

Mexican Migrants to the United States: an Alternative Methodology

José Martínez*

Fecha de recepción: 07/III/2012

Fecha de aceptación: 03/IX/2012

Abstract

Mexicans are the largest immigrant group in the United States. There is a lack of consensus about whether migrants from Mexico are positively or negatively selected. Data from the Mexican census suggest migrants are negatively selected while data from the U.S. census suggest intermediate selection. Both data sources undercount migrants, with Mexican sources systematically undercounting more educated migrants and U.S. sources undercounting less educated migrants. Net migration techniques are used to estimate migration flows during the 1990s and obtain estimates which present a more accurate characterization of Mexican immigrants. Three main conclusions are reached. First, the net flow of Mexicans to the United States during the 1990s was about 10 percent less than the U.S. census data suggest. Second, migrants are younger and less female than suggested by the U.S. census, but older and more female than suggested by the Mexican census. Third, U.S. census data significantly overstate the educational attainment of migrants. However, the disagreement in the literature on migrant selection results less from who is counted in Mexican and U.S. data and more on the responses given to the Mexican and U.S. census questions on schooling.

JEL Classification: F22, O15, J15.

Keywords: Net Migration, Mexico, Selectivity.

Resumen

Los mexicanos son el grupo de inmigrantes más grande en los Estados Unidos. No existe consenso sobre si los inmigrantes mexicanos están autoseleccionados positiva o negativamente. Datos del censo de población de México sugieren que están seleccionados negativamente, mientras que datos

* Department of Economics, University of North Texas. Address: 1155 Union Circle Denton, TX 76203, United States. Email: Jose.Martinez@unt.edu

de los Estados Unidos sugieren que están seleccionados de manera intermedia; sin embargo, ambas fuentes de datos subestiman el número de inmigrantes. Fuentes de datos de México sistemáticamente subestiman a migrantes con mayores niveles de escolaridad; mientras que fuentes de datos de los Estados Unidos, subestiman a los de menor escolaridad. Usando técnicas de migración neta para estimar los flujos migratorios durante la década de 1990, se obtienen estimadores que presentan una identificación más precisa de los migrantes mexicanos. Se llega a tres conclusiones principales. Primera, el flujo neto de mexicanos hacia los Estados Unidos durante 1990 fue cerca de diez por ciento menor al aludido por los datos de los Estados Unidos. Segunda, los migrantes representan una mayor cantidad de jóvenes y una menor cantidad de mujeres, que lo sugerido por los datos de los Estados Unidos; pero son más los adultos mayores y también más mujeres, que los sugeridos por el censo de México. Tercera, los datos de Estados Unidos sobreestiman significativamente la escolaridad de los inmigrantes. Sin embargo, acerca de la falta de consenso en cuanto a la literatura sobre la selectividad de los migrantes, los resultados muestran que es menor la escolaridad de quien es contado en México y en los Estados Unidos; pero es más de lo que se obtiene de las respuestas a las preguntas del censo sobre escolaridad.

Clasificación JEL: F22, O15, J15.

Palabras Clave: Migración neta, México, Selectividad.

Introduction

The 2000 U.S. population census counted 33 million individuals born outside the United States (U.S. Census Bureau, 2004) in 2000. The Mexican born is estimated to be 9.3 million, and represent the largest number from any single country. Those immigrants born in Mexico differ in educational attainment and age distribution from the native born population is uncontroversial. But there is less agreement on how migrants from Mexico compare with the Mexican population remaining in country.

The educational content of migration flows has been of longstanding interest to development economists concerned about “brain drain.” Docquier and Marfouk (2006) unify census data from the Organisation for Economic Co-operation and Development (OECD) countries and find that internationally, migrants are divided roughly into three equal groups having 8 or fewer years of schooling, 9 to 12 years of schooling and more than 12 years of schooling. They note that their data underrepresent flows of illegal migrants, who are not generally well measured by receiving country census agencies. Since illegal migrants are both large in number and likely to have lower education

levels, this has an effect of undetermined magnitude on the estimates of the educational content of migrant flows. The exercise we undertake in this paper provides a measure of the magnitude of the bias in these estimates for one large bilateral migration flow.

The educational attainment of Mexican migrants to the U.S. can be measured using data from either the Mexican or U.S. census. There is a clear pattern in the literature: analysis using data from Mexico finds that migrants have less schooling than those remaining in Mexico, while analysis using data from the United States finds more positive selection. Because of the nature of undercounted sectors of the migrant population, there is reason to believe that data from both the Mexican and U.S. census are biased, but in opposite directions. We show that estimates based on the method of net migration, while not without error, do not suffer from the obvious biases of the estimates relying on direct data taken from either of the two censuses.

The two data sources produce a different picture with respect to the age and gender of migrants as well. Among the Mexicans born arriving in the United States between 1995 and 2000, U.S. census data indicate that 60% are male and 40% are female. Among those who left Mexico for the United States without returning during the same years, the Mexican census reports that 75% are male and only 25% are female.

Given the issues with the current estimates of the size and characteristics of migration flows, we pursue in this paper an alternative approach of estimating the net outflow and characteristics of migrants from Mexico during the 1990s. We use data from the 1990 and 2000 Mexican and U.S. population censuses to calculate net migration from Mexico. Net migration compares the size of an age cohort in an earlier census with the size of an appropriately older cohort in a later census, adjusting for mortality. For example, we compare the number of 8-12 year olds in 1990 with the number of 18-22 year olds in 2000. We use data on the number of deaths to account for the reduction in cohort size due to mortality. The difference between the adjusted cohort sizes represents the estimate of net migration during the decade. By comparing age, education and gender cohorts, we provide a profile of migrants. Since about 98 percent of those who migrate from Mexico come to the United States, we obtain a very good estimate of the flows from Mexico to the United States.

We face several challenges in pursuing the net migration approach. For example, the percentage of the population with an unreported age is significantly higher in the 2000 Mexican census than the 1990 Mexican census, and the tendency for ages to be reported as numbers ending in 0 or 5 also changes between 1990 and 2000. The proportion of the population in

the U.S. census which is categorized as foreign born, country not specified, is much higher in the 1990 census than in the 2000 census (Cresce, Ramirez and Spencer, 2001). These changes across time result less from changes in the responses of households and more from decisions by the census bureau with respect to allocation, assigning values to non-responses. The percentage of the population in the United States uncounted by the census between 1990 and 2000 is also widely seen as having fallen.¹ We describe how we address these and related issues in more detail later in the paper. Importantly, while the solutions we pursue might add noise to the estimates, we believe that the migration estimates coming from the net migration analysis are much less subject to the biases of those relying directly on data from either the Mexican or U.S. census.

Our first main finding is that the net flow of Mexicans to the United States was about 10 percent less than the U.S. census data suggest. While the U.S. census data indicate that 4.39 million Mexicans between the ages of 3 and 72 (in 1990) came to the United States during the 1990s, we estimate the flow to be 4.04 million.² Second, we find that the age distribution of migrants lies between that obtained using U.S. data and that obtained using Mexican data, but is closer to that obtained from the U.S. data. Third, we find that the education attainment also lies between that obtained from the two censuses, but is much closer to the estimates obtained from the Mexican census. This suggests that the U.S. census significantly overstates the educational attainment of Mexican migrants. The differences in the estimates of educational attainment are too large to be explained by the differences in who is counted by the two censuses, implying the Mexicans respond differently to the questions on educational attainment in the U.S and Mexican censuses. Since there is much less reason to think that the educational attainment of the children born to Mexican migrants is similarly biased, this implies that the gain in schooling from the first to second generation is much larger than the literature currently estimates.

