The Effect of Mexican Immigration on the Wages and Employment of U.S. Natives: Evidence from the Timing of Mexican Fertility Shocks

Aaron Chalfin
Goldman School of Public Policy
UC Berkeley

Morris Levy
Travers Department of Political Science
UC Berkeley

October 19, 2012

PRELIMINARY DRAFT, SUBJECT TO REVISION

Abstract

Recent literature has used instrumental variables techniques to estimate a causal effect of immigration on labor market outcomes among U.S. natives. The majority of this literature relies on the seminal "network instrument (Altonji and Card 1991), which leverages the persistence of co-ethnic immigrant enclaves to isolate quasi-random variation in migration flows. However, if local labor market conditions that attract or repel migrants are serially correlated, then the network instrument may result in an underestimate of the true effect of immigration on natives' wages and employment. We propose a novel instrument for Mexican immigration to U.S. cities that addresses this concern. Historical data on the size of lagged state-specific Mexican birth cohorts and a time-invariant measure of the persistence of Mexican state-U.S. destination migration relations forms the basis for a decomposition of the network instrument into a portion that is explained by lagged birth cohort sizes and a portion that is not. Our identification strategy relies on two observations. First, larger Mexican birth cohorts in a given Mexican state predict larger cohorts of emigrants from that state when members of the birth cohort reach the age at which migrants typically sojourn or settle in the United States. Second, emigrants from each Mexican state exhibit distinctive, and historically determined, geographic settlement distributions across U.S. cities. Thus, in contrast to the network instrument, which may be contaminated by the endogeneity of immigrant concentrations to persistent pull factors in U.S. cities, our instrument relies exclusively on factors that push migrants out of Mexico and are unlikely to be correlated with conditions in U.S. cities. We use this framework to estimate the effect of Mexican immigration on the wages and employment of U.S. natives in race, age and skill groups. We report evidence that Mexican immigration is associated with no change in either the wages, the unemployment rate or the employment-to-population ratio of unskilled U.S. natives in any education-experience group. Estimates are precise, allowing us to rule out anything other than small effects.

JEL Classification: C14, C21, C52.

Keywords: Immigration, unemployment, wages

*For helpful comments and suggestions on an earlier draft of the paper, we thank Irene Bloemraad, Ron Lee and Steve Raphael. We also thank seminar participants from the Demography Department and the Institute for Research on Labor and Employment's Interdisciplinary Immigration Workshop at the University of California, Berkeley. This research was generously supported by funding from the National Science Foundation's IGERT Traineeship Program in Politics, Economics and Public Policy at UC Berkeley.

I. Introduction

The relationship between the level of immigration and the wages and employment of U.S. natives has been one of the most comprehensively studied topics in labor economics. However, despite the size and scope of the extant literature, researchers remain divided as to whether U.S. natives face labor market displacement as a result of immigration. While structural approaches have tended to find economically important and statistically significant effects of immigration on the economic prospects of natives (see Borjas 2003), area studies, which examine the reduced form effect of immigration in states or metropolitan areas, have tended to find either small or null effects (Card 1990; Altonji and Card 1991; Card 2001; Card and Lewis 2007).

The internal validity of the area studies approach relies on an empirical strategy that identifies plausibly exogenous variation in the timing and destination of U.S.-bound migrants. The most common area studies design analyzes longitudinal data on a large number of cross-sectional units to estimate a "national effect" of immigration. [1] This approach typically relies on an instrumental variable that assigns different numbers of immigrants to each city in each year without influencing labor market outcomes in the city through any channel other than its impact on immigration flows. The seminal instrument in this literature, first proposed in 1991 by Joseph Altonji and David Card, recognizes the salience of immigrant enclaves, and instruments for recent flows of country-specific immigration with the current national flow of migrants to the United States and the distribution of country-specific destinations of past migrants. The approach relies on the empirical observation that immigrants tend to cluster in cities where prior immigrants from their country of origin have already settled. Thus the "network" instrument achieves identification, in part, by leveraging city-specific factors that *pull* immigrants to particular locations. [2] Altonji and Card (1991), Card (2001) and Card and Lewis (2005) have used this instrument to estimate

[1] A second approach identifies a natural experiment in which there is a discrete change in the flow of immigrants to a local area and, using differences-in-differences, compares the change in natives' labor market outcomes in the treated region to natives' outcomes in a control region, defined using some heuristic. This is this approach of Card (1990), who studies the effects of the Mariel boatlift. When the labor market outcomes of natives in Miami are compared to outcomes among natives in Atlanta, Los Angeles, Houston and Tampa-St. Petersburg, there is little evidence of adverse labor market effects. While the approach is transparent, the consistency of the estimated effects hinges on the quality of the heuristic used to construct a comparison group, and there are often concerns over the generalizability of results. Recent advances in robust estimation using differences-in-differences offer an alternative and possibly more sophisticated strategy to select the heuristic. See Abadie, Diamond and Hainmuller (2010) for further discussion.

[2] The intuition is that since these factors emerged a long time ago, such variation should be uncorrelated with contemporaneous time-varying shocks which draw immigrants.

a causal effect of immigration on the labor market outcomes of U.S. natives, finding minimal effects.[3]

To the extent that lagged values of the foreign-born population stock are assumed to be related to native labor market outcomes only via their "pull" on subsequent migrants, the network instrument satisfies the exclusion restriction needed to achieve identification and returns a consistent estimate of the effect of immigration. However, there are several alternate mechanisms through which the prior location decisions of migrants might be associated with current wages. First, to the extent that there is serial correlation in unobserved city-specific factors that are correlated with labor market conditions, the network instrument might isolate not only exogenous variation in migration to that city but also migration that is drawn by persistent city characteristics. If the conditions drawing prior migrant waves to a city persist, today's migrants may be pulled to a city based on similar conditions, and those conditions, in turn, may influence or reflect natives' labor market outcomes. In this case, causal estimates using the network instrument will be inconsistent. Second, the exclusion restriction is violated if there are persistent city-specific shocks that differentially affect traditional gateway cities relative to non-gateways (Card 2001; Pugatch and Yang 2011; Chalfin 2012). For example, if higher wage growth in gateway cities continually attracts immigrant flows, then the network instrument leads to an estimate of the effect of immigration on wages that is positively biased. Put differently, the network instrument simply assumes that the total outflow of immigrants from a given country is exogenous to conditions in the destination cities where immigrants tend to settle.

To generalize, the network instrument may fail to satisfy the exclusion restriction if it taps variation in factors that *pull* immigrants to a given city rather than factors that *push* immigrants to a historically-determined set of locations. Recognizing this, Pugatch and Yang (2011) suggest that a cleaner source of identifying variation may be found in factors that differentially induce *emigration* from different source regions but that do not, in and of themselves, draw migrants to any particular destination. Such factors, situated in migrant source regions, are less likely than pull factors, situated in migrant destinations, to relate to destination labor market conditions through channels other than immigration. Following Munshi (2003), Pugatch and Yang demonstrate that positive deviations in rainfall from Mexican states' long-run rainfall means, by affecting the state of the local economy, induce emigration, providing the

[3]This is presumably in part because immigrants tend to be concentrated in industries in which they compete for employment opportunities with relatively few natives (see Card 2001 for a detailed discussion).

required push. Pugatch and Yang's key insight is that historically-determined and persistent migration networks linking particular Mexican states to U.S. states can be used to predict where a low rainfall-driven Mexican migrant is likely to sojourn or settle in the U.S. Rainfall fluctuations provide as good as random variation in emigration from Mexican states, and emigrants tend to follow pre-determined channels to particular U.S. states. [4] Thus the interaction of Mexican source-state rainfall with a set of weights that reflect Mexican state-U.S. state migration relations provides a valid instrument for variation in U.S. states' Mexican population shares. Pugatch and Yang find that these exogenous increases in the Mexican share of the labor force lead to appreciable declines in wages and increases in unemployment among non-Mexicans, particularly among those in the middle of the skill distribution, including workers whose highest level of education is either a high school diploma or several years of college.

One explanation for the sharp contrast of Pugatch and Yang's findings with the minimal effects reported in the majority of the prior area studies literature is that their instrument is cleansed of the endogeneity implicit in the network instrument. Other explanations, however, raise questions about external validity and interpretation. With respect to external validity, if the attributes of rainfall-driven migrants differ substantially from the attributes of migrants generally, then the rainfall instrument may generate an estimate of an idiosyncratic and unrepresentative local average treatment effect. There are empirical reasons to suppose this may be the case. The first-stage regressions reported indicate that rainfall predicts the concentration of Mexican migrants among high as well as low-skilled workers in the U.S., and the second-stage regressions reported indicate, counter-intuitively, that immigration from Mexico significantly reduces the wages of natives with some college and even those with a college degree. With respect to interpretation, the use of states as proxies for labor markets contrasts with previous literature that has examined the effect of immigration on metropolitan statistical areas, which are designed, in part, to reflect the geographic reach of labor markets. Moreover, while the majority of the extant literature estimates an effect of immigration using long differences, generally employing Census data, Pugatch and Yang examine year-over-year changes in immigration using data from the Current Population Survey. While this adds granularity to their analysis, it also identifies wage and employment effects that are often based on extremely small treatment dosages as it is typical for U.S. states or

[4]In order to link migration from a given Mexican state to a given U.S. state, Pugatch and Yang construct measures of regional migration patterns that developed over time in response to the construction of early 20th century railroads.

cities to add only a small fraction of a percentage point to their foreign-born Mexican share in a given year. More fundamentally, the short- and long-run effects of immigration might be quite different.

In this research, we estimate the effect of immigration on natives' labor market outcomes using a novel instrument that addresses the potential validity concerns over the network instrument while preserving its generalizability and comparability to the remainder of the area studies literature. Our innovation is to decompose the network instrument into a portion whose variation might be endogenous to destination labor market conditions and a portion that is not. In particular, we show that the level of emigration from Mexico is, in large part, a function of the size of lagged birth cohorts. Using lagged birth cohort sizes rather than the volumes of national emigration flows as a source of identifying variation eliminates the potentially endogenous component of the network instrument. The size of decades-lagged birth cohorts is assumed to be entirely exogenous to destination labor market conditions whereas the overall contemporaneous national emigrant flow may not be. Mexican state lagged birth cohort sizes are interacted with data on pre-determined network-linked Mexican state-U.S. MSA migration relations. The intuition behind the instrument is that differential temporal variation in the size of birth cohorts in different Mexican states isolates quasi-random variation in the assignment of Mexican immigrants to U.S. cities linked through historically determined migration networks to those states. After establishing the instrument's power to predict Mexican migrant share changes in a sample of seventy-six large U.S. MSAs, we use three Census cross-sections and a differencing strategy to estimate the contribution of Mexican immigration to changes in labor market outcomes among U.S. natives. We find little evidence of a net effect on the wages, unemployment rates or labor force participation rates of U.S. natives in any of sixteen age-skill groups. The results are precisely estimated allowing us to rule out anything other than very small effects.

Our births instrument explains approximately 20 percent of the variation in the network instrument, suggesting that at least one fifth of the variation in the network instrument is explained by a plausibly exogenous factor. As it turns out, IV estimates using the network instrument and those which use the births instrument lead to similar point estimates, though, if anything, the network instrument produces slightly *less* sanguine estimates of the effect of immigration on native labor market outcomes. We conclude that discrepancies between findings reported by Pugatch and Yang (2011) and the majority of the extant literature are likely explained by differences in the level of aggregation employed (states versus MSAs) or, alternatively, by differences in the short- and long-run effects of immigration on the labor market prospects of natives.

