Preliminary and incomplete

Mexicans in America

María Genoni
Duke University

Luis Rubalcava
CAMBS

Graciela Teruel
Universidad Iberoamericana

Duncan Thomas
Duke University

September 2011

Abstract

Undocumented migrants are hard to enumerate. Data from the Mexican Family Life Survey (MxFLS) are used to measure the extent and selectivity of migrants not enumerated in surveys in the United States. MxFLS is an on-going longitudinal population-representative survey of Mexicans living in Mexico in 2002. About 90% of the movers from Mexico to the U.S were interviewed in the U.S. in the first follow-up in 2005. These respondents are compared with recent migrants from Mexico interviewed in the American Community Survey (ACS) and Current Population Survey (CPS). We estimate the ACS misses about 30% of recent migrants and both ACS and CPS undercount younger, single, male and less educated migrants. Relative to migrants in ACS and CPS, MxFLS migrants are more likely to be working, earn less and have lower returns to education. This is because the ACS and CPS miss the lowest earners among the least educated migrants.

Financial support from CONACYT, the National Institute of Aging, the National Institute of Child Health and Human Development, the Fogarty International Center and the Hewlett Foundation are gratefully acknowldeged.

1. Introduction

It is estimated that about 12 million Mexican-born people are living in the United States (U.S.) which amounts to around 10 percent of the entire Mexican population and it is thought that, in recent years, approximately one third of new migrants entering the U.S. were born in Mexico. Counting the number of Mexicans living in the U.S. and the number entering in any year is not straightforward because many enter and live in the U.S. without documentation and systematic records on those who return to Mexico are not maintained. An important literature has developed and implemented methods for estimating the number of Mexicans living in the U.S. as well as the flows of migrants between the countries, drawing on data collected in the U.S. and in Mexico. (Bean and Van Hook, 1998; Bean et al., 2001; Fernandez-Huertas Moraga, 2011; Ibarraran and Lubotsky, 2007; Passel abd D'Vera Cohn, 2009; Rendall et al., 2011).

The existing empirical evidence on the size of the Mexican-origin population in the U.S. is largely based on census and survey data that has been collected on both sides of the border. However, interpreting evidence based on each of these data sources poses challenges that complicate inferences about the size and composition of the Mexican-origin population in the U.S. (Hanson, 2006). On one hand, data collected in Mexico provide information only on migrants with roots in Mexico because the informant reporting on the migrant has to be a member of a household in Mexico in order to be included in the census or eligible for the survey conducted in Mexico. Moreover, informants are usually asked about migrants who have a connection with his or her household. As a result, more permanent migrants to the U.S. and households that move entirely to the U.S. are likely to be missed in studies based on these data.

On the other hand, U.S. data sources are likely to undercount the hardest to find, particularly more transitory and undocumented migrants. This is a serious concern since 80 percent of recent Mexican-origin migrants and 55 percent of the total Mexican population living in the U.S. is thought to be undocumented (Passel and D'Vera Cohn, 2008). Moreover, given the proximity of Mexico and the U.S., many migrants are thought to be circular, crossing the border frequently.

A related concern revolves around the characteristics of those who are under-counted in these studies. If those who are not counted are selected on characteristics such as age, education or successful assimilation, inferences about the impact of migration on the domestic and migrant populations will be complicated as well estimates of the extent to which Mexican migrants in the U.S. assimilate.

The limited evidence that exists on these issues suggests they are important concerns. Ong and Houston (2002) find that neighborhoods in Los Angeles County with the highest undercount rates in the 2000 census tend to be poor and predominantly minority, and have a relatively large number of children. In addition, Ibarraran and Lubotsky (2007) compare Mexican-origin migrants counted in the U.S. with reports of migrants to the U.S. in the Mexican census and find that migrants counted in the U.S. are older and better skilled than those enumerated in the Mexican census. They conclude that the differences are likely to be explained by the undercount of young undocumented migrants and the over reporting of education in the U.S. census. Orrenius and Zavodny (2005) point out that if the likelihood of being included in the U.S. census depends on skills, results based on the U.S. census will be biased. Similarly, Fernandez-Huertas Moraga (2011) finds evidence that undercounting is selective in terms of education, by comparing the U.S. census and the American Community Survey with the Mexican Employment Survey (ENET).

This paper examines new longitudinal data designed to provide direct evidence on the extent and selectivity of Mexican-origin migrants living in the U.S. We examine a sample of recent Mexican migrants who have been interviewed in the U.S. as part of the Mexican Family Life Survey (MxFLS) a study that is representative of the population of Mexico at the time of the baseline survey which was conducted in 2002. We use data from the first follow-up conducted in 2005. We will update these estimates with evidence drawn from the second follow-up which is currently in the field.

An innovative feature of MxFLS that is key for this study is our decision to track and interview all respondents who move to the U.S. after the baseline interview. In principle, the MxFLS sample of Mexicans in the U.S. are representative of the population of all Mexicans who have moved from Mexico to the U.S. since 2002 and were living in the U.S. at the time of the 2005 follow-up. We successfully interviewed over 90 percent of those respondents.

We estimate that at least 2.250 million Mexicans moved to the U.S. between 2002 and 2005 and were living in the U.S. in 2005. This is 25% higher than estimates of the Mexican-born population who moved to the U.S. between 2002 and 2005 based on the American Community Survey (ACS) and the Current Population Survey (CPS), two nationally-representative surveys that are widely used for migration studies and form the basis of the vast majority of estimate of the Mexican-born population in the U.S.

The selectivity of respondents in these surveys, relative to those interviewed in the U.S. as part of MxFLS, is examined. Differences in the characteristics of migrants in the U.S. surveys

relative to those in MxFLS provides evidence on the characteristics of those who are likely to be under-counted in U.S. sample surveys. The comparisons also provide indications of the likely biases that arise in studies of the success of migrants in the U.S. and studies that explore the impact of migrants on the domestic population.

The demographic and socio-economic characteristics of MxFLS respondents interviewed in the U.S. are not the same as those interviewed in the ACS and CPS. Specifically, MxFLS respondents are significantly more likely to be female, they are younger, significantly less educated, more likely to be single and more likely to be renters. These differences are consistent with the hypothesis that the U.S. surveys undercount more transitory (and probably undocumented) migrants. The differences are both statistically significant and substantial in magnitude. For example, about 40 percent of recent migrants in CPS and ACS have at least 12 years of schooling. This group accounts for only 20 percent of MxFLS respondents. Migrants interviewed in MxFLS are also more likely to be working and earn substantially less than comparable respondents in CPS and ACS.

Comparisons are also drawn with the Mexican Migration Project (MMP) which has been an important source of evidence on migration between Mexico and the U.S. (Massey et al, 1990; Donato, Durand and Massey, 1992) Because the MMP is a sample of respondents in the U.S. who have been identified by family members living in Mexico, it tends to oversample migrants who have roots in Mexico. Relative to comparisons with the CPS and ACS, the differences between MxFLS and MMP are smaller although MMP migrants are significantly more likely to be male, married, and less educated than MxFLS respondents.