¹ The U.S. Census Bureau's Executive Steering Committee for Accuracy and Coverage Evaluation (ACE) estimated that the undercount was reduced from 1.61% of the population in 1990 to 1.18% in 2000. The ACE also concluded that the improvement was especially marked among Hispanics and other minority groups. The estimated undercount rate for Hispanics fell from 4.99 in 1990 to 2.85 in 2000. Farley (2001) discusses some of the methods the Census Bureau used to increase the accuracy of the 2000 count, including an advertising campaign, an increased number of enumerators, and partnerships with community organizations in difficult-to-count populations. Hogan (2001) offers corroborating evidence of the more complete count by comparing census counts with data on school and Medicare enrollments.

² The 4.39 million figure is obtained by comparing the size of the 3-72 year old cohort in 1990 and then in 2000.

1. Literature Review

Borjas (1996) argues that high returns to education in Mexico lead to negative selection of migrants, as the better educated have stronger incentives to remain in Mexico. Ibarra and Lubotsky (2005) and Fernández-Huertas Moraga (2011) provide evidence in support of negative selection using the 2000 Mexican population census and Mexican National Employment Survey, respectively. Chiquiar and Hanson (2005), on the other hand, argue that migration costs can reverse the prediction of negative selection. If all migrants pay comparable fixed costs to migrate, then lower skilled workers will have to work longer to offset those costs, making migration to the United States a less attractive option for the less skilled. Chiquiar and Hanson (2005) provide evidence based on the 2000 U.S. population census suggesting that migrants come disproportionately from the upper half of the education and wage distribution in Mexico. Cuecuecha (2003) reaches a similar conclusion using data on the Mexican born in the Current Population Survey. McKenzie and Rapoport (2006) show that the probability of migration is increasing with education in communities with low migration networks, but decreasing with education in communities with high migration networks. This might explain in part why some authors find positive selection of migrants based on education while others have found evidence of negative selection. On the other hand, Aguayo-Tellez and Martínez-Navarro (2013) find evidence that single adult males with low levels of education tend to migrate to the U.S., even when controlling for self-selectivity of migrants. Finally, migrant's remittances have been shown to promote human capital accumulation for non-migrants in the source country, Contreras (2012).

The most direct measure of migration flows from the 2000 Mexican population census data are derived from a question asking whether anyone in the household has migrated outside of Mexico within the previous five years. Those responding affirmatively are asked the age and gender of the migrant(s), the country to which each individual migrated, and whether the individual has returned to Mexico. Ibarra and Lubotsky (2005) use these data to estimate the education levels of migrants based on the education of non-migrants remaining in the household. They find that migrants have education levels which are, on average, lower than the population remaining in Mexico. Fernández-Huertas Moraga exploits the fact that the Mexican National Employment Survey is a rolling panel, and uses data on both the characteristics and earnings of individuals surveyed one quarter and reported as having migrated the next.

Chiquiar and Hanson (2005) point out that the Mexican census question fails to count households that migrate out of Mexico in their entirety. They argue

that these households are more likely to be urban and more highly educated than households sending single members to the United States. Hence, the Mexican census and employment survey data bias downward the educational attainment of migrants to the United States. Instead, Chiquiar and Hanson (2005) use data from the U.S. census on the characteristics of the Mexican-born residing in the United States, comparing them with the population resident in Mexico. The U.S. census data suggest that migrants have far higher education levels than the Mexican census, leading to a conclusion that migrants have schooling levels which are higher than the population remaining in Mexico.³

Supporting Hanson and Chiquiar's argument, the U.S. census data suggest that 2.7 million Mexicans arrived in the United States between 1995 and 2000. The Mexican census measures only 1.3 million Mexicans leaving the country for the United States during the same five-year period.⁴ This suggests that as many as half of the migrants from Mexico are not counted by the question asked in the Mexican census.

But there are reasons to believe that the data from the U.S. census present a distorted picture as well. The best estimates suggest that 10 to 15 percent of Mexican-born migrants were not counted in the 2000 census. The U.S. census is widely understood to undercount the Mexican-born population in a systematic way (Constanzo Davis, Irazi, Goodkind and Ramirez, 2001). Young, single, low-wage workers are less likely to live at a fixed address and more likely to be undercounted. The uncounted are more likely to have low levels of education relative to other migrants and relative to natives (Borjas and Katz, 2005; Bean, Corona, Tuirán, Woodrow-Lafield, and van Hook, 2001). Additionally, there are issues with the language translations for schooling levels used in the Spanish version of the U.S. census. Most importantly, the U.S. census refers to high school as *secundaria*. In Mexico, *secundaria* refers to the junior high school level. High school is referred to as *preparatoria* or *bachillerato* in Mexico.

³ Using a sample of individuals 16 years of age and older, Ibarra and Lubotsky (2005) conclude that between 16 and 26 percent of those migrating to the U.S. come as a household. Their estimates are based on comparing populations counted in Mexico and the United States. Our own estimate, discussed in more detail later in the paper, is that households migrating entirely between the two censuses account for at least 36 percent of migrants from Mexico.

⁴ This number does not include those who returned and resided in Mexico at the time of the 2000 census. Including those who returned, the Mexican census counts 1.6 million emigrants to the United States.

2. The Net Migration Methodology

The net migration methodology is widely used to overcome the lack of data on migration flows in most censuses of the 20th century. The method relies on the difference between changes in total population and changes in the natural population between two periods.

The estimated net migration flow M can be represented as:

$$M = (P_{t,a+t} - P_{o,a}) - \sum_{i=1}^t (B_i - D_i) \quad (1)$$

The first component on the right-hand side is the difference between the population of age a at time 0 and the population of age $a+t$ at time t . The second component on the right-hand side is the difference between births and deaths in this cohort over the t years. A positive value for M would imply that, on net, more people came into the country than went out and vice versa.

For reasons that we discuss later in the paper, we exclude the cohort which was 0-2 years of age in 1990 and 0-12 years of age in 2000. Thus, births are not a factor in our analysis. However, an accurate mortality rate is critical to the analysis. We use Mexican data on deaths from the vital statistics from the country, which is the most reliable method for accounting for mortality,⁵ (Siegel and Hamilton, 1952).

One of the main advantages of the net migration method is that it does not require data from migration-specific questions. There are several potential drawbacks which must be addressed. One relatively minor drawback is that the method generates only the net flow and does not allow for separate estimates of outflows and inflows. More serious concerns derive from changes in census methodologies over time. For example, as Bogue, Hinze, and White (1982) point out, changes in the percentage of the population enumerated by the census or changes (*de facto* or *de jure*) in the method of allocating non-responses from the enumerated population will affect the estimate of net migration if not properly accounted for. In addition, incorrect or incomplete measures of mortality may also affect the estimates. In our case, mortality figures are available only by age and gender; so we must

⁵ Hill and Wong (2005) also estimate migration flows using a version of the net migration method. However, their analysis is limited to overall flows without differentiating by gender and education.

assume they are independent of education levels. If mortality rates are negatively correlated with education levels, then we will likely understate the educational attainment of migrants. However, for individuals aged 10-55, mortality rates are quite low, so the lack of education-specific mortality rates is not likely to have a large impact on the estimates.

We must also be concerned with “educational drift.” Other researchers have noted a tendency for individuals report higher educational levels as they age. This is particularly common among individuals with low levels of schooling. Thomas and Muvandi (1994), with reference to Botswana and Zimbabwe, attribute the phenomenon to regression to the societal mean during a period in which educational attainment is increasing rapidly. That is, the older generation reports higher schooling attainment as the attainment of the younger generation actually increases. Thomas and Muvandi (1994) find little evidence of drift in the United States, where educational attainment has risen more slowly in recent generations. Educational drift will affect any estimate of the educational attainment of migrants obtained using net migration. We address this issue by deriving an estimate for the level of educational drift, as described in Table A1 of the appendix, and by exploiting differences between Mexican states with high and low migration rates.

