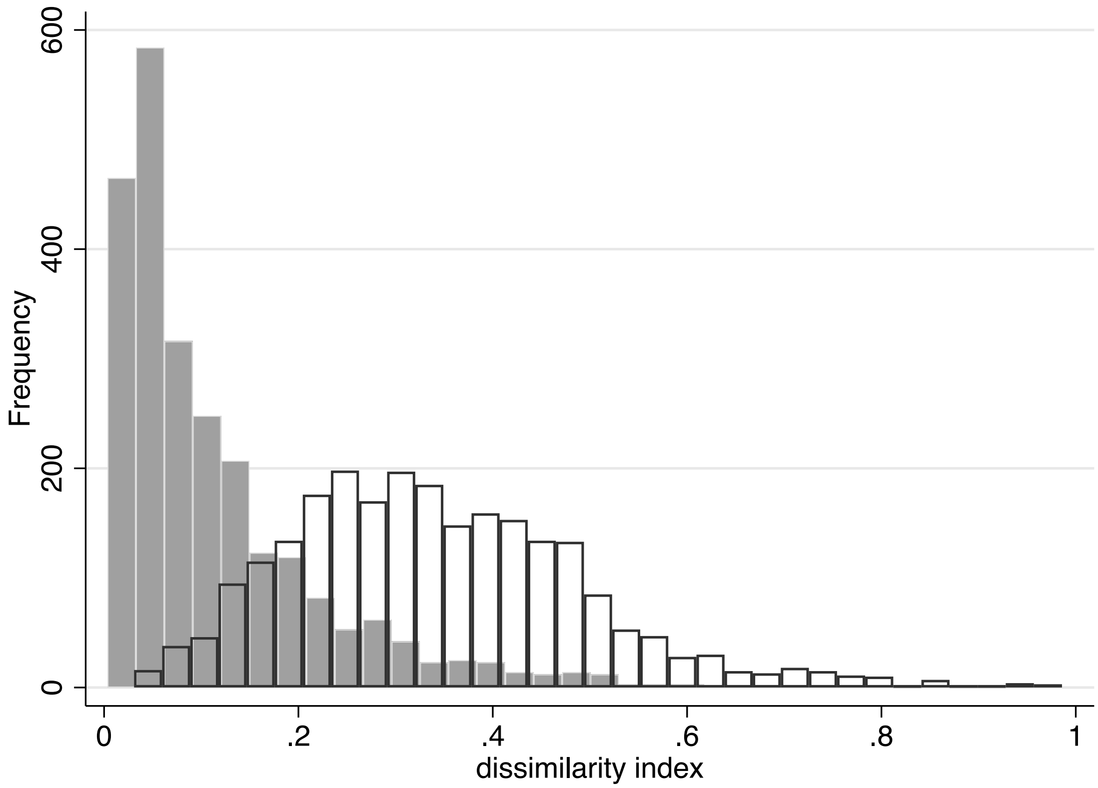
In Figure 5 panel (b), we document substantial differences in destination distributions between Mexican states and the *municipios* within those states, suggesting that assigning all migrants their source state's average destination distribution introduces substantial measurement error into an analysis of the effects of pre-existing local migration patterns. However, random variation could account for some observed differences between *municipio*-level and state-level destination distributions, particularly since some *municipios* are quite small. In this appendix, we implement a permutation test to demonstrate that the observed differences are far larger than those that could plausibly be explained by small *municipio* samples and random variation, concluding that *municipios* in the same state do in fact exhibit substantial differences in their destination distributions.

We implement a permutation test of the null hypothesis that an individual's *municipio* of birth within a Mexican state has no effect on the relevant migrants' destination distribution. Under this null hypothesis, migrants from different municipios within the same state will have the same destination distribution on average, although the observed distributions will be different in finite samples. We simulate the *municipio*-specific destination distributions under this null hypothesis 1000 times by randomly reallocating observed *municipios* of birth among individuals from each state. For each permutation, we then calculate the destination dissimilarity between these simulated *municipios* and their respective states, following Equation (1). We then compare the distribution of dissimilarity indices from these simulations to those using individuals' true birth *municipios*. If the observed differences in destination distribution reflect random variation alone, then the dissimilarity measures for the simulations will be similar to the ones derived using true birth *municipios*.