MxFLS is described in more detail in the next section. Our estimates of the number of recent migrants from Mexico are presented and followed by comparisons with ACS, CPS and MMP.

2. Data

MxFLS is an ongoing longitudinal survey that collects extensive information on individuals, households, families and communities. The baseline survey, MxFLS-1, was conducted in 2002 and collected detailed information about 35,677 individuals living in 8,440 households in 150 communities spread across 16 states in Mexico. The sample was selected by INEGI, the Mexican statistical agency, and is designed to be representative of the entire Mexican population living in Mexico at the time of the baseline survey (Rubalcava and Teruel, 2004). The

first follow-up, MxFLS-2, was conducted in 2005-2006 and interviews were completed with at least one individual from over 90 percent of the baseline households. The second follow-up, MxFLS-3, is currently in the field.

A novel feature of MxFLS, which is key for this research, is that we decided to not only follow respondents who moved within Mexico but to also follow respondents who moved to the U.S. and interview them in the U.S. [1] Those who subsequently return to Mexico are tracked back to locations in Mexico and interviewed there. Following movers is hard and few large-scale surveys have followed respondents across international borders. We conducted a series of pilot tests to assess the feasibility of tracking movers to the U.S. and to evaluate alternative strategies to tracking and recontact that would maximize the probability of success in this endeavor.

We started fieldwork in the first follow-up by returning to each baseline location and asking the whereabouts of each respondent who was living there in 2002. When a member of a baseline household was found, the respondent was asked about the location of each member of the baseline household. If one of those respondents had moved, information about the new location was recorded including physical addresses, telephone numbers, electronic mail addresses and contact information for people in the new location who may know the whereabouts of the respondent. When no members of the origin household were located, we sought information about the respondents from other MxFLS respondents in the same community and also from other informants in the community including local leaders, teachers, health providers, post office workers and employers.

Obtaining information about migrants in the U.S. is not straightforward and out pilot work highlighted two critical factors that were included in the design of MxFLS. In some cases, the respondent provided an address and phone number in the U.S. In other cases, the informant could only provide contact information about someone else living in the U.S. who would know how to contact the target respondent. In many cases, the informant did not have contact information as the migrant moved around and telephoned the informant regularly. In other cases, the informant was reluctant to provide contact information without first checking with the target respondent. In these cases, we asked the informant to obtain contact information and permission to give us that information the next time the informant spoke to the migrant. Obtaining this information required multiple re-visits to the informant's home. More generally, whenever we obtained contact information, we asked the informant to let the migrant know we had been to

[1] About 95 percent of Mexican international movers go to the U.S.

their home and that we would like to conduct the same interview with the migrant. We followed up with a phone call to the migrant to make contact, tell the migrant we have been in touch with the family in Mexico, bring news of the family and request permission for the follow-up interview. The pilots suggested – and our experience with tracking confirmed – that it was imperative that we build trust with the informant and this often required multiple visits to the home to gain their confidence. Of course, the fact that we had interviewed the respondents in the baseline and that we were re-interviewing them to find out how their lives had changed contributed to gaining their confidence and trust.

A second lesson from the pilots was that migrants in the U.S. were initally very reluctant to speak with MxFLS enumerators. Telling the migrants about their families that we had just interviewed allayed these fears and, once we had gained the confidence of one migrant, we were able to interview other migrants in the same social network – typically people living together, working together or migrants from the same village in Mexico. Key for our success is the fact that the MxFLS enumerators are Mexican and speak with an identifiable Mexican accent. To be sure, our task was facilitated by the fact that the respondent had been interviewed in the baseline in Mexico and we reminded the respondent of that interview.

This effort paid off. As shown in panel A of Table 1, in MxFLS-2, we re-interviewed 89.2 percent of the 35,000 baseline respondents. Of those respondents, 854 had moved to the U.S. and 90.6 percent of them were interviewed there. As shown in panel B of the table of the migrants who were age 16 or older in 2005, 91 percent were re-interviewed in the U.S. Respondents in this age group form our analytical sample because children are likely to have moved with one or both parents and it is the characteristics of those parents that is likely to drive selection.

MxFLS follows individuals who split off from their baseline households as well as all baseline members if everyone in the household moved to the U.S. This implies that the analyses will not be restricted to migrants with roots in Mexico but, as additional waves of MxFLS are added, it will include more settled and permanent migrants. In other words, since the baseline sample is representative of the population living in Mexico in 2002, respondents interviewed in the U.S. in the first follow-up are representative of all Mexicans who moved to the U.S. since 2002 and are still living in the U.S. in 2005.

The MxFLS respondents in the U.S. are compared with respondents in the 2005 American Community Survey (ACS) and the 2005 March Supplement of the Current Population

Survey (CPS). Both ACS and CPS have been major sources of information for studies of migrants in the U.S.

ACS is a monthly survey that collects information on demographic, social, and economic characteristics of the U.S. population and housing. Fully implemented in 2005, the ACS is the largest household survey in the United States, with a sample size of about 3 million housing unit addresses throughout the country. It has been estimated that the ACS misses about 10 percent of the undocumented migrants (Hoefer et al., 2006). A key advantage of ACS is that the sample sizes are very large and estimates are precise. However, ACS does not collect detailed information on economic characteristics. We therefore also use CPS which interviews a smaller sample but collects detailed economic data.

CPS is a monthly survey of about 55,000 households and it is the primary source of information on the labor force characteristics of the U.S. population; the sample is expanded to about 80,000 households for the March supplement. For the CPS March Supplement, the estimated undercount of unauthorized migrants is about 12 percent (Passel and Vera-Cohn, 2008 and 2009).

Both the 2005 ACS and 2005 March CPS samples are restricted to include all Mexican born respondents who report having arrived in the U.S. during the 3 years prior to the survey.[2] The samples will be directly comparable with the MxFLS sample of recent migrants assuming that recall error in the date the migrant reports moving to the U.S. in the ACS and CPS is small. This seems a reasonable assumption given the short recall period and the likely saliency of the move that brought the respondent to the U.S.

3. Results

We begin with a comparison of estimates of the number of Mexican-origin migrants who moved to the U.S. between 2002 and 2005, drawing on the 2005 ACS and MxFLS. Since the CPS sample size is relatively small, we do not draw comparisons with that source. The demographic and socio-economic characteristics of respondents in MxFLS are contrasted with the characteristics of respondents in ACS and CPS in effort to provide insights into the reasons

[2] For ACS, we use country of birth to identify the Mexican born and year of entry to identify recent migrants. The year of entry variable in ACS asks all respondents born outside the U.S. for the year in which they came to live in the U.S. For respondents who have entered the U.S multiple times, the interviewers were instructed to request the most recent year of entry. For CPS, we use place of birth to identify the Mexican-born and recent movers are identify by the response to a question about when the respondent came to the U.S. to stay.

for discrepancies in the estimated number of new migrants. We next explore estimates of earnings functions based on these data in order to assess whether inferences about assimilation and selectivity of migrants differ depending on the data source.