The remainder of the paper is organized as follows. Section II presents the theoretical foundations upon which the instrument is based, Section III provides a description of the econometric framework, includes a brief discussion of the identifying assumptions of the model and describes the data and sample. Section IV presents the empirical results and includes a discussion that links the results to those estimated in the prior literature. Section V concludes.

II. Identification Strategy

A. Motivation

The identification strategy we propose requires that variation in the size of lagged birth cohorts in Mexican states predict variation in U.S. cities' Mexican-born population shares. The mechanism underlying this relationship consists of two elements. First, lagged birth cohorts in Mexican states must be associated with higher volumes of emigration from those states. Second, emigration from each Mexican state must flow predictably into a distinctive set of U.S. destination cities, and the predictability of that flow must stem from factors that are otherwise unrelated to contemporary destination labor market conditions.

Straightforwardly, larger historical birth cohorts yield more potential migrants in Mexican sending states once each birth cohort reaches prime migration age. Therefore, all else equal, more potential migrants translates into more actual migrants unless birth cohort size is negatively associated with its members' propensity to migrate. On this point, we note that Hanson and McIntosh (2008) have documented that the rate of emigration is higher among members of larger Mexican birth cohorts, providing a second mechanism through which large lagged birth cohorts are positively associated with emigration. Large birth cohorts, they argue, yield a correspondingly large supply of workers once the cohorts reach working age, and surplus labor supply of comparable experience exerts negative pressure on the wages of the birth cohort's members. [5]

It is well established that migrants tend to travel to the same destinations that others from their source region have settled or sojourned in (Pugatch and Yang 2011; Munshi 2003; Massey 1999; Light 2006). The resulting networks confer social and informational benefits that can furnish tangible help in finding

[5] The decline of the U.S. baby boom two decades before the receding of the Mexican baby boom accentuated the potential benefits of migration for members of large birth cohorts.

work and housing. Importantly, many migration networks linking regions in Mexico to U.S. destinations were forged early in the 20th century, often tracking railroad routes along which U.S. employers brought in recruited agricultural labor (see, e.g., Cardoso 1980; Massey et al. 2002; and Woodruff and Zenteno 2007; Pugatch and Yang 2010; Chalfin 2012) [6] Thus, for reasons independent of prevailing labor market conditions in potential U.S. destinations, variation in push factors in different immigrant source regions will dependably influence the number of migrants located in different sets of U.S. destinations.

The identifying variation driving the births instrument emerges from the observation that different Mexican states experienced different levels of growth in the sizes of their birth cohorts at different times in the past century. This has been documented by Hanson and McIntosh (2008) using data from Mexican Censuses, and we begin by presenting corroborating evidence drawn from historical natality data from each Mexican state. **Figure 1** presents the time series of the changes in the number of births for each of the thirty-two states in Mexico. While, broadly speaking, Mexican birth cohort sizes tended to increase monotonically from 1930 until the 1970s, a careful review of these data reveals considerable variation, beyond the common national signal, among different source regions. For example, in Aguascalientes, the number of births increases in every decade until 1970, falling in 1980. However, in Chihuahua, growth in births peaks earlier in the 1950s. Guerrero, an important source region for cities in Texas and the Midwest experiences birth cohort sizes that are monotonically increasing, while Zacatecas, an important source region for cities in southern California peaks in 1960 and experiences sharp decreases in cohort size growth thereafter. If we posit that migration to network-linked U.S. destinations is a function of the birth cohort size-determined supply of prime age (17-52 year old) males in Mexico then in order to predict the number of Mexicans living in the U.S. in 1980, we would need an estimate of the number of Mexican males born between 1928 and 1963. [7] Likewise, the predicted number of Mexicans in the U.S. in 1990 will be a function of the number of Mexican males born

[6] In fact Pugatch and Yang (2011) use historical railroad routes as the basis for their calculation of migration network weights.

[7] Though this window reflects the age-range in which migration of Mexican males is most common, its precise upper bound is chosen for reasons of data availability. Natality data by state exist dating back to 1928 in Mexican government almanacs, defining the upper bound of the age window for the 1980 sample at 52 years old. For consistency, this upper bound is retained for the other years. This is not a serious concern because birth cohorts earlier than 1928 would have passed prime migration age by the late 1960s, prior to which most migration was seasonal and thus did not contribute to large-scale growth in the Mexican migrant population share. We have run all analyses shifting the window's lower and upper bounds one and two years earlier, shifting the lower bound one and two years later, and shifting only the upper bound one and two years earlier. None of these changes had a material effect on any of the results presented.

between 1938 and 1973 and the number of Mexicans in the U.S. in 2000 will depend on births between 1948 and 1983. Accordingly, the years which uniquely predict the change in the immigrant share of the population between 1980 and 1990 are 1928-1938 and 1963-1973 and the years which uniquely predict the change in the immigrant share between 1990 and 2000 are 1938-1948 and 1973-1983.

Using the above framework, **Table 1** shows the number of birth cohort size-determined eligible emigrants from each of Mexico's six largest migrant sending states in 1980, 1990, and 2000 and the percent change decade-to-decade. It is apparent at a glance that some states realized more dramatic growth in their eligible emigrant pools between 1980 and 1990, others between 1990 and 2000, and some showed similar growth in both periods. For example, while the number of eligible migrants from Jalisco was approximately 50,000 higher in 1990 than in 1980, in Zacatecas the number of eligible migrants declined considerably. The intuition behind our estimation strategy in this case is that U.S. destinations that are linked to Jalisco would be predicted to receive larger increases in Mexican immigration in the 1990s than in the 1980s, while the opposite pattern should be true of U.S. destinations that are linked to Zacatecas.

The empirical observation that there are important differences in the time series in fertility among Mexican sending states motivates an empirical framework in which we instrument for contemporary migrant flows using the size of birth cohorts in linked Mexican sending states. As we will show, this framework recognizes that we can decompose the network instrument into a component that is explained by variation in birth rates and a component that is a function of other, time-varying factors. The crux of our identification strategy is that birth cohort-induced migration through network-linked pathways partials out "good" variation in the network instrument. The remaining variation in the network instrument potentially taps the effects of economic conditions in U.S. receiving communities that would otherwise confound inferences as to the effect of changes in the migrant stock on U.S. natives' wages and employment rates.

B. Econometric Framework

Using data from three U.S. Censuses (1980, 1990 and 2000), we begin with a sample consisting of seventy-six metropolitan statistical areas and we generate an estimate of the proportion of each area's male labor force that is comprised of foreign-born Mexicans in a given Census year (ΔIMM_{st}). [8] By construction, ΔIMM can be disaggregated into the number of Mexicans who migrate to the United

[8] We use all MSAs for which we have sufficient data on the source region of migrants.

States from each of thirty-two Mexican states:

$$\Delta IMM_{it} = \sum_{m=1}^{32} \Delta IMM_{mit} \tag{1}$$

Thus, in (1), the change in the number of Mexicans living in city i in year t is simply the sum of Mexicans in that city in that year who migrated there from each of thirty-two Mexican states. Since ΔIMM_{mit} is not observable given currently available data on Mexican migrants living in the United States, it must be estimated. Following the functional relationship suggested by the traditional network instrument, we formulate ΔIMM_{it} as a function of the total number of Mexican migrants from each Mexican state who arrive in each year (ΔIMM_{mt}) and a set of Mexican state- U.S. city migration weights (P_{im}). Here, the weights are estimated using the mean probability that a migrant from Mexican state m migrates to each U.S. city using data from 1921-1979.[9] Equation (2) captures this relationship, with the inclusion of a time- and city-varying disturbance term that captures idiosyncratic shocks that are unrelated to the migration weights.

$$\Delta IMM_{it} = \sum_{m=1}^{32} (P_{im} * \Delta IMM_{mt}) + \epsilon_{it} \tag{2}$$

While the weights, P_{im} are strictly pre-determined, static, and likely reflect long-standing network ties that formed a century ago, ΔIMM_{mt} varies over time and potentially captures both economic conditions that push migrants out of Mexico as well as labor market conditions in linked U.S. labor markets. In order to further explore this term, we recognize that the change in the number of migrants from each Mexican state can be de-composed in the following way:

$$\Delta IMM_{it} = \left[\sum_{m=1}^{m=32} \sum_{t=t-17}^{t=t-52} BIRTHS_{mt} \right] \times Pr[MIG|BIRTH]_{mt} \tag{3}$$

In (3), the first term within the double summation is the number of births in network-linked Mexican states that occurred between 17 and 52 years ago. The second term is the average probability of migration conditional upon having been born between 17 and 52 years ago. This term varies both by Mexican state

[9] We choose 1979 as an end date to ensure that all of the migration relations contained in P_{im} are pre-determined with respect to the study sample.

and by year and, as a result, it is this term that creates a potential problem for the network instrument. In particular, the conditional probability of migration in year t will potentially be a function of push factors in Mexico as well as pull factors in network-linked U.S. cities. For example, if a particular city is experiencing positive wage growth over a given time period, this wage growth might increase the conditional probability of migration, thus building in a negative bias to the network instrument. Recognizing this, we re-formulate the network instrument in a way that partials out this potentially "bad" variation:

$$\Delta IMM_{it} = P_{im} \times \left[\sum_{m=1}^{m=32} \sum_{t=t-17}^{t=t-52} BIRTHS_{mt} \right] + \epsilon_{it} \tag{4}$$

In (4), for each of the thirty-two Mexican states, the time-invariant vector of migration weights to each state (P_{im}) is multiplied by the total number of births between years $t-17$ and $t-52$. Summing over all Mexican states, we obtain an estimate of the number of eligible migrants in network-linked states. Finally, we formulate the instrument by scaling this quantity by the size of a metropolitan area's labor force in 1980:

$$Z_{it} = \frac{\Delta IMM_{it} = P_{im} \times \left[\sum_{m=1}^{m=32} \sum_{t=t-17}^{t=t-52} BIRTHS_{mt} \right]}{LABF_{it=1980}} \tag{5}$$

Given our formulation of the instrument, equation (6) is a stylized representation of the first stage regression:

$$\Delta IMM_{it} = \alpha + \gamma \frac{\Delta IMM_{it} = P_{im} \times \left[\sum_{m=1}^{m=32} \sum_{t=t-17}^{t=t-52} BIRTHS_{mt} \right]}{LABF_{it=1980}} + \phi_t + \epsilon_{it} \tag{6}$$

Referring to (6) ψ_t represents year fixed effects which control for decadal migration shocks at the national-level. Since the Mexican share is differenced, we purge the model of between-MSA variation. In order to satisfy the requirement of instrument relevance, the instrument must predict the growth in the within-city Mexican population share that is not explained by national immigration trends. The corresponding model presented in (7) yields the relationship between the change in the outcome variable (either the log of the wage, the unemployment rate or the employment-to-population ratio, Y_{it}) and the change in the immigrant share in U.S. cities that is predicted by the size of lagged birth cohorts in network-linked Mexican states:

$$\Delta Y_{it} = \eta + \theta \Delta I\hat{M}M_{it} + \phi_s + \epsilon_{it} \tag{7}$$

The coefficient on the predicted immigrant share, θ, represents the effect of a one percentage point increase in a city's Mexican share on the change in a particular labor market outcome.