As an example, the gray bars in Figure A-11 show a histogram of the dissimilarity values for one such simulation. For comparison, we also show the actual distribution of dissimilarity values with hollow bars. It is clear that the dissimilarity values are far higher in reality than they are after randomly assigning birth *municipios*, implying that random variation alone cannot account for the large dissimilarity values that we observe in the data.

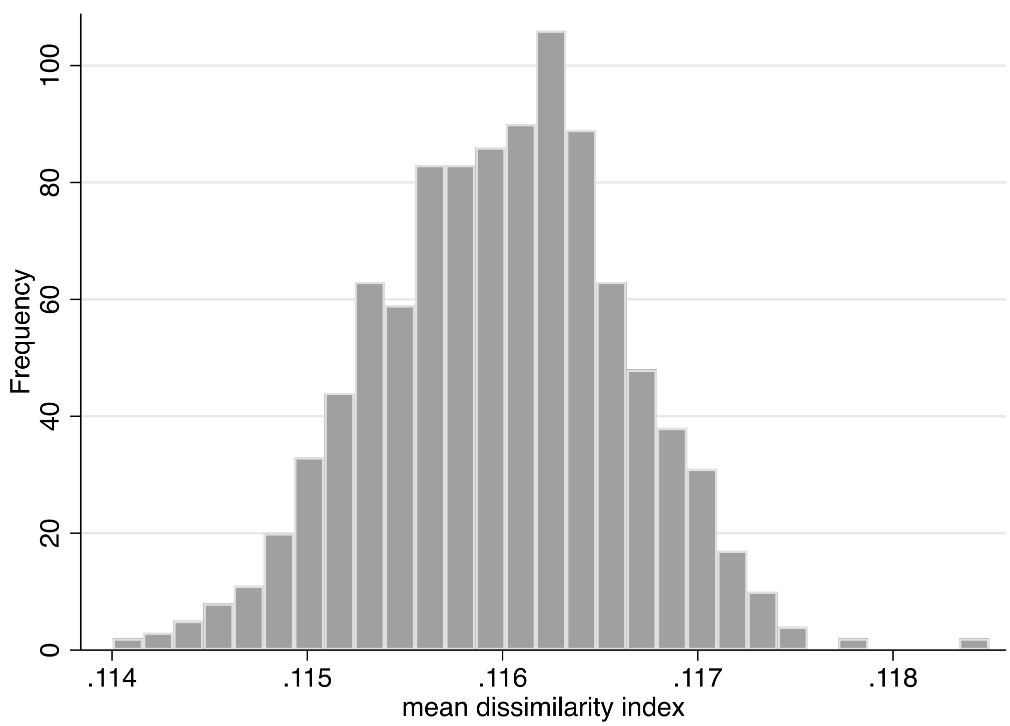
To make this point more formally, we calculate the average dissimilarity index in the data (the mean of the distribution of hollow bars in Figure A-11), yielding a value of 0.343. On average, *municipios* in the data require 34.3 percent of their migrants to be reallocated across destinations to match their state's destination distribution. We then do the same for each of the 1000 simulated permutations just described and plot a histogram of the resulting 1000 average dissimilarity index values in Figure A-12. These values range from 0.114 to 0.118, i.e. none of the 1000 simulated permutations yields an average dissimilarity index approaching the real-world mean of 0.343. We can therefore reject the null hypothesis that an individual's *municipio* of birth within a Mexican state has no effect on the relevant migrant's destination distribution, at significance levels well below 0.1 percent.

Figure A-11: Distribution of Dissimilarity for Simulated and Actual *Municipios*



See the text for a detailed description of construction and interpretation. Gray bars plot the histogram of the dissimilarity index for a random permutation of birth *municipios*, while hollow bars plot the histogram of the dissimilarity index for the real-world birth *municipios*.