The Mexican Migration Project (MMP) is a major source of evidence on migrants in the U.S. and in the final sub-section, we contrast respondents interviewed in the U.S. as part of that survey with migrants interviewed in the U.S. as part of MxFLS.

3.1 Number of recent migrants from Mexico to the U.S.

Table 2 reports estimates of the number of recent migrants from Mexico to the U.S. by age and gender, based on MxFLS (in panel A) and the ACS (in panel B). The estimates are weighted to take into account sampling probabilities and non-response (including inter-survey attrition in MxFLS).

According to MxFLS, 2.26 million Mexicans moved to the U.S. between 2002 and 2005. This estimate is 30% higher than the ACS estimate of 1.72 million new migrants. The gap between MxFLS and ACS is larger for males than females and is largest for young males (ages 10-24). For some demographic groups, the ACS estimates more migrants than MxFLS. It is possible that this reflects the impact of attrition in MxFLS that is selected on migration to the U.S. in which case the MxFLS estimates of the number of recent migrants would be a lower bound. One approach to address this concern might take the higher estimate of migrants for each demographic group. In that case, our estimate of the total number of recent migrants to the U.S. would be 2.34 million which is close to the MxFLS estimate. Either way, the evidence indicates that there is a substantial undercount of recent migrants in the ACS.

3.2 Comparison of demographic and socio-economic characteristics of migrants

To provide some insights into the people who are included in MxFLS but not in ACS, we turn next to a comparison of the demographic and socio-economic characteristics in these surveys and, also, in CPS. Univariate comparisons are followed by a comparison of conditional means in multivariate models. This is followed by a comparison of the distribution of earnings. The samples are restricted to respondents who are age 16 and older in 2005.

Comparisons across the surveys are reported in Table 3. In each panel of the table, means and standard errors are presented for MxFLS, ACS and CPS in columns 1 through 3. Panel A compares all respondents; panels B and C compare males and females, respectively. All

estimates are weighted so as to be population-representative and standard errors take into account clustering of the surveys.

As noted above, relative to ACS, Mexican migrants are more likely to be male: 66 percent of recent Mexican movers are male in MxFLS and 62 percent in ACS and CPS The average age of migrants is 28 or 29 years with males slightly younger than females. These differences across the data sources are relatively small. However, about half the respondents in the U.S. surveys are married whereas in MxFLS only 36 percent of the respodents are married.

There are important differences in the distribution of the education of respondents across the surveys. In CPS, 26 percent of recent Mexican migrants have completed primary school or less and 40 percent have completed high school or more. The ACS distrbution is slight to the left: 31 percent completed primary school or less and 38 percent completed high school or more. It is not clear what underlies these differences. The may reflect sampling error (which is larger in the CPS) or differences in the interpretation of the question about when the respondent moved to the U.S. However the difference pales in comparison with the difference between them and the MxFLS distribution which is substantially to the left of both distributions. Among MxFLS migrants, 37 percent completed primary school or less and only 20 percent report completing high school or more. The MxFLS sample is considerably less educated than the samples of migrants interviewed in the U.S. – a point made by Ibarraran and Lubotsky (2007). The education gap is larger for males than females.

While the socio-demographic characteristics discussed thus far are comparable across the surveys, measures of employment and earnings are not. In MxFLS, an individual is *employed* if the person responded that he/she is currently employed in the U.S. In CPS and ACS employed persons are those who worked at a paid job or business for at least one hour during the reference week, or worked at an unpaid family business for 15 or more hours during the reference week or who did not work last week, but held a job or owned a business from which they were temporarily absent during the reference week.

In MxFLS the earnings question asks approximately, how much do you earn per month at your main employment in the U.S.? About 6.5 percent of the respondents did not know the amount they earned at their main job. Interviewers were instructed to probe using unfolding brackets which elicited an estimate of earnings from over 40 percent of the respondents who could not provide an answer. As a result, only 3.7 percent of the MxFLS respondents did not provide information on monthly earnings.

In CPS, earnings are defined as the income from longest job held in the past year. This includes earnings from employers or net earnings from business/farm after expenses. The amounts reported are before taxes. The amounts are converted to monthly earnings based on the number of weeks worked in the past year. In ACS earnings are similarly defined but include income from all jobs in the past year and so we expect those estimates to be higher than CPS and MxFLS.

In addition to being less educated, migrants in MxFLS are more likely to be working than migrants in CPS and ACS. This may reflect differences in the nature of the question although the gaps are large enough to suggest that the differences reflect a higher rate of connection with the labor market among MxFLS respondents. Almost all the male recent migrants in MxFLS report that they were working at the time of the survey.

Recall that earnings in ACS is for all jobs but only for the main job in CPS. 8 percent of CPS migrants report having more than one job and the difference in average earnings in the two surveys is 6.7 percent. However, average earnings reported in MxFLS is 14 percent less than in CPS and the gap is bigger for females relative to males.

Relative to those interviewed in CPS and ACS, migrants in MxFLS are also more likely to be renters, which along with being less well educated, more likely to be working and earning less is consistent with the MxFLS respondents being more transitory migrants who are working in the U.S. and have not put down roots in the country.

Multivariate comparisons

We turn next to multivariate models to identify those characteristics of MxFLS migrants that are significantly different from the characteristics of recent migrants in ACS and CPS. Results are presented in Table 4 which reports estimates from linear probability models in which respondents in MxFLS and the comparison survey are stacked and we report the interaction between each covariate and an indicator that the respondent is in MxFLS. The coefficient estimate measures the difference between the MxFLS and comparison respondents. All models are weighted and standard errors take into account clustering of the surveys. Panel A of the table includes all respondents, panels B and C are restricted to males and females, respectively. In each panel, MxFLS is compared with ACS and CPS in columns 1 and 2, respectively.

Overall MxFLS migrants are more likely to be male than migrants in the comparison surveys. This is reversed after controlling age, education, employment and home ownership: relative to ACS and CPS, MxFLS migrants are statistically significanlty less likely to be male.

Comparing the conditional age distributions across surveys, MxFLS respondents are less likely to be age 25 to 34 years relative to age 16 through 24 years old. For ACS and CPS, the age differences with MxFLS are only statistically significant for males. In addition, MxFLS respondents are significantly less likely to be married than those included in ACS, CPS and MMP. These differences are significant for both males and females.

Table 3 established substantial differences in the distribution of education between MxFLS and both ACS and CPS. These differences are large and significant in the conditional distributions for both males and females. Drawing comparisons with ACS respondents, relative to those who completed primary school or less, MxFLS respondents are 5 percentage points less likely to have completed high school. This difference is large: recall that only 20 percent of MxFLS respondents reported completing high school. The difference is important given the income gaps between those who do and do not complete high school.

We also find that MxFLS respondents are significantly more likely to be working than ACS and CPS respondents. Finally, MxFLS migrants are more likely to be renting their place of residence than those included in the U.S. surveys.