C. Identifying Assumptions and LATE

In order for the instrument to return a consistent estimate of a causal effect of Mexican immigration on the labor market outcomes of U.S. natives, the instrument must be both relevant and valid. We present evidence on instrument relevance in Section IV of the paper. Here, we briefly focus on the exclusion restriction. In order for the exclusion restriction to be met it must be the case that lagged Mexican birth cohorts affect the contemporary distribution of U.S. labor market outcomes *only* through their influence on the size of flows of Mexican migrants to the United States. We defend the exclusion restriction in the following ways. First, the long lag between changes in past Mexican fertility and contemporary measurement of wages and unemployment among U.S. natives provides some assurance that the two variables are not temporally confounded. As such, we can rule out reverse causation as a source of endogeneity.

The temporal lag between the two measurements also limits the causal pathways through which the variables can be related. For example, while changes in the labor market outcomes of U.S. migrants might plausibly be affected by contemporary Mexican birth rates (for example, if both were a direct function of the health of the Mexican economy), it is difficult to see how contemporary labor market outcomes should be related to economic conditions in Mexico several decades earlier other than through the push effect of births on migration. Nevertheless one potential confounder merits attention: the size of past Mexican state birth cohorts is correlated with the contemporaneous number of births to Mexican-born parents in linked U.S. destinations. To address this issue, in all subsequent analyses, we include a control for changes in the lagged number of U.S. births to Mexican-born parents in each city.

Second, the influence of other drivers of migration common to all U.S. cities for example periodic economic crises or shifts in U.S. immigration policy associated with the Immigration Reform and Control Act (IRCA) are netted out using year fixed effects. Notably, the instrument may interact with these drivers. Finally, while the network instrument is potentially compromised by persistent labor market shocks which are long-standing drivers of network-linked migration, the births instrument isolates a portion of the variation in the network instrument that is predicted by a "push" factor in Mexico, limiting the degree to which such persistence should cloud the resulting estimates.

Though any instrumental variables strategy identifies a local average treatment predicated on the behavioral response of "compliers," there are several reasons to expect that this instrument captures generalizable effects. First, the source of variation we consider accounts for a very large share of all Mexican immigration to the United States. Hanson and McIntosh (2008) assert that Mexican labor supply growth, generated by variation in the sizes of Mexican state birth cohorts over time, has accounted for at least 40 percent of all Mexican immigration to the United States. What's more, this 40 percent reflects only the higher propensity of members of larger birth cohorts to emigrate and does not take into account the simple fact that larger birth cohorts eventually produce larger cohorts of potential migrants. Second, economic motivations are likely to undergird much of the migration that introduces Mexican worker substitutes (or complements) for U.S. native workers into the United States. Mexican workers who migrate due to the push factor of downward pressures on Mexican wages are more likely to resemble those who are motivated to migrate due to the pull factor of growing opportunities in U.S. labor markets.

III. Data

Our data come from three sources. Data on state-specific births are drawn from tabulations of registered births and male-to-female birth ratios in Mexican states included in statistical almanacs produced by the Mexican government's *Instituto Nacional de Estadistica, Geografia e Informatica* (INEGI). [10] This is the most granular source of annual natality data in Mexico. To our knowledge, these data have not yet been used in research pertaining to U.S.-Mexico migration. Several potential sources of minor measurement error are worth noting. First, it is possible that not all births are registered, despite the Mexican government's assiduous efforts to accomplish full registration. Second, some births are registered in years after the birth actually occurred. For some years, the almanacs contain a break-down of late registrations in 1-2 year intervals, but this is not always available by state. Available data indicate that approximately 90 percent of births ever registered in most states are registered no

[10] Scanned copies of these almanacs are available on INEGI's website.

more than two years late, and that over 75 percent are registered in the year of occurrence. Since our identification strategy involves aggregating births over a thirty-five year interval, these errors should have only a minimal effect on our resulting estimates.

To construct a set of weights that capture migration patterns linking Mexican sending states to U.S. destination cities, we rely on migrant-level data from the Mexican Migration Project (MMP), administered jointly by Princeton University and the University of Guadalajara. The MMP surveys Mexican households in known sending regions and includes data on when migrants embarked on their first journeys to the United States, each migrant's state of birth, and the U.S. metropolitan area in which that migrant subsequently settled. Since we wish to construct weights that are pre-determined relative to the period whose migration trends we will predict, we include data only from pre-1980 self-reports. [11]

Data used to construct P_{im}, the matrix of Mexican state-U.S. city specific time- invariant migration weights were generated from the Mexican Migration Project's migrant level file. The file contains survey data on a sample of over 7,000 individuals, each of whom migrated to the United States at least once in their lifetime. The migrants are a subset of individuals who were sampled at random within each community sampled in the dataset. Each community was sampled once and individuals who reported having migrated to the United States were asked to recall each of their prior migration experiences. Among male household heads, 23 percent reported having migrated to the United States within three years of the time of survey, with 89 percent reporting an undocumented migration spell (Hanson 2002).[12] Using data on the U.S. destination for the migrants first migration episode, we remove from this file all migrants whose first migration experience occurred after 1979 and construct a matrix of weights that represent the average propensity of a migrant from a given Mexican state to migrate to each U.S. MSA in the dataset.[13] Thus, the weights were constructed from the migration experiences of 3,981 Mexican migrants. We begin in **Table 2A** by presenting the three most prevalent U.S. destinations for each Mexican state. The percentage of migrants who settled in each area is given in parentheses next to the

[11] We use data regarding the migrant's first journey to the United States rather than the last, both of which are available in the survey. The computed weights are not at all sensitive to this decision.

[12] Hanson further notes that the MMP surveys only households in which at least one member has remained in Mexico. As such, households that have entirely moved to the United States are not counted. Moreover, the migrants who are surveyed are a selected subset of migrants who have returned to Mexico, at least temporarily. For a detailed discussion of the MMP's migrant level file, see Hanson (2002).

[13] In principle, we could have used the migrant's last migration episode. However, the first migration experience is most likely to reflect network ties between the source and destination communities. In practice, the magnitudes of the elements of the matrix are virtually invariant to the choice of migration episode.

name of the metropolitan area. For example, the top two U.S. destinations for migrants from Baja California del Norte, located along the border with San Diego, CA are San Diego and Los Angeles. Likewise, the top three U.S. destinations for migrants from Nuevo Leon, a state in eastern Mexico are Houston, Dallas and McAllen, TX. While there is a fair amount of spread in the number of U.S. destinations in the dataset, the leading cities are predictably Los Angeles, Chicago, Houston, Dallas and San Diego. In **Table 2B**, we present data on each of the largest MSAs in our sample. Here, we see a large amount of variation, with each MSA relying on a markedly different combination of Mexican sending states. Consequently, weights derived from the MMP sample will indeed predict that large lagged birth cohorts in different Mexican states will increase the Mexican population share in different sets of U.S. cities.

Table 3 presents summary statistics on the three dependent variables examined in the paper — wages, the unemployment rate and the employment-to-population ratio. For each variable and each skill group, the table presents the mean and the standard deviation (disaggregated into the between- and within-unit variation) along with the values of the minimum and maximum observations.[14] Among the prime-age native-born males in our sample, those with less than a high school degree earned a wage of \$18.24 per hour while those with a high school degree, some college and a college degree earned, on average, \$20.27, \$23.21 and \$31.91, respectively. The majority of the variation in the wage is between-city variation. Turning to unemployment, the rate is highest (11.5 percent) among high school dropouts and lowest (2 percent) among college graduates, with a substantial amount of cross-sectional variation at the city level. Summary statistics on the employment-to-population ratio indicate that while 92 percent and 87 percent of college-educated male natives and those with some college, respectively, were in the labor force, just two thirds of those with less than a high school education were in the labor force. Data on the foreign-born Mexican share of U.S. MSA labor markets and labor market outcomes are derived from the U.S. Census for 1980, 1990 and 2000. **Table 4** provides details on the MSAs included in our sample and on the growth in each MSA's Mexican share over time. We compute each MSA's mean log wages, unemployment rate and employment-to-population ratio in each year for the sample of U.S. natives as a whole as well as for four education categories of natives – those who hold less than a high school degree, those who graduated high school, those with some college but no degree, and those with a college degree – and for four categories of natives' race/ethnicity white non-Hispanics, black non-Hispanics, other non-

[14] Wages are expressed in 2010 constant dollars.

Hispanics, and Hispanics. We also compute each outcome for each of sixteen age-education groups. As is standard in the literature on immigration and wages, we confine the sample to males aged 18-64 employed full time (or unemployed but seeking full time work) and not self-employed. The wage is computed by dividing annual wage and salary income by the product of usual weekly hours worked and the number of weeks worked in the previous year. Consistent with a strategy pursued by Altonji and Card (1991), for each of the Census years, we residualize each of our three outcomes by regressing the outcomes on an exhaustive set of MSA dummies and a flexible model of a city's age, race and education composition. This adjustment controls for changes in a city's composition which could potentially be correlated both with a change in the Mexican immigrant share and the change in labor market outcomes. [15]

IV. Results

We begin our discussion of the results by presenting evidence on the first stage relationship between growth in the Mexican population share and the lagged births instrument. In Panel A of **Table 5** we present regression evidence on the strength of the first stage. In column (1), we begin by presenting least squares estimates, using first differences, of the effect of the instrument on a city's Mexican population share for the 1980-1990 sample. Column (2) presents the same estimates using 1980 MSA population weights. In columns (3) and (4) we present coefficients from regressions on the 1990-2000 sample. Finally, in columns (5) and (6), we present estimates that use the all three years of Census data. The coefficients presented in Table 4 have a useful interpretation. Since both the instrument and the outcome variable are scaled by the city's population, the first stage coefficient can be interpreted as the estimated probability that a migration-eligible male birth can be found in a network-linked U.S. destination 17-52 years later. Between 1980 and 1990, approximately 6 percent of males born in the eligible window migrated to a network-linked U.S. destination. Between 1990 and 2000, this proportion fell to approximately 4 percent. [16] Overall, between 1980-2000, an estimated 5 percent of male births in Mexico ended up in migration network-linked cities in the United States. The instrument explains an extraordinarily high proportion of the variation (76 percent) in the immigrant share from 1980-1990 and a non-trivial proportion (10

[15] See p. 217 Altonji and Card (1991) for additional details.

[16] Again, this does not imply that the *number* of Mexican migrants to the United States declined. Rather it reflects a decline in the strength of the network linkages between Mexican states and U.S. destinations.

percent) in the 1990-2000 sample. This drop-off likely reflects the fact that, after the late 1980's – and possibly in part as a consequence of the 1986 Immigration Reform and Control Act and the economic saturation of traditional gateway labor markets – Mexican immigrants' reliance on traditional migration networks diminished dramatically (Light 2006; Massey 1999). These figures suggest that most immigration from Mexico to the U.S. during the 1980s was exogenous to contemporary labor market conditions in particular U.S. destinations. It is unclear whether this remained true during the 1990s. Aggregating the three Census years, the instrument overall explains nearly 30 percent of the variation in the immigrant share. In nearly all of the specifications, the F-statistic on the excluded instrument meets standard criteria for instrument relevance, though the predictive power of the instrument is consistently stronger in the weighted models. This is sensible as migration relations are measured with greater precision in the largest receiving cities for Mexican immigration. Since these cities also tend to possess very large populations, the instrument is strongest in the weighted models. In our preferred specification which uses population weights and all three Census cross-sections, the F-statistic on the excluded instrument is 27.9.

In Panel B of Table 5, we present regression evidence on the strength of the relationship between the traditional network instrument applied to Mexican states and the births instrument. While the relatedness of the two instruments varies by Census year, overall, the two candidate instruments are moderately correlated with the births instrument explaining approximately 20 percent of the variation in the network instrument in the full sample.