Mexican immigrants may be more likely to overstate their educational attainment after arrival from a tendency to revert to the community mean. Black, Sanders and Taylor (2003) compare responses to educational attainment from a subsample of the 1990 U.S. population census with responses to a resurvey with much more detailed questions about educational attainment. They find that just over 7 percent of those who report having a bachelor’s degree in the census are found not to have a college degree (BA) when asked the more detailed questions. Among Hispanics, the error rates are much higher: 17 percent of those who report having a BA in the census do not have a college degree. Black, Sanders and Taylor (2003) find that the larger error among Hispanics (and Asians) is almost entirely attributable to low levels of English language ability and the large number of immigrants among the Hispanic population, concluding: “Although this is suggestive that language skills may play an important role in the measurement error, it may simply reflect that immigrants often have a lack of familiarity with the U.S. higher education system.” Language and differences in the educational system suggests that the educational attainment levels reported in the Mexican census are likely to be more accurate than those reported by Mexican immigrants in the U.S. census.

An additional issue is that the net migration method gives us an estimate of all international migrants from Mexico. We are interested in migration from

Mexico to the United States. There is no way to isolate migration to the United States. However, this concern is reduced because the United States is the destination of about 98 percent of the migrants reported in the Mexican census. The Mexican census data do indicate that among households reporting migrants going to Europe, the education level of the household head is much higher (12.7 years) than the level of household heads with migrants to the United States (4.9 years). If European-bound migrants themselves have higher schooling levels than U.S-bound migrants, this will result in a small upward bias of the estimated educational attainment of migrants from Mexico to the United States.

3. Results

3.1. Migration Flows and Gender Estimates

Both in numerical counts and in characteristics of migrants, the estimates of migration obtained from the U.S. and Mexican censuses are very different. Table 1 shows the age and gender distribution of migrants leaving Mexico between 1995 and 2000 -estimated using the Mexican population census (Column 2, including migrants who have returned)- and arriving in the United States -estimated using the U.S. population census (Column 3)- between 1995 and 2000. Since the Mexican census fails to count families migrating as a unit, we should expect the Mexican data to understate the number of children and females.⁶ Since migrants residing in the United States as a family are more likely to be counted in the U.S. census (because they are more likely to reside in a fixed residence, for example). We should expect these same groups to be over-represented in the U.S. census data. Consistent with these expectations, Table 1 shows that while only five percent of migrants in the Mexican census are 14 years of age or younger, 22.6 percent of those enumerated in the U.S. census are 14 or younger. Similarly, only a quarter of migrants in the Mexican census are females, while females comprise almost 42 percent of the Mexican born population in the U.S. census. On the other hand, the Mexican data indicate much larger participation in migration by the 15-24 year olds: 54 percent of migrants are in this age range in the Mexican data compared with only 37 percent in the U.S. data. This is the age range which is most likely to live outside of traditional housing -such as migrant camps- or in multi-family households, and hence be undercounted by the U.S. census.

⁶ Both females and younger children are more likely to migrate as a part of an entire household or for reasons of marriage (United Nations, 2006), more likely to live in apartments or houses once arriving in the United States and/or more likely to be counted by the U.S. census.

We now turn to estimating the net migration from Mexico to the United States. We derive three estimates. The first utilizes the Mexican census data, and the second the U.S. census data. The two censuses produce different estimates even of the total flow of migrants because the percentage of the population uncounted in the census falls between 1990 and 2000 in each country. Our third net migration estimate accounts for improvements in census coverage between 1990 and 2000 in both countries. We return to this estimate below.

Table 1
Mexican Migrants to the U.S.

	(1)	(2)	(3)
	Mexican Population	Mexican Census Migrated 1995-2000	U.S. Census Arrived 1995-2000
Age / Gender	%	%	%
0-4	11.4	1.5	7.1
5-9	11.7	1.3	9
10-14	11.2	2.2	6.5
15-19	10.4	27.2	13.8
20-24	9.4	27.1	23.3
25-29	8.4	15.1	15.9
30-34	7.4	8.6	8.7
35-39	6.6	5.6	5.5
40-44	5.4	3.5	3.4
45-49	4.2	2.2	2.2
50+	13.5	3.2	4.1
Male	48.7	74.8	58.1
Female	51.3	25.2	41.9

Notes: Data in columns (1) and (2) from 2000 Mexican Population Census. Data in column (3) from the 5% Public Use Micro Sample of the 2000 U.S. Population Census. Source: calculations by author.

There are several data issues we must address in making the net migration estimates. We list the issues here, and describe how we address them:

- 1) In the Mexican data, there is a tendency for people to report ages ending in 0 or 5. For example, in 1990 around 300,000 individuals are reported as being 59 years old, 640,000 as 60 years old and 170,000 as 61 years old. The clumping suggests that people tend to

ound their age to 60 years even if they are 59 or 61 years old.⁷ Clumping increases with age, that is, it is more severe around 60 than it is around 30, and clumping is more apparent in 1990 than it is in 2000. We address this by grouping ages into five-year cohorts centered around the ages ending in 5 and 0. That is, we compare the number of males aged 3-7, 8-12, etc. in 1990 with the number of males aged 13-17, 18-22, etc. in 2000.

- 2) The percentage of the Mexican sample with missing age is higher in 2000 than in 1990. We address this issue by increasing the percentage of the 1990 sample with missing ages so that it matches the percentage with missing age in 2000. These adjustments are discussed in greater detail in the appendix.
- 3) In the U.S. census, the percentage of the foreign-born population with the country of birth not specified is much higher in 1990 than in 2000. In 2000, the census bureau assigned a country of birth to individuals reporting themselves as foreign born, but not reporting a country of birth. The allocation was based on the place of birth of members of nearby households. We allocate the place of birth in the 1990 census using the percentage of the foreign born population (by age and gender) in an individual's Public Use Micro Area (PUMA) which was born in Mexico.
- 4) The net migration calculations indicate large net in migrations of children aged 0-2 in 1990 into both Mexico and the United States. Indeed, summing up the population born in Mexico from the population censuses of both countries, we find about 1 million more children aged 10-12 in 2000 than there were children aged 0-2 in 1990. This appears to result from households not reporting very young children in the population census.⁸ We therefore exclude this cohort from the discussion, and focus on those 3-72 years of age in 1990.

The first two columns of Table 2 show the raw out/in migration flows from the Mexican and U.S. censuses, respectively. The Mexican data show much smaller migration flows, reflecting the fact that households leaving in their entirety are not counted in the Mexican census. The lower percentage of

⁷ A similar issue exists with U.S. census data but to a lesser extent, perhaps because ages are smoothed by allocations made by the census bureau. See figures A1 and A2 in the appendix for more detail.

⁸ A similar pattern is evident in the census data from Brazil, Chile, Argentina, and even from the U.S., to a lesser extent. Looking at the population by each single year of age suggests that the phenomenon is limited to those under age 2.

female migrants in the Mexican data reflects the same phenomenon. The third and fourth columns of Table 2 show our estimates of migrant flows developed using the net migration methodology applied to both the Mexican (Column 3) and U.S. (Column 4) censuses, adjusted for the issues just discussed.

Table 2
Mexico-U.S. Migration Flows 1990-2000, by Age and Gender

Age	Gross Outflow/Inflow		Net Migration		
	Raw Data		Unadjusted		Adjusted*
	Mexico	U.S.	Mexico	U.S.	
0-2	-25,806	229,014	864,607	205,761	NA
3-7	-216,091	412,118	-219,772	408,162	-328,583
8-12	-615,474	820,351	-1,074,803	837,573	-933,176
13-17	-698,257	982,373	-1,326,294	1,054,981	-1,163,421
18-22	-363,469	638,111	-697,656	761,116	-723,452
23-27	-221,390	356,441	-346,915	514,632	-436,191
28-32	-127,866	215,399	244,545	308,327	-70,552
33-37	-91,502	128,896	-265,799	184,597	-212,560
38-42	-63,529	86,347	-125,645	119,597	-114,862
43-47	-38,915	51,266	19,387	67,803	-24,019
48-52	-26,806	37,778	-90,495	49,567	-58,253
53-57	-16,306	23,800	105,840	29,894	35,178
58-62	-5,092	15,886	-30,655	29,398	-19,453
63-67	-3,655	10,687	-100	18,123	1,637
68-72	-1,326	5,899	2,849	11,176	11,455
0-72	-2,515,484	4,014,366	-2,940,905	4,600,678	NA
3-72	-2,489,678	3,785,352	-3,805,511	4,394,917	-4,036,251
% Female	26.8	43.4	40.2	43.6	39.4

Notes: Calculations using data from 1990 and 2000 Population Census for Mexico and the U.S.* Data are adjusted for the improvement in the coverage of the census in Mexico and the U.S. The improvement is allocated between the two countries in a manner described in the text. NA: not available.