Figure A-12: Distribution of Average Dissimilarity for Simulated Permutations



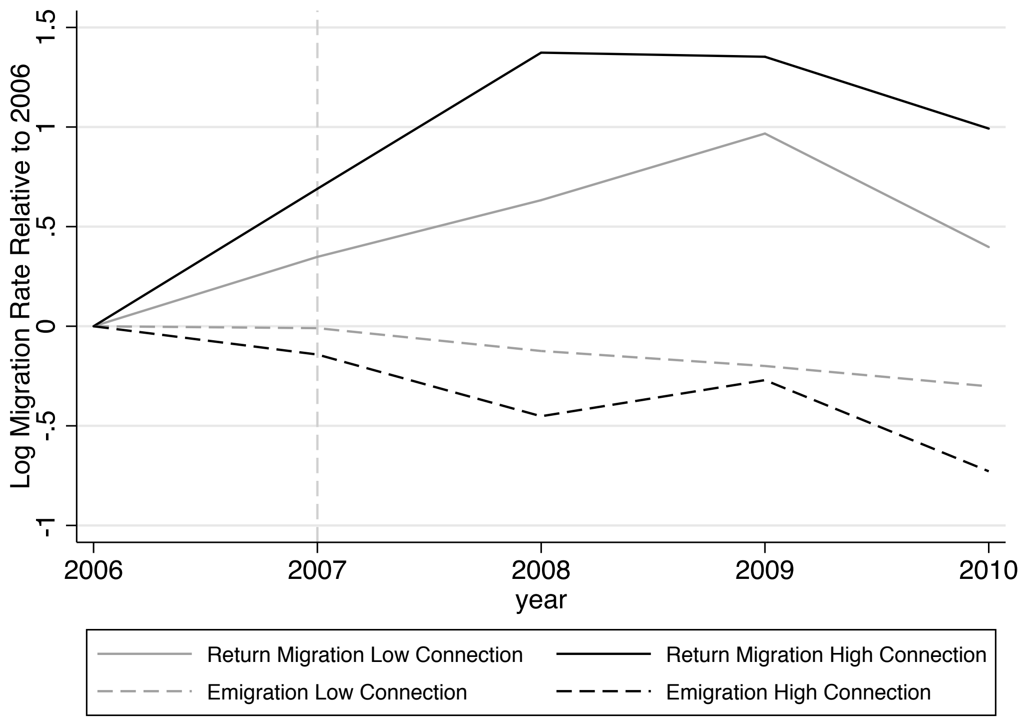
See the text for a detailed description of construction and interpretation. Gray bars plot the histogram of the average dissimilarity index values across *municipios* for each of 1000 simulated permutations.

A.8 Time Series of Return Migration and Emigration Rates

Here, we present information on the time series evolution of return migration rates to and emigration rates from Mexican *municipios* with high and low initial connection to Arizona, to put the results from Section 5 in a broader context. High connection *municipios* are the top 5 percent of *municipios* in terms of their share of 2006 MCAS issuances in Arizona. These "high connection" *municipios* had Arizona as the place of residence for at least 22.4 percent of their 2006 MCAS issuances.

Figure A-13 shows log migration rates, normalized to zero in 2006. Specifically, it shows $\ln(\text{migration rate}_t) - \ln(\text{migration rate}_{2006})$. First, focus on the solid lines, which show return migration rates. After 2006, return migration increases on average for all source *municipios*, reflecting declining economic opportunities in the U.S. following the housing bust and subsequent recession. However, return migration increased more for *municipios* that initially sent a larger share of migrants to Arizona, as seen by the black solid line being above the gray solid line. The dashed lines show emigration rates. Again due to declining economic opportunity in the U.S., emigration rates fall on average for all *municipios*. However, emigration fell more for *municipios* that initially sent a larger share of migrants to Arizona, as seen by the black dashed line being below the gray dashed line. These differences corroborate those in the the scatter plots in Figures 6 and 7 and regression results in Tables 2 and 3.

Figure A-13: Time Series of Return Migration and Emigration Rates by Initial Connection to AZ



See the text for a detailed description of construction and interpretation of the time series. High Connection *municipios* are those with Arizona as the destination for at least 22.4 percent of their identity cards issuances in 2006. Low Connection *municipios* are all others. Values represent the difference in log migration rates between the indicated year and 2006.