Table 6 presents both least squares and 2SLS results for models in which the mean log wage, unemployment rate, and employment-to-population ratio of U.S. natives are the dependent variables. Results are shown for the full sample as well as for four skill groups, disaggregated by educational attainment. In particular, separate regression coefficients are reported for individuals with less than a high school education, individuals whose highest level of education is a high school degree, individuals with some college and individuals with at least a four-year college degree. Columns (1), (3) and (5) present results for our first differences specification without the inclusion of MSA population weights while columns (2), (4) and (6) present results for the weighted models.

The first two columns of Table 6 present results for the effect of Mexican immigration on natives' mean log wages. Throughout Table 6, results are estimated for the full sample using population weights. Results are disaggegated by skill group and race. Beginning with the least squares results, we find some evidence that increases in Mexican immigration are associated with an increase in the wage of U.S. natives. A one percentage point increase in the Mexican population share is associated in the full sample of natives with a roughly three-quarters percent increase in natives' wages. [17] The positive association between immigration and wages holds for all skill and race groups, broadly reflecting the empirical tendency for migrants to settle in regions with higher wages. When the births instrument is employed, the pattern of results changes considerably. All of the coefficients shrink, and the sign on the coefficients for the low-skilled groups becomes negative. However, the estimated effects are extraodinarily small, with a one percentage point increase in a city's Mexican share predicting only a 0.02 percent change in natives' wages overall.

Columns (3) and (4) of Table 6 present results for natives' unemployment rates. In both the least squares models and the IV models there is little evidence of any substantively significant positive association between the change in the immigrant share and the change in the unemployment rate. In the OLS models, a one percentage point increase in the Mexican immigrant share is associated with a 0.05 percentage point increase in the unemployment rate among all natives with slightly larger but still very small and insignificant effects among high school graduates and blacks. In the IV models, the point estimates remain very close to zero for all groups, though they are largest in magnitude for natives with a high school diploma and for blacks.

In the final two columns of Table 6 we present evidence on the effect of Mexican immigration on the employment-to-population ratio among U.S. natives. OLS models indicate no meaningful association of the Mexican immigrant share with the employment rate for most groups. On the other hand, there is evidence of a modest but statistically significant positive effect on the employment rate of college graduates and of a larger disemployment effect among blacks. Results from the IV models corroborate estimates for the overall sample and for most subgroups. The estimated effect for the full sample of natives is a precisely estimated zero: we can reject even a quarter-percentage point immigration-driven decrease in the employment-to-population ratio with 95% confidence. However, instrumenting for immigration erases the negative effect on the black employment-to-population ratio. It still appears in the IV models that immigrants are complements to college-educated labor, as the coefficient for college graduates' employment rates more than doubles over the OLS estimate. This implies that a one percentage point increase in a

[17]In analyses not shown, we disaggregate the results by decade. The positive association was substantially more pronounced for the 1990-2000 changes than for the 1980-1990 changes, perhaps reflecting the greater endogeneity of the later period's immigrant location decisions to destination labor market conditions.

city's Mexican immigrant share yields a substantively significant 0.3 percentage point increase in college graduates' employment-to-population ratio. If immigration from Mexico has distributional effects on natives' labor market outcomes, these effects appear to emerge more from disproportionately positive effects on the outcomes of highly skilled natives than from negative effects on the outcomes of less skilled natives.

Tables 7, 8 and 9 further disaggregate effects on natives' wages, unemployment rates, and employment rates, respectively, into sixteen age-education groups, meant to reflect both skill and labor market experience. Though the estimates are less precise than those presented in Table 6, the pattern of results is strikingly consistent with the view that immigration from Mexico exerts at most a small negative influence on the labor market outcomes of less skilled and less experienced natives and, if anything, a positive influence on the outcomes of more skilled and more experienced natives.

Table 7 presents estimates of the effect of immigration from Mexico on the mean log wages of each skill-experience group of natives. The OLS models in the left panel indicate across-the-board positive associations of immigration with wages that are close to uniform across cells. Instrumenting turns many of these point estimates negative, though they are small and insignificant. However, positive estimates remain among the more skilled groups, and these are concentrated among those with more experience. Negative point estimates are concentrated among less skilled and less experienced natives, particularly those aged 19-25 with a high school education or less and, to a degree, less skilled workers aged 41-55.

Turning to the unemployment rate,**Table 8** indicates a similar pattern of results. Both OLS and IV models imply that immigrants are substitutes for less skilled and less experienced natives and are complements for more skilled and more experienced natives. Only the point estimates for the youngest and least skilled workers suggest more than a vanishing effect on unemployment, and there is, if anything, a negative effect on the unemployment rate of older and more educated natives. This pattern repeats in **Table 9**, which describes the effect of immigration on natives' employment-to-population ratios. The OLS results show a substantial negative association between immigration and younger, less skilled natives' employment rates. IV results show a greatly diminished effect, though the coefficients remain generally negative in the low-skilled, low-experience upper left hand corner of the right panel of the table. For natives with more experience and more education, OLS models suggest a positive but small and insignificant association of immigration from Mexico with the employment rate. Instrumenting, however, shows that the causal effect of immigration on the employment rates of older and more educated natives is positive, significant, and, for the most educated groups, quite large — a one percentage point increase in the migrant share is estimated to lead to a one percentage point increase in the labor force participation rate for older college-educated natives.

It is natural to compare our results directly to those that would have been obtained in our sample using a Mexico-specific version of the network instrument employed in Altonji and Card (1991) and Card (2001). Ideally, with data on aggregate migration flows from each Mexican state to the U.S., and with data on the existing distribution across cities of Mexican migrants from each Mexican state, we could replicate a Mexican state-specific version of the network instrument that would parallel the births instrument but include variation that we have argued could reflect destination pull factors and thus bias the causal estimates. Lacking these data, we construct the network instrument by interacting the existing distribution of Mexican migrants across cities in 1980 with the overall decadal net influx of Mexican immigration into the United States from 1980-1990 and from 1990-2000. In **Table 10**, we present our IV estimates for the weighted 1980-2000 sample alongside the results that are obtained using the network instrument. Though the estimated coefficients are less precisely estimated using the network instrument, which is not surprising given that it is based on far less source-geography granularity, the coefficients are extremely similar in magnitude. Though few results are statistically significant, the network instrument returns estimates of the effect of immigration on natives' labor market outcomes that are substantially more concerning than those the births instrument returns. This is especially true for low-skilled natives. Ironically, addressing the latest critique of the network instrument only makes any negative effects of immigration from Mexico on natives' labor market outcomes appear even smaller.

V. Conclusion

We have proposed a novel instrument for changes in the Mexican share of U.S. city labor forces, demonstrated its predictive power, and used it to estimate the effect of Mexican immigration on the wages and employment of U.S. natives. The validity of the instrument is rooted in two well-understood features of Mexican migration to the United States. Increases in the sizes of state-specific Mexican birth cohorts are associated with increased emigration and perhaps also decreased return migration. Migrants from each Mexican sending state, in turn, tend to settle or sojourn in particular U.S. destination cities, a persistent relic of networks initiated in early 20th century labor recruitment patterns. Together, these two phenomena allow us to predict decennial changes in the Mexican share of a U.S. city's labor force. Our findings indicate little evidence of meaningful effects of Mexican wages on wages or unemployment rates among U.S. natives apart from positive effects on the employment of more skilled and more experienced natives.

For several reasons, our estimation strategy is a promising advance over prior econometric techniques used to estimate the effects of immigration on U.S. natives' employment outcomes. By partialing out variation in the network instrument which is explained by factors other than the size of state-specific Mexican birth cohorts, the instrument is grounded in historical variation in Mexico rather than solely in "pull factors in the U.S. destination region. This innovation reduces the chances that serial correlation in destination city conditions that have historically drawn migrants and continue to draw migrants today, confounds the recovery of a consistent estimate of the effect of Mexican immigration. Finally, given that lagged birth cohort sizes appear to account for a very large share of total growth in the Mexican migrant shares of U.S. cities in recent decades, the treatment effects we identify, while local, should be more broadly applicable to debates over immigration's effects on the United States than estimates derived from instruments based on phenomena that account for only a small, and perhaps selective, share of aggregate emigration.

Despite evidence that our instrument is more plausibly exogenous than the network instrument, we replicate the generally minimal effects found in most of the area studies literature. Regarding discrepancies between the findings of Altonji and Card (1991), Card (2001) and Card and Lewis (2005) and those of Pugatch and Yang (2011), our best guess is that these are explained by differences in the short-run and long-run effects of migration, by differences in the effect in states versus MSAs, or by an unrepresentative LATE of weather-driven migration.

References

Abadie, Alberto, Alexis Diamond and Jens Hainmuller. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of Californias Tobacco Control Program," *Journal of the American Statistical Association*, 105 (2010): 493-506.

Altonji, Joseph G. and David Card. "The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives," in *Immigration, Trade and the Labor Market*, edited by John Abowd and Richard Freeman. Chicago: University of Chicago Press, 1991.

Borjas, George J. "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market,"]textitQuarterly Journal of Economics, 118 (2003): 1335-1374.

Card, David. "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review* 43 (1990): 245-257.

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

Card, David and Ethan G. Lewis. "The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts," in *Mexican Immigration to the United States*, edited by George J. Borjas. Chicago: University of Chicago Press, 2007.

Cardoso, Lawrence A. Mexican Emigration to the United States, 1897-1931: Socioeconomic Patterns. Tucson, Arizona: University of Arizona Press, 1980.

Chalfin, Aaron. "What is the Contribution of Mexican Immigration to U.S. Crime Rates? Evidence from Rainfall Shocks in Mexico," Working Paper (2012).

Hanson, Gordon and Craig McIntosh. "The Great Mexican Emigration," *Review of Economics and Statistics* 92(4) (2010): 798-810.

Light, Ivan. Deflecting Immigration: Networks, Markets, and Regulation in Los Angeles. New York: Russell Sage Foundation, 2006.

Massey, Douglas S. "International Migration at the Dawn of the Twenty-First Century: The Role of the State," *Population and Development Review* 25 (2) (1999): 303-322.

Massey, Douglas S., Jorge Durand, and Nolan J. Malone. Beyond Smoke and Mirrors: Mexican Immigration in an Era of Economic Integration. New York: Russell Sage Foundation, 2002.