Source: calculations by author.

For the Mexican data, we use the 10 percent samples from the 1990 and 2000 Mexican censuses. The 1990 sample is unweighted and representative at the state level. The 2000 census is weighted and representative at the *municipio* (county) level. Smaller *municipios* are over-sampled, and sampling weights are provided. We use the sampling weights for all of our calculations. Following equation (1), and given that we are going to follow age cohorts over time, equation (1) can be rewritten as:

$$M_{aj} = (P_{(a+10)j}^{2000} - P_{aj}^{1990}) + \sum_{t=1990}^{2000} D_{ajt} \quad (2)$$

here a and $a+10$ refer to age cohort and j refers to gender. Annual data on deaths were obtained from the Mexican National Statistical Institute (INEGI). These data contain deaths by age cohort, gender and *municipio* of usual residence of the deceased. For example, we have the number of deaths for males age 0 to 4 for each year from 1990 to 2000 for each *municipio*.

For the U.S. data, we use the 5% Public Use Micro Survey data, aggregated using the weights provided in the census. Mortality data comes from the National Center for Health Statistics, for the entire U.S. population. Mortality rates by cohort are then applied to the cohort size to obtain a number of deaths. The results are not sensitive to the choice of mortality rates. We obtain very similar results if we use mortality rates for the Hispanic population or even the Mexican mortality rates. The Mexican data are shown as negative numbers, reflecting net outflows, while the U.S. data are shown as positive numbers, reflecting net inflows.

Comparing the net migration estimates based on Mexican data (column 3) and those based on U.S. data (column 4), the Mexican data measure larger flows than the U.S. data among those 8-17 years of age in 1990. This is consistent with those in this age range being more likely to be uncounted in the U.S. census. The U.S. data, on the other hand, suggest larger flows than the Mexican data among those 23-32 years of age in 1990. The distribution of flows across age ranges is much smoother in the U.S. data than in the Mexican data. We expect this reflects more frequent data allocation in the U.S. census. The U.S. data also suggest that a larger portion of the migrants are female, consistent with the greater likelihood of undercounting single males in the count.

Note that the estimates using the different censuses produce not only a different distribution of migrants, but a different total flow of migrants. The U.S. data suggest a net flow of almost 4.4 million Mexican migrants aged 3-72 (in 1990), while the Mexican data show only 3.8 million migrants in the same age range. The difference in total flow results from an improvement in the census coverage over the decade. In other words, the census bureaus did a better job of counting the Mexican-born population in 2000 than in 1990. An improved count in Mexico will tend to reduce estimates of net migration, since we will find people in the 2000 data who were not covered in 1990. An improved count in the United States will have the opposite effect.

Combining data from the Mexican census with data on the Mexican-born population resident in the United States, we find an “extra” 614,547 individuals born in Mexico in the 2000 census. This represents the combined reduction in the undercount in the two countries, which must be allocated between the two countries. The (U.S.) Census Bureau’s Executive Steering Committee for Accuracy and Coverage Evaluation (ACE) estimated that the total undercount was reduced from 1.61% of the population in 1990 to 1.18% in 2000. The ACE also concluded that the improvement was especially marked among Hispanics and other minority groups.⁹ We know of no independent estimate of the improvement of coverage in Mexico,¹⁰ so instead we use the estimates in the improvement in coverage in the United States. Based on the ACE analysis, we estimate that 25% of the Mexican-born population was uncouned in 1990 and 12.5% was uncouned in 2000. Together, these suggest that 256,797 of the improvement in coverage should be assigned to the United States. That implies that 357,749 of the improved coverage should be assigned to Mexico.¹¹ We assign these improvements proportionately by age/gender cohort. After making these adjustments, the two censuses produce a unified estimate of net migration flows totaling 4.04 million during the 1990s. That estimate is shown in Column 5 of Table 2.

Table 3 shows the percentage distribution of the net migration flow by age and gender using the age cohorts which straddle the 0’s and 5’s. The first column shows the age distribution of the population residing in Mexico in 2000. Column 2 shows the distribution of those reported as having migrated in the Mexican census data, Column 3 those reporting in the US census as having arrived in the U.S. between 1990 and 2000, and Column 4 the unified net migration flows, repeated from Column 4 of Table 2. The data give an indication of the magnitude of the bias in counting migrants in both the Mexican and the U.S. census. These data also provide us with an estimate the number of migrants arriving from households leaving Mexico in their entirety between the 1990 and 2000 census. Both the 2000 census and the

⁹ The undercount rate for Hispanics went down from 4.99 in 1990 to 2.85 in 2000. Farley (2001) discusses some of the methods the Census Bureau used to increase the accuracy of the 2000 count, including an advertising campaign, an increased number of enumerators, and partnerships with community organizations in difficult-to-count populations.

¹⁰ The Mexican Population Bureau (CONAPO) estimates that the undercount fell from 3.64 percent in 1990 to 2.54 percent in 2000. However, the CONAPO estimates incorporate data on the number of Mexicans in the United States from the Current Population Survey, and hence are not based on improvements in collection methods in Mexico.

¹¹ The improvement in the coverage in Mexico apparently occurred without a major increase in resources. The only major change between 1990 and 2000 was an increase in the collection time from one week to two weeks. This allowed for a reduction in the number of interviewers from 500,000 in 1990 to 260,000 in 2000, suggesting only a minor increase in person/days.

1995 inter-census *Conteo* ask households if any member has migrated outside of Mexico in the previous five years. Summing those individuals aged 3-72 in 1990 who have left Mexico and not returned within five years, we find 2.56 million migrants. This suggests that 1.47 million migrants in 2000 come from households not counted.¹² In other words, just over a third of migrants (36 percent) come from households leaving in their entirety between censuses. Combined with estimates of the undercounts in the U.S. census, this suggests that each census fails to count a significant percentage of the migrant population.

Table 3
Percentage Distribution of Migrants by Age and Gender

Age / Gender	Population-2000 Mexican Census	Migrants		
		Migrants counted in Mexican Census	U.S. Census arrivals 1990-2000	Adjusted net migration flows*
3-7	13.2	3.7	9.1	8.1
8-12	13.0	17.2	18.0	23.1
13-17	12.5	26.5	21.6	28.8
18-22	10.4	17.9	14.0	17.9
23-27	8.7	11.2	7.8	10.8
28-32	7.2	6.5	4.7	1.7
33-37	5.9	4.5	2.8	5.3
38-42	5.1	3.1	1.9	2.8
43-47	3.7	1.8	1.1	0.6
48-52	3.3	1.3	0.8	1.4
53-57	2.6	0.9	0.5	-0.9
58-62	2.2	0.5	0.3	0.5
63-67	1.7	0.2	0.2	0.0
68-72	1.2	0.2	0.1	-0.3
Male	49.0	74.8	55.9	60.6
Female	51.0	25.2	44.1	39.4

Notes: Calculations using data from 1990 and 2000 Population Census for Mexico and the U.S. * Table 2, Column 4.

Source: calculations by author.