3.3 Comparisons of migrant earnings

We turn next to an examination of earnings reported by migrants in MxFLS, ACS and CPS. This comparison is important given the central role that earnings, and wages, have played in studies of the economic success and assimilation migrants to the U.S. We compare monthly earnings in the main job received during the past year. Figure 1, presents box-and-whisker plots of earnings by gender which provide a simple graphical representation of the earnings' distributions. The colored box represents the 25^{th} to 75^{th} percentiles of the distibution and the line inside the box represents the median. The upper and lower whiskers extend to the 95^{th} and 5^{th} percentiles of the distributions, respectively.

For both males and females, median monthly earnings are very similar across all three surveys but the distributions of earnings in MxFLS are far more concentrated than in either ACS or CPS. For example, among males, median earnings is about US$ 1,400 per month and the inter-quartile range is about US$600 in MxFLS but about US$900 in both ACS and CPS.[3] The right tails of the distributions are considerably longer in CPS, relative to ACS and those tails are longer than in MxFLS. It is this difference that drives the lower mean earnings of MxFLS

[3] The 25^{th} percentile in MxFLS is US$ 1,110 compared with US$ 1,075 in ACS and US$ 992 in CPS. The 75^{th} percentile in MxFLS is US$ 1,700 compared with US$ 1,900 in ACS and US$ 1,943 in CPS.

respondents reported in Table 3. Studies that use ACS or CPS to compare the mean earnings of Mexican migrants with that of the native born will substantially overstate the economic success of migrants in the U.S.

In order to explore whether the differences in the earnings distributions can be attributed to the differences in age and education, we have estimated Mincer-type models of (the logarithm) of earnings. Results are reported in Table 5. The covariates included in the model are indicators for gender, different age groups, education levels, marital status, and fixed effects for the state of residence in the U.S. at the time of the survey. Each covariate is interacted with an indicator for the respondent being in MxFLS. The models are estimated to draw comparisons between MxFLS and ACS (in column 1) and between MxFLS and CPS (in column 2). OLS models are reported for all respondents in panel A and for males in panel B. Least absolute deviation models are reported in panel C.

The interactions between MxFLS and each covariates provide information on how the estimated 'returns' to that characteristic differs in a study that relies on migrants who are interviewed in the comparison survey, relative to migrants interviewed in MxFLS. The only significant interactions are for the education indicator variables. Among ACS and CPS respondents, the difference between the earnings of those who only completed primary school or less and those who attended but did not complete high school is very small. (The estimated difference is 1.3 percent for all migrants in ACS and CPS.) However, among MxFLS respondents, those who attended but did not complete high school earn 17 to 18 percent more on average (or 12 to 15 percent more at the median). The difference for males is 15 to 21 percent. These are all large and significant.

According to ACS and CPS, migrants who completed high school or more earn about 13 percent more than those who completed primary school or less. According to MxFLS, these respondents earned 28 percent more than the least educated – the returns to completing high school or more is between 14 and 15 percent higher according to MxFLS, relative to the US-based sample. This difference is significant in the median regressions and in the regressions that are restricted to males.[4]

In sum, the relationship between earnings and education is much flatter for recent migrants in the ACS and CPS samples relative to recent migrants interviewed in MxFLS. There are two possible explanations. High earning well-educated migrants from Mexico are under-

[4] The small sample in CPS has very few respondents who completed primary school or less and so it is difficult to estimate the return to some high school and completing high school, relative to this group.

counted in ACS and CPS relative to MxFLS. It is clear from Table 3 and Figure 1 that is not the case. An alternative explanation is that low earning, less educated migrants are under-counted in ACS and CPS relative to MxFLS. This is far more plausible explaination. It is consistent with the interpretation that transitory, circular migrants are under-counted in the US-based samples and suggests that US-based samples overstate the economic success of Mexican-origin migrants, particularly among those with low levels of education. [5]

3.4 Comparisons with MMP

It is instructive to compare MxFLS with respondents in the MMP which is a ethnographic survey of communities in Mexico that are expected to have high rates of migration to the U.S. In the MMP two to five Mexican communities are surveyed each year during the months of December and January of successive years, when most U.S. seasonal migrants return to Mexico. Within these communities a random sample of about 200 households is selected.

The MMP is not designed to be representative of the Mexican population or the population of Mexicans living in the U.S. However, the communities included in the survey are selected to provide a broadly representative cross-section of communities where out-migration is likely. Moreover, the MMP is not a representative of Mexican migrants. It does not necessarily represent all migrants who returned to Mexico or all who continue to reside in the United States. By design, the MMP would be more likely to include seasonal migrants and migrants with roots in Mexico than MxFLS.

The MMP sample we consider includes all household members who were reported as being on a trip to the U.S. at the time of the interview. The definition of a U.S. trip in this context is a visit to the U.S. that involves work, an active job search, or a reasonably stable residency. This means that the respondents who were currently on a U.S. trip were still living in the U.S. at the time of the interview. Some of them were back in Mexico for a visit and others were reported to be in the U.S. by another household member. We consider the respondents from communities interviewed after 2000 who report that the current trip lasted 3 years or less. Appendix Table 1 presents summary statistics for the MMP sample.

[5] We have estimated quantile regressions for the 25th and 75th percentiles of the earnings' distribution. The flatter association between education and earnings in the CPS and ACS still holds, except for the CPS at the 75th percentile. In addition, we check whether the differences in the relationship between education and earnings are driven by imputed values to non-responses. In the ACS and CPS samples, a significant number of cases have imputed earnings. Conditional on working, 34 percent of ACS cases and about 20 percent of CPS cases have imputed earnings. We run the earnings' regressions excluding the imputed cases and the results are not affected.

Relative to the MxFLS respondents, this sample of MMP respondents is more likely to male, married and much more likely to be working. The distribution of education of MxFLS and MMP respondents is similar although the MMP sample is less well educated and includes a larger fraction of males who did not progress beyond primary school. Recall that, relative to ACS and CPS, MxFLS respondents are more likely to be male, unmarried, working and are less well educated. In terms of these characteristics, apart from marriage, the MxFLS sample lies between the MMP sample and the ACS/CPS. Recall the MMP respondents tend to be more transitory and the higher marriage rates probably reflects the fact that the MMP respondents are more likely to have maintained roots in Mexico. Taken together, these comparisons suggest that at least part of the difference between the MxFLS respondents and the ACS/CPS respondents is driven by the latter undersampling more transitory migrants in the U.S.

4. Assessment of the evidence

We turn now to a discussion of some of the key the assumptions underlying the analyses discussed above and an assessment of the quality of the data that we have examined.

4.1 Sample sizes

A legitimate concern in that the sample of migrants interviewed in MxFLS is small. This reflects the fact that the baseline MxFLS interviewed about 35,000 respondents and, of them, it is estimated that about 750 moved to the U.S. In principle, the size of the sample of migrants should not affect out estimates of location but will affect inferences about significant differences. The fact that many of the differences between MxFLS and ACS or CPS are statistically significant suggests that the samples are large enough to be informative. Note that the CPS sample of recent migrants is about the same size as MxFLS.