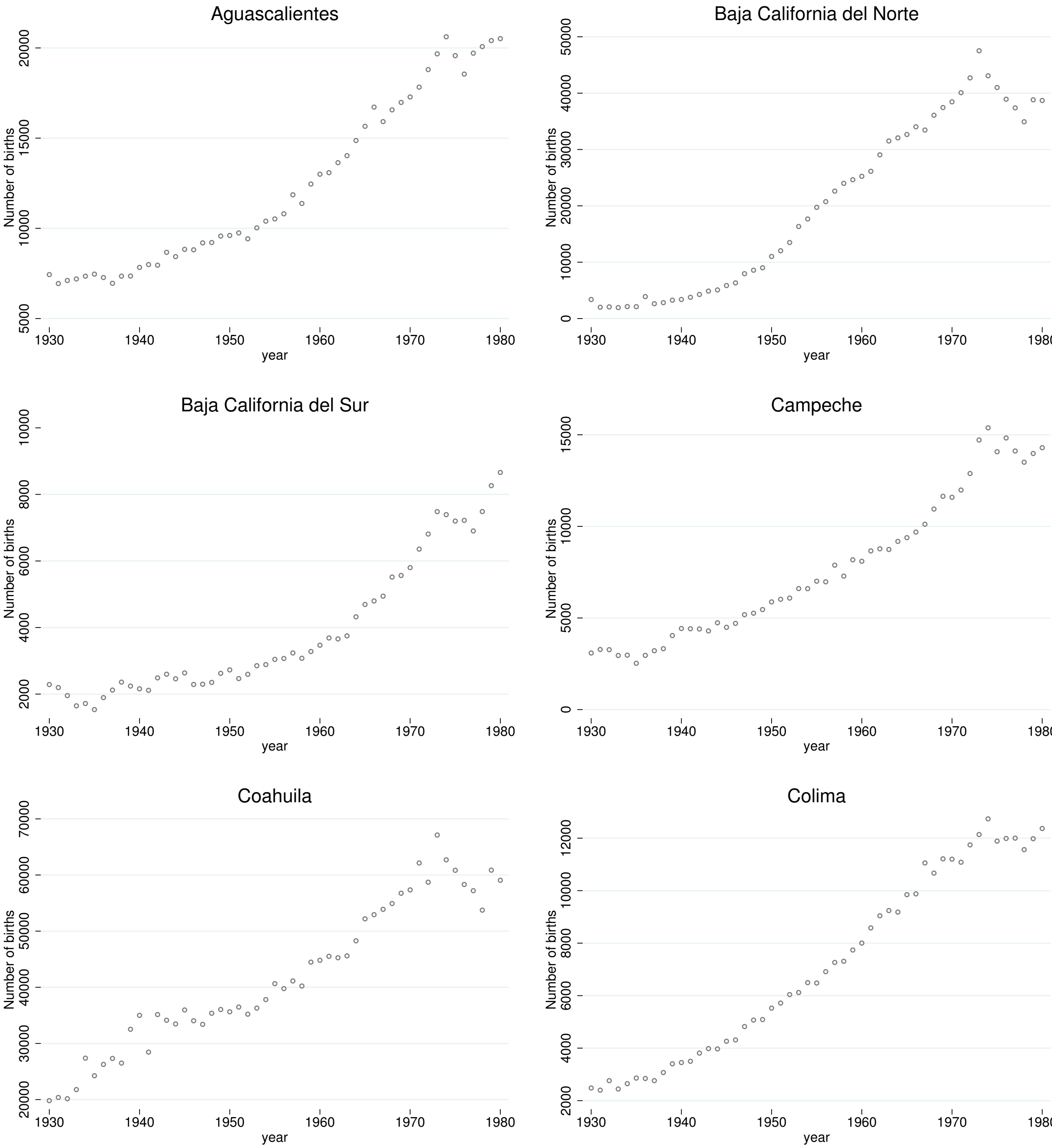
Munshi, Kaivan. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics* 118(2) (2003): 549-597.

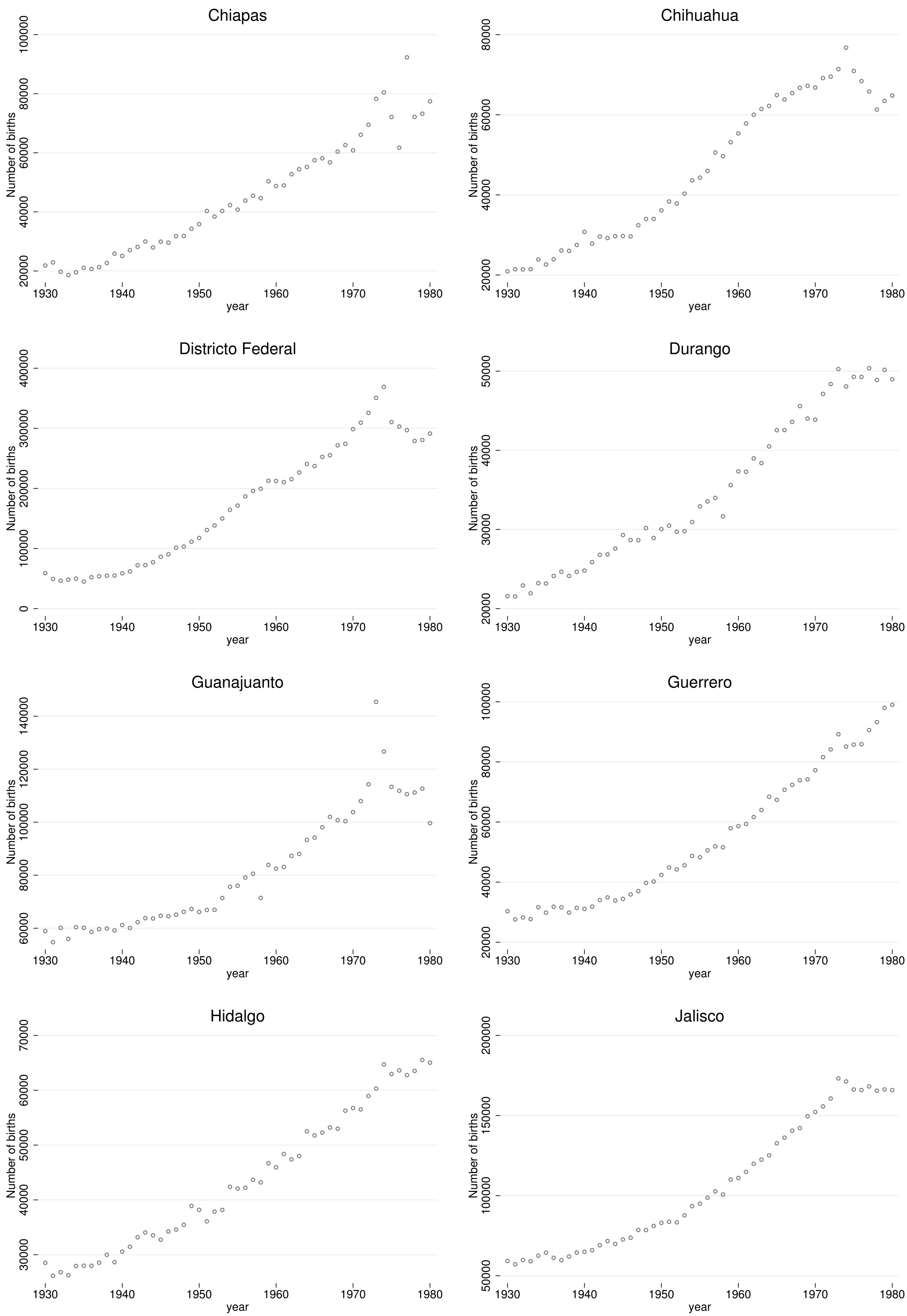
Pugatch, Todd and Dean Yang. "The Impact of Mexican Immigration on U.S. Labor Markets:

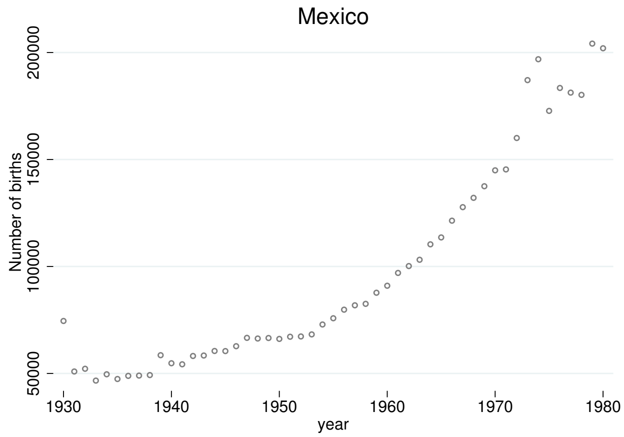
Evidence from Migrant Flows Driven by Rainfall Shocks," Working Paper (2010).

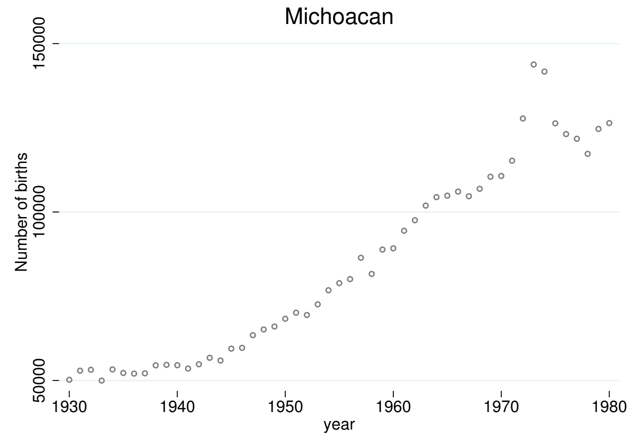
Woodruff, Christopher and Rene Zenteno. "Migration networks and Microenterprises in Mexico." *Journal of Development Economics* 82 (2007): 509-528.

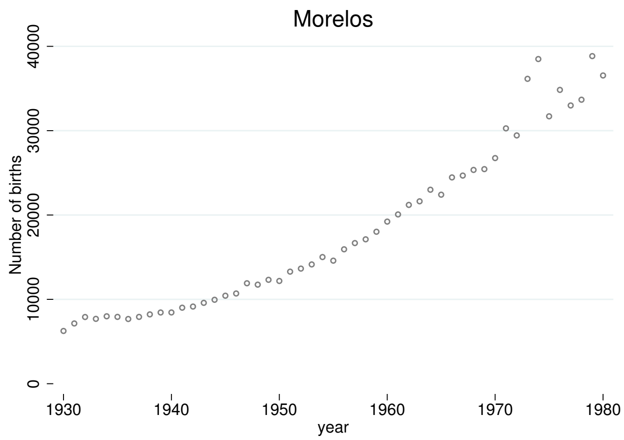
Figure 1. Number of Births by Mexican State

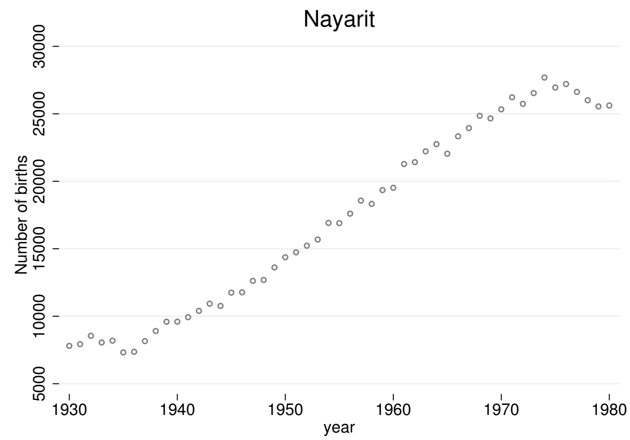


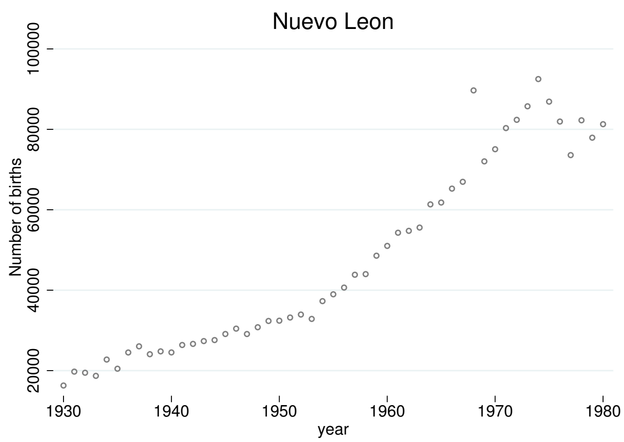

Chiapas
Chihuahua
Districto Federal
Durango
Guanajuanto
Guerrero
Hidalgo
Jalisco
Number of births
year