¹² There are two sources of error in this estimate. First, the 4.04 million net outflow certainly includes some returning migrants. Hence, the gross outflow of migrants not returning to Mexico before 2000 was certainly more than 4.04 million, and hence more than 1.47 million must have left with their household. On the other hand, some migrants leaving as individuals within the previous five years may not be reported by the household members remaining in Mexico.

How are the undercounts in the two censuses reflected in the characteristics of migrants? Consider first the cohort aged 3-7 in 1990. These individuals are more likely to have migrated as children in a household leaving in its entirety, and hence be undercounted by the Mexican census. They are also more likely to live with their parents in the United States, and live at a fixed address where they are more likely to be counted in the US census. Consistent with these expectations, we find that the net migration method increases the share of this age group in the total migration flow compared to the Mexican census (by 4 percentage points), but decreases the share compared with the U.S. census (by 1 percentage point). With respect to the gender of migrants, we also find results consistent with the expected biases in the census data from the two countries. Only 25 percent of the migrants counted in the Mexican census are females, compared with 44 percent in the U.S. census. Since females are more likely to migrate as a household and live in standard housing in the U.S., we expect the former understates the share of females in the flow, while the latter overstates the share. Using net migration, we find that 39 percent of the flow is female, between the two estimates.

In sum, there are large differences between U.S. and Mexican data on migrants in terms of age, gender, educational attainment and number of migrants. These differences are consistent with the prior expectations of the nature of bias in both the U.S. and the Mexican census data. The net migration calculations illustrate the problem with identifying the nature of immigrant selection relying on direct estimates of flows from either the Mexican census or the U.S. census. Both are likely to be biased in important ways and in opposite directions. We next turn to the issue of estimating the educational attainment of migrants using net migration.

3.2. The Educational Attainment of Migrants

Net migration estimates of the age and gender of migrants are intermediate between the estimates based on Mexican and U.S. census data. This is consistent with correcting for expected biases in estimates coming directly from either census. But the debate on the selectivity of Mexican migrants has focused less on age and gender and more on the educational attainment of migrants. What does net migration tell us about the educational attainment of migrants?

The data from the U.S. census and CPS indicate that the education level of Mexican migrants is much higher than is suggested by the Mexican census

data.¹³ While not ruling out the possibility that migrants obtain additional schooling after arriving in the United States -a point we discuss in more detail below- there are reasons to believe the U.S. data overstate the educational attainment of migrants. First, lower-educated migrants are more likely to have lower-paying jobs, and to live in non-traditional housing. Hence, those with lower education are more likely to be undercounted. Second, there are issues with the translation of the education categories in the Spanish version of the U.S. census. In particular, *secundaria* is used to indicate high school, while *secundaria* refers to the junior high school level in Mexico. Third, migrants may experience higher levels of “educational drift” after crossing the border, influenced by the higher education levels of those in their new community.

We use the net migration data to estimate the educational attainment of the net flows of migrants from Mexico to the United States. Ideally, we would obtain the education distribution for a given cohort in 1990 and then compare it to the corresponding cohort in 2000, adjusted for mortality. However, four issues complicate the ability to do this.¹⁴ First, non-migrants in Mexico may increase their actual educational attainment between the 1990 and 2000 censuses. Ignoring this increase in education would yield estimates that understate the educational attainment of migrants. We adjust for additional schooling using the percentage of individuals of a given age and educational attainment who report they are attending school. Table A1 in the appendix describes these adjustments in more detail. Moreover, we minimize this issue by carefully constructing the age and education cohorts. For example, most individuals who will complete lower secondary school have done so by age 18, so we construct our first cohort to include people age 18 to 22 in 1990 and adjust for the percentage of population still attending school. In this age cohort, we split the sample into three groups: those with 0 to 4 years of schooling, those with 5 to 8 years of schooling, and those with 9 or more years of schooling.

A second issue is that, as with age, there are differences in the 1990 and 2000 Mexican censuses in the percentage of individuals with a missing education level. For example, for those aged 18 to 22, a smaller percentage has missing responses on schooling attainment in 1990 than in 2000. Again, failing to account for this would lead to incorrect estimates of the

¹³ The Mexican census does not ask for the education level of migrants. Ibararan and Lubotski (2005) estimate the educational attainment of the migrants reported in the Mexican census using the characteristics of the remaining members of the household, along with the age and gender of the migrants. Their estimates suggest the educational attainment of migrants is substantially lower than that suggested by the U.S. data.

¹⁴ There is an extra issue related to people overstating their educational attainment as they get older. We look at it in more detail when we present the results on education.

educational attainment of migrants. We address this issue by implementing a method similar to the one we used for addressing differences in the frequency of missing age: We estimate education for enough of those missing education in 2000 so that the percentage of the samples with missing education data is comparable in the two samples. Details of this are given in the appendix.

Third, we must allocate the improvement in the count between 1990 and 2000 to different levels of schooling. We do this in a strictly proportional manner. If anything, we expect that the improvement in coverage might have been greater among those with lower levels of schooling. If that is the case, then the proportional distribution will tend to overstate the educational attainment of migrants. Given the numbers involved, the effect of this is small. Finally, we use the same mortality figures for all schooling levels. There is some evidence that education is negatively correlated with the rate of preventable deaths (Kenkel, Lillard and Mathios, 2006), which would imply that our estimates understate the educational attainment of migrants. Given that we use only individuals aged 18-37 in our education estimates, the mortality rates are very low. Differences in mortality by education level are unlikely to have a material effect on our estimates.

Given sensitivity of estimated educational attainment to these adjustments, we also pursue an alternative method of estimating educational levels. We calculate the difference-in-difference between Mexican states with high and low migration rates. The difference-in-difference has the advantage of not requiring any of the adjustments discussed above, so long as these factors have similar impacts among those remaining behind in high- and low-migration states. Table 4 presents the education of distribution and the difference-in-difference estimates comparing the upper and lower quartile of states ranked by migration rate. In the eight states in the upper quartile, an average of 10.1 percent of households reported a migrant to the United States between 1995 and 2000; in the eight states in the lower quartile, only 1.3 percent of households reported a migrant.¹⁵ We make calculations for two age groups, those 18-22 and 23-27 years of age in 1990. For both groups, the difference-in-differences estimates suggest that those with 0-4 years of schooling are less likely to migrate and those with 5-8 years of schooling are more likely to migrate. There are modest differences in the two cohorts at the highest schooling levels. The difference-in-differences estimates indicate that those 18-22 in 1990 with 9 or more years of schooling

¹⁵ Migration rates are based on percentage of households with at least one migrant to the United States from 1995-2000. Low migration states are Tabasco, Quintana Roo, Chiapas, Campeche and Yucatan with an average migration rate of 1 percent. High migration states are Aguascalientes, Durango, Guanajuato, Michoacán and Zacatecas with an average migration rate of 12 percent.

are more likely to have migrated, while in the 23-27 age group, those with 9-11 years of schooling are more likely to have migrated but those with 12 or more years of schooling are much less likely to have migrated.¹⁶ The simple difference-in-difference analysis suggests that migrants are more likely to have intermediate levels of education. But the distribution of schooling attainment is also shifted far to the left compared to the distribution reported in the U.S. census.