4.2 Attrition in MxFLS

A key assumption at the foundation of the interpretation of the comparisons between MxFLS and the other surveys is that the sample in MxFLS is representative of the population of recent migrants in the U.S. This assumption may not be true if attrition in MxFLS is selective. In this section we explore the extent to which attrition can explain the results.

Table 1 presents the tracking results for the follow up done in 2005-2006. The number of adult individuals interviewed in 2002 who were not found in 2005 is approximately 2,800. This implies an 11 percent attrition rate for the entire adult sample independently of their location.

During the tracking period for MxFLS-2, interviewers relied on informants from the members of same household and community to obtain information about the potential destination areas of migrants. Special attention was paid to identify potential movers to the U.S. Thus, although some respondents were not found, we know their likely location from the tracking data. Assuming that this information is reliable, we can assess whether attrition in MxFLS is a problem by looking at the baseline characteristics of the group that attrited but was known to be in the U.S. in 2005.

In Appendix Table 2 we compare the characteristics measured at baseline for the sample known to be in the U.S. in 2005. The first column presents the summary statistics for the entire sample, the second column for those who were found in the U.S. and the last column for those who could not be found. Since the contact rates were very high, only 64 adult respondents were known to be in the U.S. but could not be contacted in 2005. Note that, in 2002, the respondents who attrited were more likely to be female (45 percent versus 36 percent of those found), older (27 years versus 24 years), and more educated (about 23 percent had more than 12 years of schooling, compared to 9 percent of those found).

We can adjust the comparisons presented in Tables 2 and 3 using the baseline characteristics for the lost sample and make assumptions about the indicators measured in 2005. To be conservative, we assume that all individuals who were lost to follow-up completed an additional 4 years of education, they all got married, none of them work in the U.S and none of them rents their home. These assumptions would make the differences between MxFLS and the U.S. samples smaller. The test is to check whether, when we add the sample that attrited under these extreme assumption, our findings change substantially.

Appendix Table 3 presents summary statistics pooling the lost and found samples. We can see that, although the differences become smaller between the MxFLS and U.S. surveys the main patterns do not change much. Appendix Table 4 presents a table analogous to Table 3 but uses the extended MxFLS sample that adjusts for attrition. We find that some of the differences become statistically insignificant. Now, MxFLS respondents are as likely to be female, single and they are not more likely to be renting when we compare MxFLS and the two U.S. surveys. However, even under the most conservative assumptions, we still find that MxFLS respondents

are significantly younger, less educated and more likely to be working than respondents in ACS and CPS.

Taken as a whole, given that attrition was relatively low, the differences between the sample that attrited and the one found do not seem large enough to explain the key discrepancies between the MxFLS and U.S. surveys.

4.3 Measurement error

An additional issue that could drive part of the differences between the MxFLS and U.S. surveys is measurement error in arrival dates in ACS and CPS. There are legitimate concerns about the accuracy of restrospective dates of entry to the U.S. collected in survey data. To assess the extent to which this is an important problem for our comparison, we test alternative windows for date of arrival. Clearly, extending the window back will tend to result in the sample including more longer-term Mexican-origin people who are living in the U.S. Reducing the size of the window will shift the sample in the opposite direction. Moving the window has little impact on the overall results or our inferences about the differences between MxFLS and the U.S. samples. This is the case even when we restrict attention to only those Mexicans who report having arrived in the U.S. during the year immediately prior to the ACS and CPS interviews..

It is difficult to know whether measurement error in earnings explains the differences in the estimated returns to education in MxFLS and the U.S.-based samples. It is possible that the poorest educated in the U.S. samples overstate their earnings. It is also possible that MxFLS respondents with little education understate their earnings. We can provide some evidence on this question by examining data from the third wave of MxFLS, which was conducted between 2009 and 2011.

4.4 Preliminary evidence from the third wave of MxFLS

The third wave of MxFLS is currently in the field. Over 70 percent of the respondents who are thought to be living in the U.S. have been interviewed and we have compared this sample with migrants in the 2010 ACS and 2010 March CPS. The same socio-economic and demographic differences emerge in these comparisons: the U.S.-based surveys appear to systematically miss less educated, younger, unmarried and more transitory migrants. With these preliminary data, we do not find evidence that the returns to education are higher for migrants in MxFLS.

5. Conclusion

Undocumented migrants are hard to enumerate in surveys. This project provides evidence on the extent and nature of this problem using uniquely rich data from the Mexican Family Life Survey. MxFLS, an on-going longitudinal survey of Mexicans, follows and interviews movers to the U.S. In the first follow-up, we interviewed 91 percent of movers to the U.S. We compare respondents in MxFLS with recent migrants from Mexico interviewed in the ACS and CPS.

According to MxFLS, in 2005, there were about 2.25 million Mexicans who have moved to the U.S. in the prior three years. This is about 30% more than estimates based on the ACS. Moreover, the MxFLS respondents interviewed in the U.S. do not look the same in terms of demographic and socio-economic characteristics as those interviewed in the ACS and CPS. In particular, we find that MxFLS respondents are significantly more likely to be female, they are younger, significantly less educated, more likely to be single, more likely to be working and more likely to be renters. These differences support the idea that the U.S. surveys are undercounting undocumented and more transitory migrants. Not only are the differences statistically significant but they are also substantively important. For instance, MxFLS respondents are significantly more likely to have less than high school complete and be working.

Comparisons of the relationship between earnings and education indicate that the returns to education are estimated to be higher in MxFLS. If this is correct, it suggests that estimates of assimilation and economic success among the least educated may be overstated in U.S. samples.

Overall, our analyses provide evidence that the undercounting of migrants in U.S. surveys is substantial and selected on both demographic and socio-economic characteristics. These results have important implications for interpretation of evidence on the economic success and assimilation of migrants to the U.S.

References

Bean, F. D., Corona R., Tuirán R., Woodrow-Lafield K. A., and J. Van Hook (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): 411–22.

Bean, F. D., and J. Van Hook (1998). Estimating Unauthorized Migration to the United States: Issues and Results. In Migration between Mexico and the United States, Binational Study, vol. 2, ed. Mexican Ministry of Foreign Affairs and U.S. Commission of Immigration Reform. Austin: Morgan Printing, 511–50.

Donato, K. M., Durand, J., Massey, D. S. 1992. "Stemming the Tide? Assessing the Deterrent Effects of the Immigration Reform and Control Act," *Demography*, Vol. 29, No. 2, pp. 139–157.

Fernández-Huertas Moraga, J. (2011). New Evidence on Emigrant Selection. *The Review of Economics and Statistics,* 93(1):72-96

Hanson, G.H. (2006). Illegal Migration from Mexico to the United States. *Journal of Economic Literature*, 44(4), 869-924

Hoefer, M., Rytina, N., and C. Campbell (2006). Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2005. Washington, DC: Office of Immigration Statistics, Policy Directorate, U.S. Department of Homeland Security.

Ibarraran, P. and D. Lubotsky (2007) .Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican census. In G. J. Borjas (ed) *Mexican Immigration to the United States.* University of Chicago Press.