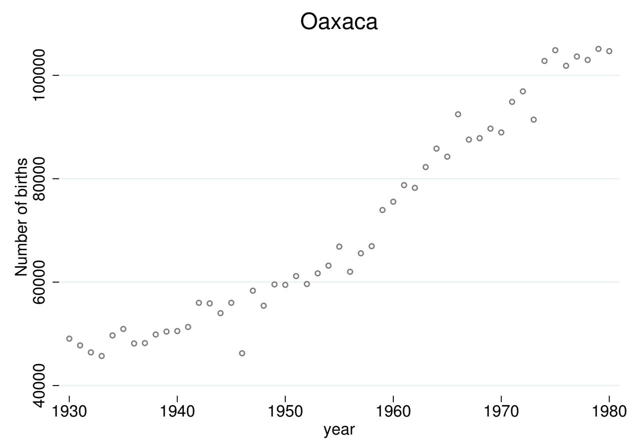

Mexico
Number of births
50000
100000
150000
200000
1930
1940
1950
1960
1970
1980
year

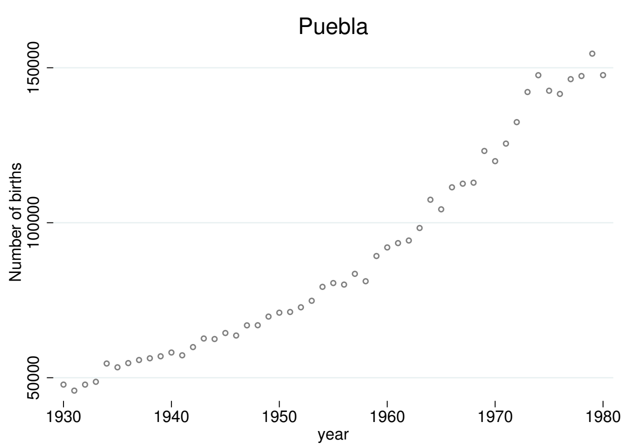

Michoacan
Number of births
50000
100000
150000
1930
1940
1950
1960
1970
1980
year

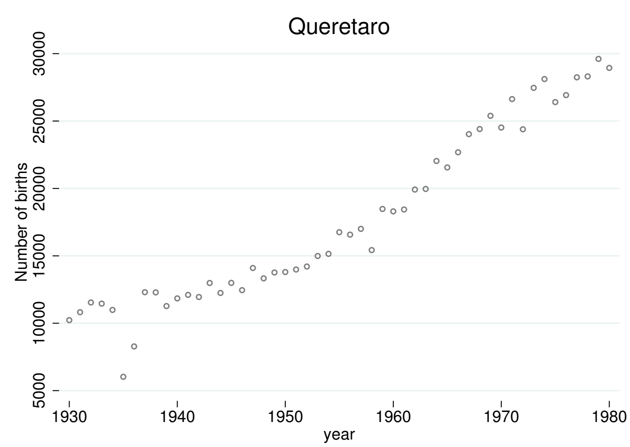

Morelos
Number of births
0
10000
20000
30000
40000
1930
1940
1950
1960
1970
1980
year

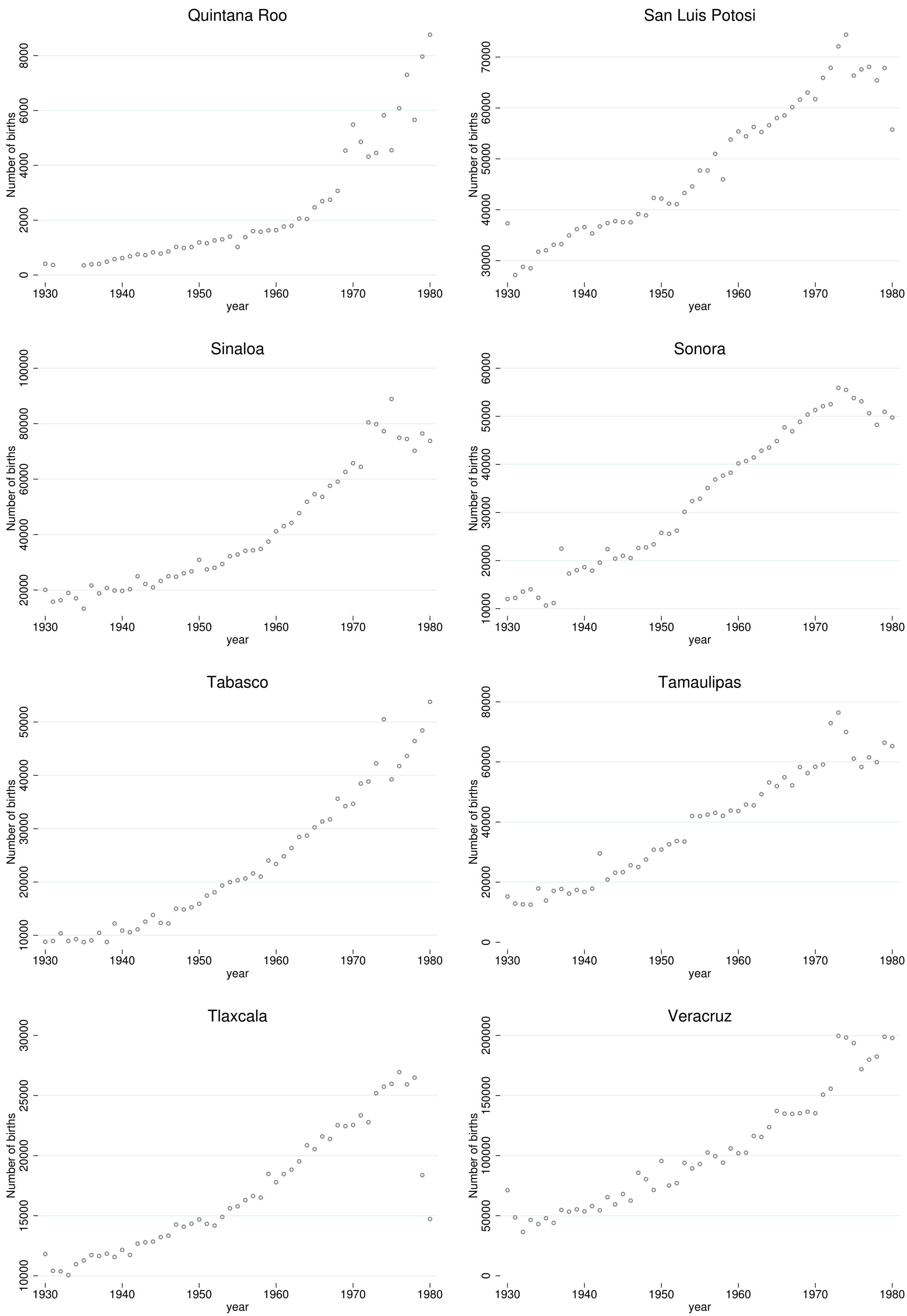

Nayarit
Number of births
5000
10000
15000
20000
25000
30000
1930
1940
1950
1960
1970
1980
year

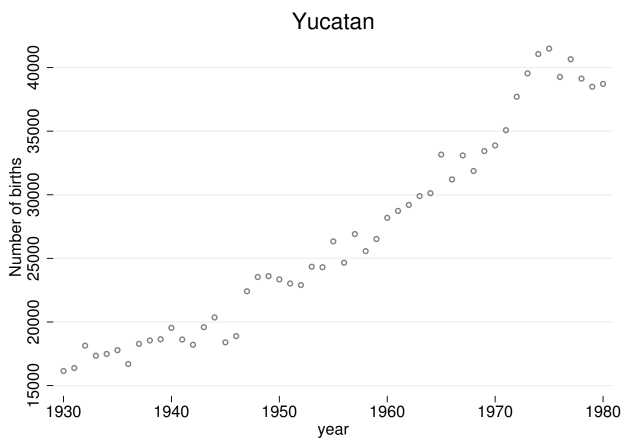

Nuevo Leon
Number of births
20000
40000
60000
80000
100000
1930
1940
1950
1960
1970
1980
year


Oaxaca
Number of births
40000
60000
80000
100000
1930
1940
1950
1960
1970
1980
year


Puebla
Number of births
50000
100000
150000
1930
1940
1950
1960
1970
1980
year


Queretaro
Number of births
5000
10000
15000
20000
25000
30000
1930
1940
1950
1960
1970
1980
year


Quintana Roo
Number of births
0 2000 4000 6000 8000
1930 1940 1950 1960 1970 1980
year
San Luis Potosi
Number of births
30000 40000 50000 60000 70000
1930 1940 1950 1960 1970 1980
year
Sinaloa
Number of births
20000 40000 60000 80000 100000
1930 1940 1950 1960 1970 1980
year
Sonora
Number of births
10000 20000 30000 40000 50000 60000
1930 1940 1950 1960 1970 1980
year
Tabasco
Number of births
10000 20000 30000 40000 50000
1930 1940 1950 1960 1970 1980
year
Tamaulipas
Number of births
0 20000 40000 60000 80000
1930 1940 1950 1960 1970 1980
year
Tlaxcala
Number of births
10000 15000 20000 25000 30000
1930 1940 1950 1960 1970 1980
year
Veracruz
Number of births
0 50000 100000 150000 200000
1930 1940 1950 1960 1970 1980
year


Yucatan
Number of births
15000
20000
25000
30000
35000
40000
1930
1940
1950
1960
1970
1980
year

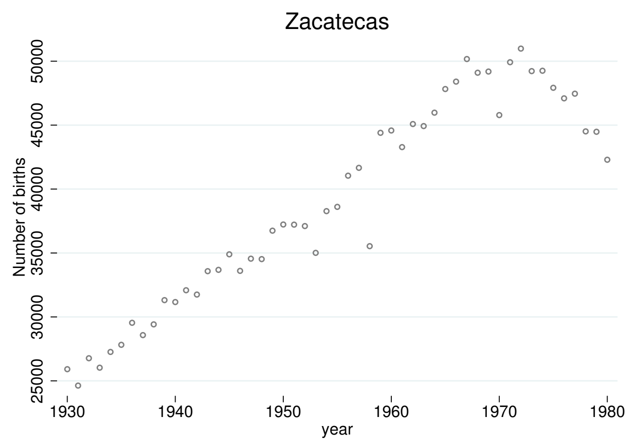

Zacatecas
Number of births
25000
30000
35000
40000
45000
50000
1930
1940
1950
1960
1970
1980
year

Table 1. Migration-Eligible Births by Mexican State, 1980-2000

Year	State	Number of Migration-Eligible Births (thousands)	Change in Migration-Eligible Births (thousands)
1980	Durango	526.2	
1990		642.1	115.9
2000		747.1	105.0
1980	Guanajuanto	1235.6	
1990		1488.1	252.5
2000		1753.6	265.5
1980	Jalisco	1451.0	
1990		1903.3	452.3
2000		2397.3	494.1
1980	Michoacan	1200.6	
1990		1536.6	336.0
2000		1880.4	343.8
1980	San Luis Potosi	739.8	
1990		905.9	166.1
2000		1048.4	142.5
1980	Zacatecas	623.0	
1990		741.3	118.4
2000		800.2	58.8

Note: The table reports both the number of migration-eligible births and the change in the number of migration-eligible births by decade for the six most prominent states of origin among Mexican migrants to the United States.