Table 4
High-migration vs. Low-migration States*

		Age 18-22 in 1990			
		Educational distribution (%)			
		0-4	5-8	9+	
Low Migration States	1990	24.70%	28.40%	47.00%	
	2000	22.40%	25.90%	51.80%	
Difference		-2.30%	-2.50%	4.80%	
High Migration States	1990	18.70%	35.40%	45.90%	
	2000	19.20%	31.60%	49.20%	
Difference		0.50%	-3.80%	3.30%	
Difference-in-Differences		2.80%	-1.30%	-1.50%	

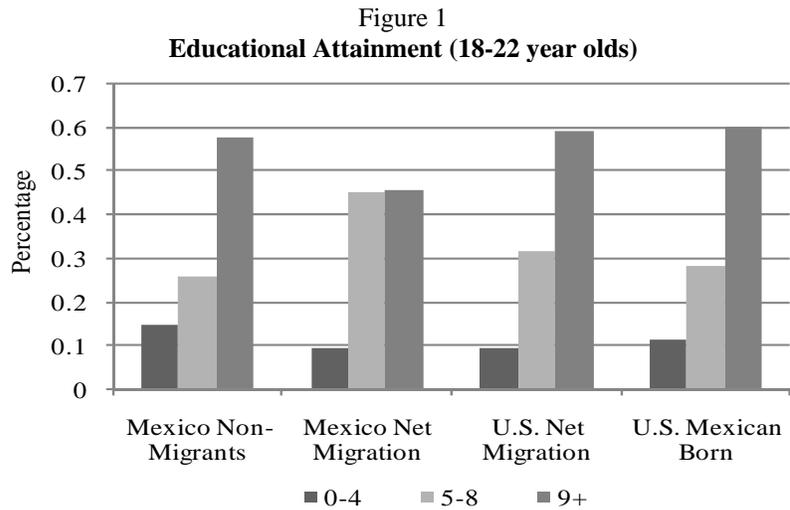
		Age 23-27 in 1990			
		Educational distribution (%)			
		0-4	5-8	9-11	12+
Low Migration States	1990	33.00%	25.70%	19.60%	1.80%
	2000	30.70%	24.30%	19.90%	5.10%
Difference		-2.30%	-1.40%	0.30%	3.30%
High Migration States	1990	28.70%	32.10%	20.10%	9.10%
	2000	27.50%	28.80%	19.50%	4.10%
Difference		-1.20%	-3.30%	-0.60%	5.00%
Difference-in-Differences		1.10%	-1.90%	-0.90%	1.70%

Notes: *It includes the 8 high-migration and the 8 low-migration states. A negative difference-in-difference implies more out-migration of this group. Calculated using data from the 1990 and 2000 Mexican Population Census. Individuals classified by state of birth rather than state of residence so that migration within Mexico does not affect the results.

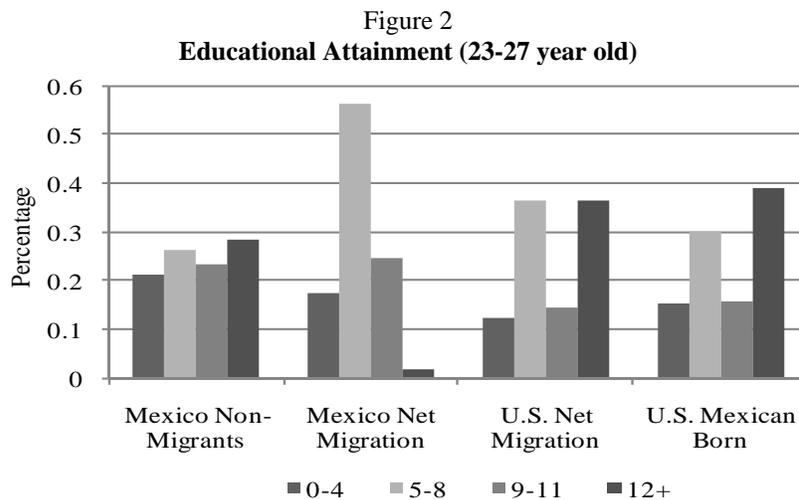
Source: calculations by author.

¹⁶ We find generally similar results if we use all 32 federal entities and compare the upper quartile with the lower three quartiles. The differences are smaller in all categories, and for the 23-27 year olds, we find more migration among those with 0-8 years of schooling and less among those with 9 or more years of schooling.

Figures 1 and 2 show the results of the full net migration analysis, with all of the adjustments discussed above and in the appendix. For each age group, we estimate four separate distributions. The figures show the distribution of educational attainment for those residing in Mexico in 2000, for those identified in the U.S. census as having arrived from Mexico between 1990 and 2000, and for the flows estimated by net migration using both the Mexican and U.S. census data.



Source: calculations by author.



Source: calculations by author.

Compared with the population remaining in Mexico, the net migration data based on the Mexican data tell a story which is generally consistent with the difference-in-difference (dif-in-dif) results reported above. Migrants are less likely to have very low or high levels of schooling, and are much more likely to have 5 to 8 years of schooling. There is some difference between the dif-in-dif and net migration estimates of those with 9-11 years of schooling, with the net migration estimates suggesting that this group is equally represented in the migrant and non-migrant population, and the dif-in-dif estimates suggesting that those with 9-11 years of schooling are more likely to migrate. This may reflect either differences in the data, or the fact that the dif-in-dif estimates are based on 16 of Mexico 32 federal entities, while net migration covers the entire country. Because migrants are less likely to have 0-4 years of schooling, these data provide some support for the view that there is intermediate selection of migrants. However, relative to the population remaining behind, the net migration estimates suggest that migrants have lower schooling levels on average.

An equally interesting feature of the figures is the difference in the implied educational attainment of migrants using the Mexican and U.S. census data. Both estimates we report are based on the same number of migrants going from Mexico to the United States, that is, we use the unified net migration flows. Notice that the net migration estimates based on Mexican data are very different from those based on U.S. data. The U.S. data imply that a much larger share of the migrants are high school graduates (or more), and a much smaller share have 5-8 years of schooling.

Since the unified net migration flows match migrants by age and gender, the differences in the educational distribution suggest that individuals report different levels of schooling in the U.S. than they do in Mexico. Why might this be the case? One possibility is that migrants to the United States might obtain schooling after arrival. However, available data suggest that few Mexican-born arriving in the United States as adults obtain a high school education after arrival. Since we focus on those arriving after age 18 (or age 23), the main channel for increasing educational attainment is by obtaining a Graduate Education Degree (GED). There are two sources of data suggesting that obtaining a GED is rare among these groups. First, the 1992 wave of the Legalized Population Survey includes a category for GED.¹⁷ Less than one percent of those in the 18-27 age range, and only two percent who report

¹⁷ The Legalized Population Survey interviewed a sample of immigrants who had applied for legal status following the passage of the Immigration Reform and Control Act (IRCA) in 1987. Individuals were interviewed twice, once in 1989 just after they applied for legal status, and once in 1992, after legal status had been obtained. The full LPS sample size is 6,193 in 1989 and 4,012 in 1992. The sample includes 608 Mexican immigrants aged 18-27 in 1989.

having a high school education or more, obtained a GED. Among those ages 28-37 saying they have completed high school, less than five percent did so through obtaining a GED. Second, the October 2000 CPS also has questions about the GED. Among those aged 28-37 (in 2000), only nine percent of the Mexican-born with exactly a high school education, and less than two percent of the entire age group, say they have obtained a GED.¹⁸ Since the Mexican data show the largest difference among those with 5-8 years of schooling, we find it extremely unlikely that 23-27 year old migrants would obtain additional schooling by attending junior high school.

A second possibility is that those not counted by the census may have a different education profile than those counted. If the uncounted in the U.S. census have lower schooling levels, then the U.S. census is likely to overstate the educational attainment of migrants. But the undercount rates are not high enough to explain much of the gap. Therefore, it appears likely that a substantial part of the difference is caused by individuals responding differently to the U.S. and Mexican census questions. The Spanish version of the U.S. census refers to high school as *secundaria* and college as *bachillerato*. In Mexico, *secundaria* refers to junior high school and *bachillerato* to high school. For these reasons, we believe the Mexican net migration data present the more accurate picture of the educational attainment of migrants from Mexico.¹⁹

In sum, net migration estimates of the educational attainment of migrants using U.S. data differ greatly from estimates based on Mexican census data. Net migration estimates for Mexico suggest migrants from Mexico tend to have low or medium levels of education, while U.S. data suggests they tend to have medium to high levels of education. The data suggest that the differences in educational attainment of migrants are driven less by who is counted or uncounted, and more by the responses to the educational attainment questions in the Mexican and U.S. censuses.