Massey, D. *Mexican Migration Project*, http://mmp.opr.princeton.edu/home-en.aspx

Massey, D. S., R. Alarcon, J. Durand, H. Gonzalez. 1990. *Return to Aztlan: The Social Process of International Migration from Western Mexico,* University of California Press. ...

Ong, P.M., and D. Houston. 2002. The 2000 Census Undercount in Los Angeles County. The Ralph & Goldy Lewis Center for Regional Policy Studies. Working Paper # 42

Orrenius, P.M. and M. Zavodny (2005). Self-selection among undocumented immigrants from Mexico. *Journal of Development Economics* 78, 215–240

Passel, J.S., and D'Vera Cohn 2008. *Trends in Unauthorized Immigration: Undocumented Inflow Now Trails Legal Inflow*. Washington, DC: Pew Hispanic Center.

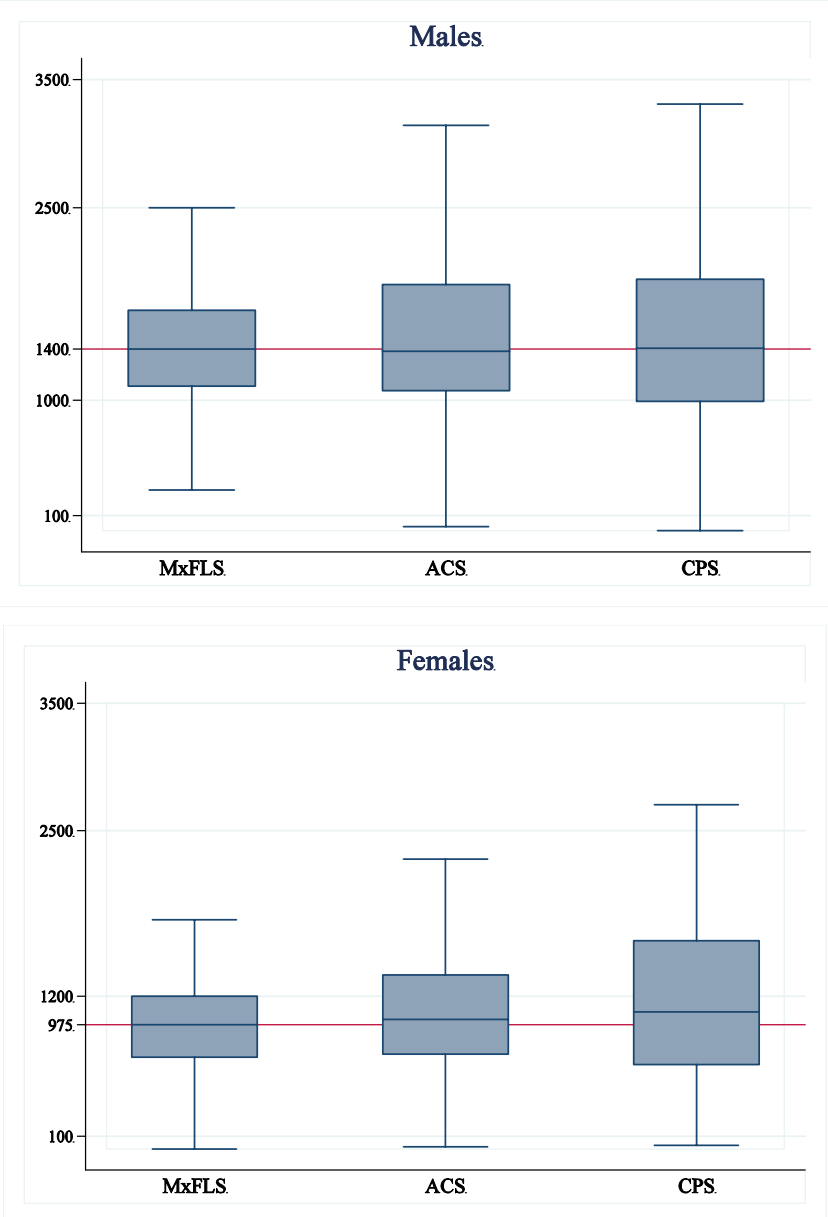
Passel, J.S., and D'Vera Cohn. 2009. *A Portrait of Unauthorized Immigrants in the United States*. Washington, DC: Pew Hispanic Center.

Passel, J.S., and D'Vera Cohn. 2009. *Mexican immigrants: How many come? How many leave?* Washington, DC: Pew Hispanic Center.

Rendall, M., P Brownell and S. Kups. 2011. Declining return migration from the United States to Mexico in the late 2000s Recession. *Demography.* 48.3:1049-58.

Figure 1

Box Plot for Earnings by Gender



Notes.- Sample consists of Mexican-born individuals 16 years or older who arrived to the U.S. between 2002 and 2005 and received earnings in the past year. Earnings are defined as monthly income from the main job received in the past year. The horizontal lines of each box show, from bottom to top, the 25th, 50th and 75th percentiles of earnings. The `whiskers' that extend from each box show the upper and lower `adjacent values'. The former is calculated as the largest data point smaller than the 75th percentile plus 1.5 x IQR, where IQR is the inter-quartile range. Similarly, the lower adjacent value is defined as the smallest data point greater than the 25th percentile minus 1.5 x IQR. The red line plots the 50th percentile for the MxFLS sample.

Table 1
Sample sizes and recontact rates in MxFLS-2005

	A. All ages		B. Age at least 16 in 2005	
	Number	%	Number	%
Baseline respondents re-interviewed				
Total	31,338	89.2	22,990	89.0
In Mexico	30,564	87.0	22,343	86.5
In US	774	2.2	647	2.5
% of migrants interviewed		90.6		91.0
Baseline respondents not re-interviewed				
Total	3,796	10.8	2,841	11.0
In Mexico	3,716	10.6	2,777	10.8
In US	80	0.2	64	0.2
Total respondents in MxFLS-1	35,134	100.0	25,831	100.0
who were subsequently in Mexico	34,280	97.6	25,120	97.2
who were subsequently in US	854	2.4	711	2.8

Notes.- Excludes panel respondents who died between 2002 and 2005.