Table 2A. U.S. Destinations of Mexican Immigrants

Mexican State	Destination #1	Destination #2	Destination #3
Aguascaliente	Los Angeles (20%)	Reno (6%)	Tulsa (6%)
Baja California del Notre	San Diego (60%)	Los Angeles (22%)	
Baja California del Sur			
Campeche			
Coahuila de Zaragoz			
Colima	Los Angeles (41%)	Fresno (9%)	
Chiapas			
Chihuahua	El Paso (16%)	Los Angeles (9%)	Dallas/Phoenix (9%)
Districto Federal	Los Angeles (20%)	Chicago (11%)	Orange County (CA) (8%)
Durango	Chicago (23%)	Los Angeles (19%)	Dallas (7%)
Guanajuanto	Los Angeles (15%)	Chicago (11%)	Houston (7%)
Guerrero	Chicago (29%)	Los Angeles (15%)	Phoenix (12%)
Hidalgo	Las Vegas (12%)	Dallas (9%)	Houston (7%)
Jalisco	Los Angeles (26%)	San Diego (6%)	San Jose (4%)
Mexico (Estado)	Chicago (32%)	Stockton (10%)	los Angeles (7%)
Michoacan	Los Angeles (20%)	Fresno (8%)	Chicago 96%)
Morelos	Los Angeles (29%)	Minneapolis (18%)	Chicago (10%)
Navarit	Los Angeles (29%)	San Jose (10%)	Orange County (CA) (7%)
Nuevo Leon	Houston (16%)	McAllen (15%)	Dallas (11%)
Oaxaca	Los Angeles (51%)	San Diego (9%)	
Puebla	New York (56%)	Los Angeles (23%)	
Querataro			
Quintana Roo			
San Luis Potosi	Houston (16%)	San Diego (16%)	Dallas (6%)
Sinaloa	Los Angeles (48%)	San Diego (10%)	Riverside (8%)
Sonora			
Tamaulipas			
Tabasco			
Tiaxcala	Los Angeles (9%)		
Veracruz	Los Angeles (14%)	Chicago (13%)	San Jose (8%)
Yucatan	Portland (31%)	San Francisco (29%)	Los Angeles (11%)
Zacatecas	Los Angeles (28%)	Fresno (5%)	Merced (5%)

Note: The table reports the three largest U.S. metropolitan area destinations for migrants from each Mexican state, among migrants in the Mexican Migration Project's Migrant File, 1921-1979.

Table 2B. Mexican State Sources of U.S.-Bound Immigrants
Selected U.S. Metropolitan Areas

U.S. MSA	Source #1	Source #2	Source #3
Atlanta	Jalisco (23%)	Nuevo Leon (12%)	Veracruz (11%)
Austin-San Marcos	San Luis Potosi (33%)	Veracruz (26%)	Guerrero (21%)
Chicago	Durango (30%)	Jalisco (25%)	Guanajuanto (19%)
Dallas	Guanajuanto (28%)	Durango (26%)	Jalisco (11%)
Denver	Yucatan (58%)	Chihuahua (14%)	Districto Federal (7%)
El Paso	Chihuahua (64%)	Zacatecas (9%)	Veracruz (5%)
Fresno	Jalisco (44%)	Michoacan (15%)	Guanajuanto (14%)
Houston	San Luis Potosi (50%)	Guanajuanto (15%)	Michoacan (7%)
Las Vegas	Jalisco (43%)	Nayarit (14%)	Districto Federal (13%)
Los Angeles-Long Beach	Jalisco (23%)	Michoacan (10%)	Guanajuanto (9%)
Merced	Nayarit (43%)	Jalisco (23%)	Michoacan (18%)
Minneapolis-St. Paul	Morelos (100%)		
New York	Puebla (56%)	Morelos (22%)	Tlaxcala (5%)
Oakland	Jalisco (58%)	Michoacan (36%)	Districto Federal (2%)
Orange County (CA)	Jalisco (25%)	Guerrero (20%)	Guanajuanto (13%)
Philadelphia	Guanajuanto (91%)	Districto Federal (4%)	
Phoenix	Chihuahua (30%)	Guanajuanto (16%)	Durango (12%)
Portland	Yucatan (91%)		
Riverside-San Bernardino	Michoacan (22%)	Jalisco (20%)	Yucatan (9%)
San Diego	Baja California del Norte (61%)	San Luis Potosi (16%)	Jalisco (7%)
San Francisco	Yucatan (54%)	Jalisco (13%)	Nayarit (10%)

Note: The table reports the three most prevalent source regions among Mexican immigrants to selected U.S. metroplitan areas. These data are based upon the experiences of migrants surveyed in the Mexican Migration Project's Migrant File, 1921-1979.

Table 3. Summary Statistics on Natives' Wage and Employment Outcomes

Variable		**Mean**	**S.D.**	**Min.**	**Max.**
Panel A. Wages					
< High School	O	$18.24	$2.55	$11.95	$25.24
	B		$2.03		
	W		$1.55		
High School	O	$20.27	$2.28	$13.29	$25.74
	B		$2.05		
	W		$1.02		
Some College	O	$23.21	$2.45	$16.91	$31.95
	B		$2.29		
	W		$0.93		
College +	O	$31.91	$3.68	$22.84	$46.53
	B		$2.88		
	W		$2.30		
Panel B. Unemployment %					
< High School	O	11.5	3.9	3.5	21.1
	B		3.2		
	W		2.3		
High School	O	6.6	2.4	2.4	15.1
	B		1.9		
	W		1.5		
Some College	O	4.2	1.6	1.4	1.3
	B		1.2		
	W		1.1		
College +	O	2.0	0.8	0.4	6.1
	B		0.6		
	W		0.5		
Panel C. Employment %					
< High School	O	66.9	7.5	43.7	84.3
	B		5.6		
	W		5.0		
High School	O	81.8	5.3	65.9	92.2
	B		4.0		
	W		3.6		
Some College	O	87.3	4.1	73.1	94.8
	B		3.3		
	W		2.5		
College +	O	92.1	3.6	74.1	97.7
	B		3.1		
	W		1.8		

Note: This table reports descriptive statistics for the mean wage, the unemployment rate and the employment-to-population ratio for the seventy-six MSAs in our sample. For each variable, we report the overall mean, the standard deviation decomposed into overall ("O"), between ("B"), and within ("W") variation, as well as the minimum and maximum values.

Table 4. Foreign-Born Mexican Population Share
U.S. Metropolitan Areas, 1980-2000

Metropolitan Area	Primary State	1980	1990	2000
Albuquerque	NM	1.6	3.3	6.1
Allentown	PA	0.0	0.1	0.3
Atlanta	GA	0.1	0.9	5.1
Austin-San Marcos	TX	1.5	4.1	9.5
Bakersfield	CA	7.5	13.1	19.5
Beaumont-Port Arthur	TX	0.5	1.4	3.7
Benton Harbor	MI	0.2	0.2	1.5
Boston	MA	0.0	0.1	0.2
Brownsville	TX	21.3	26.6	29.2
Charlotte	NC	0.0	0.2	4.0
Chicago	IL	3.7	6.8	10.3
Chico	CA	2.1	3.8	4.8
Cleveland	OH	0.0	0.0	0.4
Colorado Springs	CO	0.2	0.3	2.2
Corpus Christi	TX	2.9	4.3	3.4
Dallas	TX	2.6	7.1	14.6
Denver	CO	0.9	1.9	7.4
Detroit	MI	0.2	0.2	1.0
Eugene	OR	0.1	0.7	1.8
Fort Myers	FL	0.4	1.5	5.1
Fort Wayne	IN	0.1	0.2	1.5
Fresno	CA	8.8	17.4	23.1
Galveston	TX	1.9	2.6	5.5
Grand Rapids	MI	0.3	1.0	2.8
Greenboro-Winston Salem	NC	0.0	0.2	4.5
Houston	TX	5.1	9.4	14.0
Indianapolis	IN	0.1	0.1	1.7
Kansas City	MO	0.3	0.6	2.5
Lakeland-Winter Haven	FL	0.5	1.6	4.9
Las Vegas	NV	1.5	4.1	11.1
Lexington	KY	0.1	0.1	2.4
Los Angeles-Long Beach	CA	12.6	19.5	21.7
Lubbock	TX	1.0	1.8	2.0
McAllen	TX	22.7	30.8	37.1
Medford	MA	1.2	2.7	4.0
Miami	FL	0.3	0.8	1.6
Milwaukee	WI	0.4	0.7	2.8
Minneapolis-St. Paul	MN	0.1	0.2	1.6
Modesto	CA	6.0	11.5	17.1
Nashville	TN	0.0	0.1	2.7
New Orleans	LA	0.1	0.1	0.5
New York	NY	0.2	1.0	2.8
Ocala	FL	0.0	0.5	1.8
Oklahoma City	OK	0.5	1.5	4.0
Omaha	NE	0.4	0.6	3.4
Orlando	FL	0.3	0.5	2.0
Philadelphia	PA	0.1	0.2	0.5
Phoenix-Mesa	AZ	2.3	5.3	12.4
Pittsburgh	PA	0.0	0.0	0.1
Portland-Vancouver	OR	0.2	1.6	5.1
Pueblo	CO	0.7	1.2	2.8
Raleigh-Durham-Chapel Hill	NC	0.0	0.4	5.3
Reading	PA	0.0	0.7	1.2
Reno	NV	1.3	6.4	8.7
Riverside-San Bernardino	CA	3.6	11.9	17.5
Rockford	IL	0.6	1.7	4.7
Sacramento	CA	2.1	3.1	5.0
St. Louis	MO	0.1	0.1	0.4
Salinas	CA	11.5	18.1	28.8
Salt Lake City-Ogden	UT	0.4	1.0	5.0
San Antonio	TX	5.5	7.8	9.7
San Diego	CA	5.2	9.3	11.9
San Francisco	CA	2.3	4.6	5.8
San Jose	CA	3.6	6.9	10.6
Santa Rosa	CA	2.6	6.2	11.4
Seattle	WA	0.2	0.4	2.3
Stockton-Lodi	CA	6.2	11.5	16.3
Tacoma	WA	0.2	0.3	1.6
Tampa-St. Petersburg	FL	0.2	0.8	2.0
Tucson	AZ	3.6	6.3	8.7
Tulsa	OK	0.1	0.8	3.5
Visalia-Tulare-Porterville	CA	11.2	22.6	28.6
Washington	DC	0.1	0.4	1.1
West Palm Beach-Boca Raton	FL	0.4	1.7	3.0
Wichita	KS	0.7	1.0	3.6
Wilmington-Newark	DE	0.0	0.4	1.8

Note: Each column reports, for a given decade, each MSA's foreign-born Mexican share. Data were tabulated using the 1980, 1990 and 2000 U.S. Census.