¹⁸ The CPS data are not comparable to the LPS data because the CPS question on the GED is asked only of those with a high school degree, not those with more than high school.

¹⁹ In results not shown, we account for the potential negative correlation between education and mortality by adjusting the rate of lower education cohorts to be twice the rate for higher education cohorts. Such adjustments increase the level of positive selection of migrants, but not considerably due to the size of the adjustment relative to the size of the cohort. We also look at the issue of people overstating their educational attainment as they get older, the educational drift. This phenomenon would imply that results tend to understate the level of positive selection, so the adjustment to the results is meant to represent an upper bound for the positive selection. The results for the 23 to 27 year olds in 1990 are included in the appendix, including a brief explanation on how the adjustment was done.

Conclusions

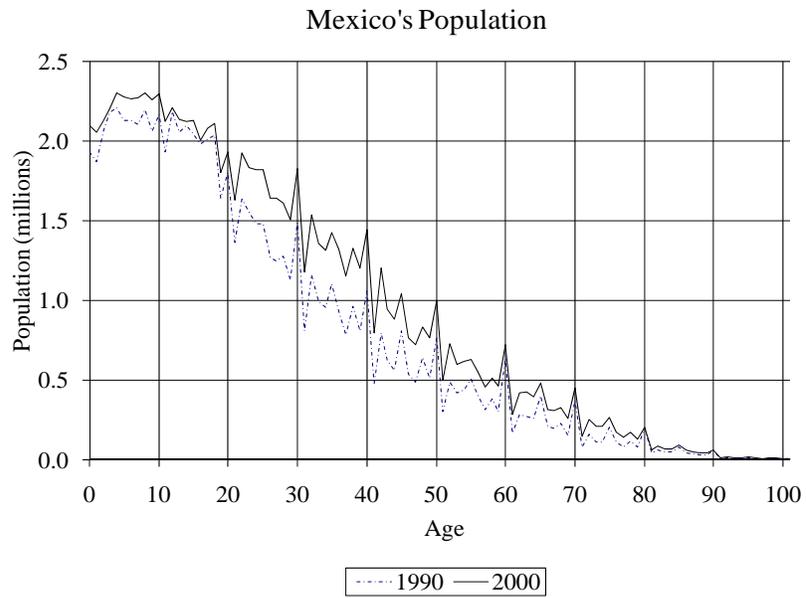
We use the net migration methodology to obtain estimates of the age, gender, educational attainment and number of migrants from Mexico. Our first finding is that migration from Mexico to the U.S. during the 1990s was about 300,000 less than the estimate obtained from the U.S. census data. We also find that migrants from Mexico are somewhat younger than suggested by Mexico's census data on migration but older than estimates based on U.S. census data. Likewise, net migration estimates suggest females have a higher migration participation rate than it is suggested by Mexican data on migrants, but a lower rate than it is suggested by U.S. data.

Using the Mexican census data, we also find that migrants have lower schooling levels on average than the population remaining in Mexico, though they are less likely than non-migrants to have the very low levels of schooling. Using the U.S. data, the implied educational attainment of migrants is much higher, and above the average for those remaining in Mexico. There are reasons to believe that the responses to the questions in the U.S. census are overstated. There is less reason to think the responses to the Mexican census questions are understated to the same degree. Hence, we view the net migration estimates based on Mexican data as likely to be more accurate.

These results have two important implications. First, regarding the selectivity of migrants from Mexico to the U.S., the results suggest that the differences between the negative selectivity found when using Mexican data (e.g., Ibararan and Lubotsky, 2005) and the positive selectivity found with U.S. data (e.g., Chiquiar and Hanson, 2005) arise less from differences which migrants are counted by the U.S. and Mexican censuses and more from differences in responses to the education questions in the U.S. and Mexican census. The results imply that migrants from Mexico have lower education levels than the U.S. data indicate. Second, since at least some of the factors contributing to the overestimate of education are limited to the first generation respondents, our results suggest that the gain in educational attainment from first to second generation Mexican immigrants is larger than is found using U.S. census or CPS data. The rate of educational assimilation during the first generation, therefore, is more rapid than commonly portrayed.

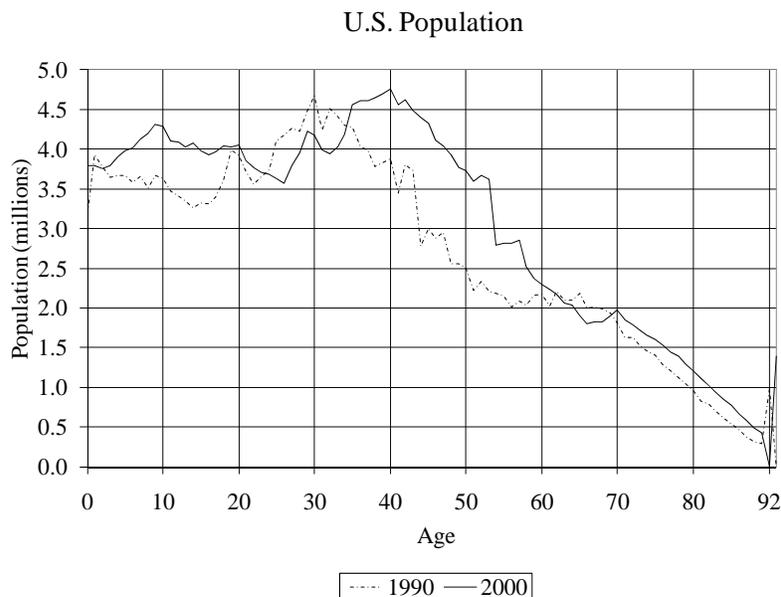
Appendix

Figure A1
Age clumping around the 0's and 5's



Source: calculations by author using the 1990 and 2000 Mexico's Population Census 10% sample data.

Figure A2
Age clumping around the 0's and 5's



Source: calculations by author using the 1990 and 2000 U.S. Population Census 5% sample data.

Predicted Education

We follow Ibarra and Lubotski (2005) method to obtain predicted levels of education. They use data from the 2000 Mexican census to run regressions on years of education for non migrants living in non migrant households on indicator variables for age, town size and number of children age 0 to 8. Regressions include also indicators for the number of adult men and women, and they are performed separately by gender and municipality. Predicted values for years of education are obtained for the whole population, including migrants. Given than obtained levels are not exact numbers, we construct the cohorts using decimal points. For example, from 0 to 4.499 are classified as 0-4. 4.5 to 8.499 as 4-9, and so forth.

Age cohorts' mortality adjustment

Mortality data from Mexico for each year between 1990 and 2000 is available by five-year age group, gender and *municipio* of usual residence of the deceased. The data show, for example, the number of deaths in each *municipio* for males age 5-9 in 1991. We use age cohorts which straddle the

0's and 5's, and so use a weighted average of both the annual data and the five-year cohorts available in the mortality data. For example, our mortality number for the 3-7 year old cohort in 1990 uses 2/5ths the number of deaths of the age 1-4 cohort and 3/5 the number of deaths of the age 5-9 cohort in 1990. By 1991, this cohort is aged 4-8, so we use 1/5th of the age 1-4 deaths and 4/5ths of the 5-9 deaths. A similar procedure is used up to 1999.

Adjustment for missing age

In the 1990, approximately 0.103 percent of individuals have no reported age from the 10% sample of the Mexican census. In 2000, 0.30 percent of the sample has a missing age. INEGI claims this is just the natural result of the data collection process. The only reported changes between 1990 and 2000 were the increase in the collection time from one to two weeks and the reduction in the number of interviewers by almost 50 percent. As we discuss in the text, failing to adjust the data so that the same percentage of observations are missing ages in 1990 as in 2000 would result in a mis-estimation of net migration.

We adjusted the data by first regressing an indicator for missing age on personal characteristics, including family size, number of children, relation to household head, marital status using the 2000. We then used the probability a person with given other characteristics had a missing age to remove the ages of the 0.197 percent of the 1990 population with the highest probability of having a missing age. With this adjustment, both samples had 0.30 percent of observation with no age reported.