Table 2: Estimates of size of population of recent migrants from Mexico to the United States

A. MxFLS 2005

	All		Males		Females	
Age group	**No**	**%**	**No**	**%**	**No**	**%**
Age 0-9	222,830	9.9	121,730	8.4	101,100	12.5
Age 10-14	180,731	8.0	118,447	8.2	62,284	7.7
Age 15-19	496,751	22.0	357,700	24.7	139,051	17.2
Age 20-24	615,728	27.3	402,457	27.8	213,271	26.5
Age 25-29	258,420	11.5	166,889	11.5	91,531	11.4
Age 30-34	165,029	7.3	71,509	4.9	93,520	11.6
Age 35-39	119,329	5.3	82,828	5.7	36,501	4.5
Age 40-44	85,952	3.8	58,242	4.0	27,710	3.4
Age 45+	111,290	4.9	70,112	4.8	41,177	5.1
Total	**2,256,059**	100.0	1,449,914	100.0	806,145	100.0

B. ACS 2005

	All		Males		Females	
Age group	**No**	**%**	**No**	**%**	**No**	**%**
Age 0-9	250,145	14.5	135,047	13.0	115,098	16.8
Age 10-14	101,213	5.9	55,540	5.4	45,673	6.7
Age 15-19	200,921	11.7	122,952	11.9	77,969	11.4
Age 20-24	376,097	21.9	242,726	23.4	133,371	19.5
Age 25-29	284,498	16.5	177,318	17.1	107,180	15.7
Age 30-34	184,775	10.7	113,135	10.9	71,640	10.5
Age 35-39	113,046	6.6	70,314	6.8	42,732	6.2
Age 40-44	86,497	5.0	51,675	5.0	34,822	5.1
Age 45+	122,028	7.1	66,382	6.4	55,646	8.1
Total	**1,719,220**	100.0	1,035,089	100.0	684,131	100.0

Table 3
Summary Statistics

	A. Males and Females			B. Males			C. Females		
	MxFLS	**ACS**	**CPS**	**MxFLS**	**ACS**	**CPS**	**MxFLS**	**ACS**	**CPS**
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Female	0.34	0.38	0.38						
	(0.02)	(0.01)	(0.02)						
Age	27.68	29.19	28.74	27.18	28.64	28.12	28.63	30.07	29.76
	(0.46)	(0.14)	(0.44)	(0.57)	(0.17)	(0.57)	(0.77)	(0.23)	(0.68)
Married	0.36	0.50	0.47	0.32	0.45	0.39	0.43	0.57	0.61
	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.03)	(0.04)	(0.01)	(0.03)
Education attainment									
Primary complete or less	0.37	0.31	0.26	0.40	0.32	0.23	0.33	0.29	0.31
	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.03)	(0.04)	(0.01)	(0.03)
High school incomplete	0.43	0.31	0.34	0.46	0.31	0.35	0.37	0.33	0.32
	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.03)	(0.04)	(0.01)	(0.03)
High school complete or more	0.20	0.38	0.40	0.14	0.37	0.42	0.30	0.38	0.37
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.03)	(0.04)	(0.01)	(0.03)
Employment									
Employed	0.78	0.65	0.66	0.94	0.84	0.85	0.49	0.33	0.36
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.04)	(0.01)	(0.03)
Earnings main job [(1)]	1418	1704	1596	1524	1814	1631	1027	1253	1453
	(36)	(41)	(76)	(43)	(50)	(73)	(43)	(44)	(246)
Renting	0.89	0.82	0.84	0.93	0.85	0.88	0.81	0.77	0.79
	(0.02)	(0.00)	(0.01)	(0.02)	(0.01)	(0.02)	(0.04)	(0.01)	(0.03)
Observations	645	8533	723	411	5044	421	234	3489	302

Notes.- The sample consists of Mexican born individuals 16 years or older who arrived to the U.S. in the past 3 years of the survey interview date. Summary statistics calculated using each survey's sampling weights.

Table 4

Multivariate models of differences between migrants in MxFLS and ACS or CPS

Linear probability models of probability respondent is interviewed in MxFLS

Coefficients on interaction between MxFLS respondent and covariate

	A. Males and Females		B. Males		C. Females	
	MxFLS and ACS (1)	MxFLS and CPS (2)	MxFLS and ACS (1)	MxFLS and CPS (2)	MxFLS and ACS (1)	MxFLS and CPS (2)
Female	0.106	0.093				
	[0.027]**	[0.035]**				
Age 25-34 [1]	-0.035	-0.036	-0.081	-0.104	0.046	0.088
	[0.032]	[0.040]	[0.038]*	[0.049]*	[0.053]	[0.063]
Age 35-44	-0.021	-0.003	-0.043	-0.040	0.040	0.054
	[0.042]	[0.054]	[0.054]	[0.073]	[0.062]	[0.076]
Age 45+	-0.025	-0.014	0.045	0.019	-0.104	-0.015
	[0.049]	[0.062]	[0.060]	[0.077]	[0.072]	[0.093]
Married	-0.092	-0.082	-0.093	-0.049	-0.115	-0.164
	[0.029]**	[0.037]*	[0.038]*	[0.049]	[0.044]**	[0.053]**
High school incomplete [2]	0.008	-0.033	0.033	-0.038	-0.039	-0.003
	[0.027]	[0.036]	[0.033]	[0.044]	[0.045]	[0.059]
High school complete or more	-0.198	-0.245	-0.270	-0.356	-0.107	-0.059
	[0.037]**	[0.045]**	[0.045]**	[0.054]**	[0.059]+	[0.071]
Employed	0.184	0.167	0.280	0.283	0.115	0.086
	[0.034]**	[0.042]**	[0.053]**	[0.064]**	[0.043]**	[0.053]
Renting	0.083	0.044	0.167	0.12	0.024	0.019
	[0.048]+	[0.057]	[0.050]**	[0.063]+	[0.068]	[0.078]
MxFLS respondent [3]	0.411	0.500	0.267	0.366	0.558	0.564
	[0.057]**	[0.073]**	[0.066]**	[0.091]**	[0.073]**	[0.092]**
Observations	9142	1332	5425	802	3717	530
R-squared	0.12	0.12	0.18	0.19	0.09	0.09

Notes.- Sample consists of Mexican-born individuals 16 years or older who arrived to the U.S. between 2002 and 2005. Robust standard errors in brackets. Regressions are weighted using surveys' weights. Dependent variable is an indicator variable that takes value 1 if the observation comes from MxFLS and 0 if it comes from the other survey.

+ significant at 10%; * significant at 5%; ** significant at 1%

[1] Omitted 16-24

[2] Omitted primary complete or less

[3] Indicator (1) if respondent in MxFLS. Represents difference for young, unmarried males with no schooling who are unemployed and not renting their home.

Table 5:Characteristics that predict migrant income in MxFLS, ACS and CPS

Dependent variable in log(monthly income)

Interaction of covariate with MxFLS reflects difference in association among MxFLS migrants relative to contrasting survey.