Table 5. First Stage Models
Proportion of Migration-Eligible Births in
Network-Linked U.S. Metropolitan Statistical Areas

	(1)	(2)	(3)	(4)	(5)	(6)
	1980-1990		1990-2000		1980-2000	
Panel A. Effect of Births Instrument on Mexican Population Share						
Births cohort instrument	0.043***	0.056***	0.018**	0.035***	0.030***	0.054***
	(0.008)	(0.009)	(0.009)	(0.010)	(0.008)	(0.010)
F-statistic	28.7	41.6	4.1	11.9	14.2	27.9
R^2	0.401	0.759	0.078	0.101	0.220	0.293
Panel B. Effect of Births Instrument on the Network Instrument						
Births cohort instrument	0.217***	0.179***	0.447***	0.358***	0.334***	0.258***
	(0.019)	(0.026)	(0.062)	(0.068)	(0.058)	(0.046)
Partial R^2	0.171	0.171	0.177	0.177	0.203	0.203
1980 Population Weights	no	yes	no	yes	no	yes

Note: Each column reports results of a least squares regression of the change in an MSA's foreign-born Mexican population share on the change in the predicted foreign-born Mexican share as informed by the number of eligible lagged births in network-linked Mexican states. Columns (1)-(2) report estimates obtained via first differences for the 1980-1990 sample while columns (3)-(4) report first differenced estimates 1990-2000 sample. In columns (5)-(6) we estimate the models via first differences using the entire sample. Columns (1), (3) and (5) report coefficient estimates and standard errors that are estimated via ordinary least squares while columns (2), (4) and (6) report estimates generated via WLS, weighted according to 1980 MSA population. For each time period, the coefficient is can be interpreted as the proportion of migration-eligible births in linked Mexican states who migrate to the United States. The F-statistic that we report is the square of the t-statistic on the birth cohort instrument and is a sufficient statistic to assess the strength of the first stage relationship between Mexican immigration and the birth cohort instrument. The sample size for each set of first differenced regressions is 76 cities. In columns (5) and (6), the sample size is 152 city-years. We report heteroskedasticity-robust standard errors in parentheses below the coefficient estimates. Statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6. The Effect of Mexican Immigration on the Wages and Employment Outcomes of U.S. Natives

	(1)	(2)	(3)	(4)	(5)	(6)
	Wages		Unemployment %		Employment %	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
Overall	0.787*	0.020	0.054	0.065	-0.027	0.000
	(0.426)	(0.691)	(0.073)	(0.136)	(0.088)	(0.197)
< High School	0.922**	-0.190	0.143	0.042	-0.124	0.111
	(0.413)	(0.737)	(0.117)	(0.260)	(0.145)	(0.367)
High School	0.866*	-0.110	0.131	0.182	-0.100	-0.119
	(0.467)	(0.841)	(0.094)	(0.174)	(0.110)	(0.241)
Some College	0.786*	0.160	0.008	0.095	0.001	-0.124
	(0.399)	(0.686)	(0.070)	(0.127)	(0.083)	(0.207)
College +	0.747**	0.344	-0.041	-0.060	0.128*	0.297***
	(0.314)	(0.484)	(0.039)	(0.049)	(0.070)	(0.115)
White	0.702*	-0.122	0.056	0.039	0.007	0.026
	(0.386)	(0.650)	(0.063)	(0.115)	(0.083)	(0.184)
Black	1.091**	-0.355	0.176	0.205	-0.313**	-0.045
	(0.447)	(1.072)	(0.136)	(0.273)	(0.145)	(0.338)
Hispanic	0.810*	-0.089	0.060	-0.066	-0.071	-0.124
	(0.444)	(0.714)	(0.101)	(0.222)	(0.193)	(0.345)

Note: Each column reports results of either a least squares regression or a corresponding IV regression of the change in a given outcome (the log wage, the unemployment rate or the employment-to-population ratio) on the change in the (instrumented) Mexican population share, conditional on the change in the number of U.S. births and year fixed effects. Columns (1)-(2) report coefficient estimates for the log wage, columns (3) and (4) report estimates for the unemployment rate and columns (5) and (6) report estimates for the employment-to-population ratio. All models are weighted by 1980 MSA population. As they are conservative relative to standard errors that are clustered at the MSA level, we report heteroskedasticity-robust standard errors in parentheses below the coefficient estimates. Statistical significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7. The Effect of Mexican Immigration on the Log Wages of U.S. Natives
Estimates by Age-Skill Groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS				2SLS			
	Age 19-25	Age 26-40	Age 41-55	Age 56+	Age 19-25	Age 26-40	Age 41-55	Age 56+
< High School	0.739	1.105**	0.824**	0.730	-0.558	0.084	-0.331	-0.334
	(0.529)	(0.443)	(0.376)	(0.461)	(1.032)	(0.779)	(0.726)	(0.861)
High School	0.990	1.135**	0.630*	0.401	-0.411	0.384	-0.454	-0.456
	(0.672)	(0.483)	(0.350)	(0.390)	(1.301)	(0.757)	(0.868)	(0.661)
Some College	0.935	0.809*	0.705***	1.147***	-0.497	-0.166	0.774	1.638*
	(0.604)	(0.443)	(0.236)	(0.401)	(1.068)	(0.755)	(0.497)	(0.931)
College +	0.694	0.725**	0.625***	1.045***	0.799	0.049	0.649	0.859
	(0.537)	(0.329)	(0.206)	(0.397)	(1.258)	(0.547)	(0.427)	(0.802)

Note: Each column reports results of either a least squares or a 2SLS regression of the change in the log wage on the change in the (instrumented) Mexican population share, conditional on the change in the number of U.S. births and year fixed effects. Columns (1)-(4) pertain to least squares models while columns (5)-(8) report IV results. Estimates are presented for sixteen age-skill groups using an exhaustive combination of four skill grpups (< High School education, High School education, Some College, College + education) and four age groups (19-25, 26-40, 41-55 and 56+). All models are weighted by 1980 MSA population. As they are conservative relative to standard errors that are clustered at the MSA level, we report heteroskedasticity-robust standard errors in parentheses below the coefficient estimates. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 8. The Effect of Mexican Immigration on the Unemployment Rate of U.S. Natives Estimates by Age-Skill Groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS				2SLS			
	Age 19-25	Age 26-40	Age 41-55	Age 56+	Age 19-25	Age 26-40	Age 41-55	Age 56+
< High School	0.412*	0.241	-0.060	-0.089	0.463	-0.187	0.019	0.094
	(0.213)	(0.158)	(0.128)	(0.104)	(0.475)	(0.350)	(0.259)	(0.310)
High School	0.307*	0.139	-0.040	-0.011	0.233	0.150	0.182	-0.040
	(0.173)	(0.094)	(0.073)	(0.079)	(0.288)	(0.215)	(0.137)	(0.203)
Some College	-0.056	0.054	-0.073	0.027	0.017	0.146	0.016	0.129
	(0.128)	(0.067)	(0.062)	(0.097)	(0.276)	(0.137)	(0.116)	(0.219)
College +	0.041	-0.071	-0.053*	0.087	-0.033	-0.028	-0.105*	-0.113
	(0.102)	(0.056)	(0.206)	(0.030)	(0.297)	(0.065)	(0.057)	(0.118)

Note: Each column reports results of either a least squares or a 2SLS regression of the change in the unemployment rate on the change in the (instrumented) Mexican population share, conditional on the change in the number of U.S. births and year fixed effects. Columns (1)-(4) pertain to least squares models while columns (5)-(8) report IV results. Estimates are presented for sixteen age-skill groups using an exhaustive combination of four skill grpups (< High School education, High School education, Some College, College + education) and four age groups (19-25, 26-40, 41-55 and 56+). All models are weighted by 1980 MSA population. As they are conservative relative to standard errors that are clustered at the MSA level, we report heteroskedasticity-robust standard errors in parentheses below the coefficient estimates. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 9. The Effect of Mexican Immigration on the Employment-to-Population Ratio of U.S. Natives
Estimates by Age-Skill Groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS				2SLS			
	Age 19-25	Age 26-40	Age 41-55	Age 56+	Age 19-25	Age 26-40	Age 41-55	Age 56+
< High School	-0.520**	-0.244	-0.046	0.092	-0.143	0.149	0.154	0.027
	(0.206)	(0.155)	(0.196)	(0.216)	(0.473)	(0.397)	(0.481)	(0.444)
High School	-0.328*	-0.247**	0.127	0.301	-0.246	-0.328	-0.015	0.580
	(0.195)	(0.104)	(0.110)	(0.190)	(0.369)	(0.243)	(0.235)	(0.554)
Some College	0.081	-0.134*	0.159*	-0.056	-0.141	-0.297*	0.057	0.126
	(0.151)	(0.073)	(0.086)	(0.263)	(0.345)	(0.169)	(0.237)	(0.760)
College +	-0.100	0.144*	0.105	0.278	0.308	0.123	0.315**	1.139**
	(0.153)	(0.073)	(0.068)	(0.182)	(0.369)	(0.093)	(0.153)	(0.556)

Note: Each column reports results of either a least squares or a 2SLS regression of the change in the employment-to-population ratio on the change in the (instrumented) Mexican population share, conditional on the change in the number of U.S. births and year fixed effects. Columns (1)-(4) pertain to least squares models while columns (5)-(8) report IV results. Estimates are presented for sixteen age-skill groups using an exhaustive combination of four skill grpups (< High School education, High School education, Some College, College + education) and four age groups (19-25, 26-40, 41-55 and 56+). All models are weighted by 1980 MSA population. As they are conservative relative to standard errors that are clustered at the MSA level, we report heteroskedasticity-robust standard errors in parentheses below the coefficient estimates. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1

Table 10. 2SLS Estimates of The Effect of Mexican Immigration on the Wages and Employment of U.S. Natives
Comparison Between the Births Instrument and the Network Instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Log Wage			Unemployment %			Employment %		
Overall	0.020	-0.585	-0.622	0.065	0.414	0.259	0.000	-0.446	-0.210
	(0.691)	(1.484)	(1.154)	(0.136)	(0.341)	(0.210)	(0.197)	(0.414)	(0.289)
< High School	-0.190	-0.775	-0.546	0.042	0.383	-0.087	0.111	-0.090	0.841
	(0.737)	(1.480)	(1.174)	(0.260)	(0.365)	(0.394)	(0.367)	(0.371)	(0.730)
High School	-0.110	-0.826	-0.646	0.182	0.490	0.324	-0.119	-0.654	-0.155
	(0.841)	(1.730)	(1.381)	(0.174)	(0.350)	(0.256)	(0.241)	(0.535)	(0.398)
Some College	0.160	-0.417	-0.019	0.095	0.469	0.356	-0.124	-0.560	-0.501
	(0.686)	(1.365)	(0.978)	(0.127)	(0.407)	(0.258)	(0.207)	(0.462)	(0.383)
College +	0.344	-0.455	-0.309	-0.060	0.126	0.000	0.297***	-0.175	0.053
	(0.484)	(1.298)	(0.936)	(0.049)	(0.211)	(0.139)	(0.115)	(0.341)	(0.208)
Instrument	Births	Network	Network	Births	Network	Network	Births	Network	Network
control for U.S. births	yes	no	yes	yes	no	yes	yes	no	yes

Note: Each column reports results of a 2SLS regression of the change in a given outcome (the log wage, the unemployment rate or the employment-to-population ratio) on the change in the instrumented Mexican population share, conditional on year fixed effects and, in some cases, the change in the number of U.S. births. Columns (1)-(3) report coefficient estimates for the log wage, columns (4)-(6) report estimates for the unemployment rate and columns (7)-(9) report estimates for the employment-to-population ratio. For each dependent variable, the first column pertains to IV regressions using the births instrument, the second column pertains to IV regressions using the Mexico-specific network instrument while the final column pertains to IV regressions using the Mexico-specific network instrument, controlling for U.S. births. All models are weighted by 1980 MSA population. As they are conservative relative to standard errors that are clustered at the MSA level, we report heteroskedasticity-robust standard errors in parentheses below the coefficient estimates. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1