Education Adjustments

As with age, the percentage of the 1990 and 2000 samples with missing years of schooling also differs. Failure to account for this would lead to errors in the estimates of the educational attainment of migrants. In the case of education, the percentage with missing data is higher in 1990 than in 2000. For example, among those 18-27 years of age in 1990, only 1.7 percent has missing education in 1990 and 1.16 percent in 2000.

We adjusted for missing education in a manner similar to our adjustment for missing age. However, we do the adjustment for the 18-22 year old and 23-27 year old cohorts separately. That is, we perform the regressions on a variable indicating missing education and use the predicted values for each cohort separately. Again, regressions were made using personal characteristics, including family size, number of children, relation to household head, marital status, and age. With the adjustment, each cohort had the same percentage of individuals with no education reported.

A second issue with the education data relates to people actually acquiring more the schooling between 1990 and 2000. This applies to both the U.S. and Mexican data. We estimate the educational attainment of migrants by comparing the educational distribution of those residing in Mexico in 1990 and 2000 and adjusting for mortality. For example, if after adjusting for mortality there were 1,000 fewer people with less than 5 years of education in the 18 to 22 age cohort in 2000 than in 1990, then we could say that the number of migrants with less than primary schooling was 1,000. However, if individuals actually obtain more schooling between 1990 and 2000, then part of the smaller number of individuals would be explained not by migration but by the additional schooling. Attributing all of the drop in the size of the low-schooling group to migration would result in a downward bias in the estimated education of migrants.

Our first strategy for addressing this concern is to construct age cohorts and education levels to try to minimize this problem. We consider only those aged 18-27 in 1990. Among those aged 18-22, we use 9 or more years of schooling as the top education level. Our expectation is that those reaching age 18 are unlikely to continue to attend high school. There are some individuals age 18 to 22 with less than 9 years of schooling who report that they are still attending school in 1990. By 2000, these individuals may have attained the next level of schooling. Given that we find much lower levels of schooling for migrants using the Mexican census data, the conservative approach is to assume that every such individual completes 9 or more years of schooling by 2000. We therefore reduce the size of the 5-8 education cohort and increase the size of the 9+ cohort for every such individual for the 1990 data. If instead some of the individuals attending school in 1990 drop out of school before completing 9 years, our estimates will overstate the educational attainment of migrants. Given the age cohorts we use, the effect of this adjustment is modest, as only 2.5 percent of those 18-22 have 5-8 years of schooling and say they are still attending school.

Educational Drift

To adjust for the possibility of people overstating their educational attainment as they get older, and therefore understating the level of positive selection, we use data from people living in low-migration states in Mexico: Campeche, Chiapas, Quintana Roo and Tabasco. We take the 1995 Mexican inter census count and use only the data for people who claim are no longer attending school at each age cohort. Then, we use data from the 2000 census for the same states for non-migrants who claim they were not attending school at that time and were living in the same state in 1995.

The education distribution for the 22 to 27 year olds in 1990 for 1995 and 2000 is shown below. An increase in proportions was observed for the less than four and nine to eleven years of education cohort, while a reduction was observed in the remaining two groups.

Table A1
Education distribution

Education	1995	2000	Difference
<=4	25.0	26.4	1.4
5-8	30.3	28.6	-1.7
9-11	18.6	19.9	1.3
12+	25.9	24.9	-1.0

Source: calculations by author

The results suggest more people claimed they had less than five years of schooling in 2000 than in 1995. This is somewhat puzzling because we would expect people to overstate their education, and not the other way. However, this could be in part due to the data itself, which shows a way higher percentage of people without education specified in 1995 than in 2000.

References

- Aguayo-Tellez, E. and Martinez-Navarro, J. (2013). "Internal and International Migration in Mexico: 1995-2000". *Applied Economics*, 45(13), 1647-1661.
- Bean, F. D., Corona, R., Tuirán, R., Woodrow-Lafield, K. A. and van Hook, J. (2001). "Circular, Invisible, and Ambiguous Migrants: Components of Difference in Estimates of the Number of Unauthorized Mexican Migrants in the United States". *Demography*, 38(3), 411-422.
- Black, D., Sanders, S. and Taylor, L. (2003). "Measurement of Higher Education in the Census and Current Population Survey". *Journal of the American Statistical Association*, 98(463), 545-554.
- Bogue, D. J., Hinze, K., White, M. (1982). *Techniques of Estimating Net Migration*. Chicago: Community and Family Study Center, University of Chicago.
- Borjas, G. J. (1996). "The Earnings of Mexican Immigrants in the United States". *Journal of Development Economics*, 51(1), 69-98.
- Borjas, G. J. and Katz, L. F. (2005). "The Evolution of the Mexican-Born Workforce in the United States". *NBER Working Papers 11281*.

- Chiquiar, D. and Hanson, G. (2005). "International Migration, Self-Selection, and The Distribution of Wages: Evidence from Mexico and the United States". *Journal of Political Economy*, 113(2), 239-281.
- Contreras, S. (2012). "The Influence of Migration on Human Capital Development". *International Economic Journal*, iFirst Article, 1-20.
- Costanzo, J., Davis, C., Irazi, C., Goodkind, D. and Ramirez, R. (2001). "Evaluating Components of International Migration: The Residual Foreign Born". *Population Division Working Paper No. 61*.
- Cresce, A., Ramirez, R. and Spencer, G. (2001). "Evaluating Components of International Migration: Quality of Foreign-Born and Hispanic Population Data". *Population Division Working Paper No.65*.
- Cuecuecha, A. (2003). "The Educational Characteristics of the Immigrants from Mexico to the U.S.". *Working paper ITAM-CIE*. Mexico City.
- Docquier, F. and Marfouk, A. (2006). International Migration by Educational attainment. In Ozden and Schiff (Eds.), *International Migration, Remittances and the Brain Drain*. New York: McMillan and Palgrave.
- Farley, R. (2001). "The Unexpectedly Large Census Count in 2000 and its Implications". *PSC Research Report No. 01-467*.
- Fernández-Huertas Moraga, J. (2011). "New Evidence on Emigrant Selection". *The Review of Economics and Statistics*, 93(1), 72-96.
- Hill, K. and Wong, R. (2005). "Mexico-US Migration: Views from Both Sides of the Border". *Population and Development Review*, 31(1), 1-18.
- Hogan, H. (2001). "Accuracy and Coverage Evaluation: Data and Analysis to Support the ESCAP Report". *DSSD Census 2000 Procedures and Operations Memorandum Series B-1*.
- Ibarraran, P. and Lubotsky, D. (2005). "Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census". *NBER Working Paper No. 11456*.
- Kenkel, D., Lillard, D. and Mathios, A. (2006). "The Roles of High School Completion and GED Receipt in Smoking and Obesity". *Journal of Labor Economics* 24(3), 635-660.
- McKenzie, D. and Rapoport, H. (2006). "Can Migration Reduce Education Attainments? Depressing Evidence from Mexico". *CREAM Discussion Paper Series 01/06*. London, UK: Centre for Research and Analysis of Migration.

30 *Ensayos Revista de Economía*

- Siegel, J. S. and Hamilton, H. C. (1952). "Some Considerations in the Use of the Residual Method of Estimating Net Migration". *Journal of the American Statistical Association*, 47(259), 475-500.
- Thomas, D. and Muvandi, I. (1994). "The Demographic Transition in Southern Africa: Another Look at the Evidence from Botswana and Zimbabwe". *Demography*, 31(2), 177-199.
- United Nations (2006). "2004 World Survey of the Role of Women in Development: Women and International Migration". New York: United Nations.
- United States Census Bureau (2004). "The Foreign-Born Population in the United States, 2003". United States: US Department of Commerce, Economics and Statistics Administration, US Census Bureau.