	A. OLS Males and females		B. OLS Males		C. LAD Males and females	
	MxFLS and ACS (1)	MxFLS and CPS (2)	MxFLS and ACS (1)	MxFLS and CPS (2)	MxFLS and ACS (1)	MxFLS and CPS (2)
Female	-0.405	-0.312			-0.336	-0.231
	[0.027]**	[0.107]**			[0.000]**	[0.039]**
Female* MxFLS	-0.038	-0.134			-0.069	-0.186
	[0.065]	[0.120]			[0.000]**	[0.054]**
Age 25-34 (1)	0.188	0.114	0.209	0.184	0.115	0.094
	[0.025]**	[0.096]	[0.028]**	[0.107]+	[0.000]**	[0.040]*
Age 25-34* MxFLS	-0.070	0.008	-0.054	-0.029	0.011	0.046
	[0.056]	[0.110]	[0.064]	[0.122]	[0.000]**	[0.054]
Age 35-44	0.237	0.180	0.261	0.301	0.164	0.208
	[0.035]**	[0.107]+	[0.041]**	[0.122]*	[0.000]**	[0.053]**
Age 35-44* MxFLS	-0.110	-0.037	-0.103	-0.154	0.018	-0.025
	[0.084]	[0.135]	[0.095]	[0.152]	[0.000]**	[0.076]
Age 45+	0.247	0.384	0.266	0.284	0.147	0.361
	[0.046]**	[0.215]+	[0.055]**	[0.160]+	[0.000]**	[0.069]**
Age 45+* MxFLS	-0.276	-0.403	-0.273	-0.298	-0.209	-0.397
	[0.088]**	[0.223]+	[0.078]**	[0.168]+	[0.000]**	[0.095]**
High school incomplete (2)	0.013	0.013	0.031	-0.044	0.006	-0.018
	[0.030]	[0.074]	[0.034]	[0.076]	[0.000]**	[0.040]
High school incomplete* MxFLS	0.165	0.184	0.152	0.211	0.120	0.148
	[0.047]**	[0.079]*	[0.050]**	[0.081]**	[0.000]**	[0.045]**
High school complete or more	0.134	0.134	0.151	0.007	0.072	0.076
	[0.028]**	[0.090]	[0.031]**	[0.092]	[0.000]**	[0.037]*
High school complete or more* MxFLS	0.141	0.150	0.143	0.267	0.151	0.181
	[0.050]**	[0.094]	[0.058]*	[0.100]**	[0.000]**	[0.059]**
Married	0.023	0.002	0.012	0.038	0.025	0.045
	[0.023]	[0.082]	[0.026]	[0.091]	[0.000]**	[0.037]
Married* MxFLS	-0.029	-0.016	-0.012	-0.048	-0.054	-0.080
	[0.055]	[0.099]	[0.054]	[0.105]	[0.000]**	[0.056]
Constant	7.075	7.099	7.048	7.109	7.090	7.090
	[0.052]**	[0.071]**	[0.061]**	[0.078]**	[0.000]**	[0.041]**
Observations	5776	871	4531	671	5776	871
R-squared	0.15	0.13	0.08	0.08	0.1	0.1

Notes.- Sample consists of Mexican-born individuals 16 years or older who arrived to the U.S. between 2002 and 2005. Robust standard errors in brackets. Regressions include state/region fixed effects. Regressions are weighted using surveys' weights. Dependent variable is the log of monthly income from the main job received in the past year.

+ significant at 10%; * significant at 5%; ** significant at 1%

(1) Omitted 16-24

(2) Omitted primary complete or less

Appendix Table 1
Summary Statistics for Mexican Migration Project

	A. Males and Females	B. Males	C. Females
Female	0.21		
	(0.02)		
Age	29.16	29.08	29.45
	(0.37)	(0.44)	(0.68)
Married	0.60	0.57	0.70
	(0.02)	(0.03)	(0.04)
Education attainment			
Primary complete or less	0.46	0.48	0.40
	(0.02)	(0.03)	(0.04)
High school incomplete	0.38	0.39	0.37
	(0.02)	(0.02)	(0.04)
High school complete or more	0.16	0.14	0.23
	(0.01)	(0.01)	(0.04)
Employment			
Employed	0.84	0.95	0.44
	(0.02)	(0.01)	(0.04)
Observations	1002	769	233

Notes.- The sample consists of Mexican born individuals 16 years or older who are currently on a U.S. trip.

Appendix Table 2
Summary Statistics adding Lost Migrants

	Sample known to be in U.S.		
Variables measured in 2002	**All**	**Found in U.S.**	**Not Found**
Female	0.37	0.36	0.45
	(0.02)	(0.02)	(0.06)
Age	24.30	24.00	27.30
	(0.42)	(0.43)	(1.65)
Married	0.28	0.28	0.28
	(0.02)	(0.02)	(0.06)
Education attainment			
Primary complete or less	0.44	0.45	0.34
	(0.02)	(0.02)	(0.06)
High school incomplete	0.46	0.46	0.44
	(0.02)	(0.02)	(0.06)
High school complete or more	0.10	0.09	0.23
	(0.01)	(0.01)	(0.05)
Observations	709	645	64

Notes.- Sample consists on panel respondents 16 years old or older in 2005 who were known to be in the U.S. during the 2005-2006 follow up. The variables are measured at baseline in 2002.

Appendix Table 3
Summary Statistics adding Lost Migrants

	A. Males and Females			B. Males			C. Females		
	MxFLS	**ACS**	**CPS**	**MxFLS**	**ACS**	**CPS**	**MxFLS**	**ACS**	**CPS**
	(1)	**(2)**	**(3)**	**(5)**	**(6)**	**(7)**	**(9)**	**(10)**	**(11)**
Female	0.37	0.38	0.38						
	(0.02)	(0.01)	(0.02)						
Age	28.43	29.91	29.29	28.08	29.30	28.67	29.03	30.90	30.28
	(0.43)	(0.14)	(0.44)	(0.52)	(0.18)	(0.58)	(0.73)	(0.23)	(0.69)
Married	0.42	0.50	0.47	0.39	0.45	0.39	0.47	0.57	0.61
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.01)	(0.03)
Education attainment									
Primary complete or less	0.35	0.32	0.27	0.36	0.33	0.24	0.32	0.30	0.32
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.01)	(0.03)
High school incomplete	0.46	0.29	0.32	0.48	0.29	0.33	0.43	0.30	0.29
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.01)	(0.03)
High school complete or more	0.19	0.39	0.41	0.15	0.39	0.43	0.25	0.40	0.39
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.03)	(0.03)	(0.01)	(0.03)
Employment									
Employed	0.71	0.67	0.68	0.86	0.86	0.87	0.45	0.35	0.37
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.03)
Renting	0.79	0.82	0.84	0.83	0.85	0.88	0.72	0.77	0.80
	(0.02)	(0.00)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.03)
Observations	707	8038	684	446	4741	394	261	3297	290

Notes.- The sample consists of Mexican born individuals 16 years or older who arrived to the U.S. in the past 3 years of the survey interview date.
Summary statistics calculated using each survey's sampling weights.
(1) Earnings are defined in monthly US$

Appendix Table 4
OLS regression for the probability of being included in MxFLS adding Lost Migrants

	A. Males and Females		B. Males		C. Females	
	MxFLS and ACS (1)	MxFLS and CPS (2)	MxFLS and ACS (1)	MxFLS and CPS (2)	MxFLS and ACS (1)	MxFLS and CPS (2)
Female	0.009	-0.005				
	[0.006]	[0.030]				
Age 25-34 [1]	-0.014	-0.069	-0.016	-0.063	-0.011	-0.073
	[0.006]*	[0.031]*	[0.008]+	[0.041]	[0.010]	[0.050]
Age 35-44	-0.014	-0.040	-0.018	-0.028	-0.010	-0.070
	[0.008]+	[0.042]	[0.011]+	[0.059]	[0.012]	[0.061]
Age 45+	-0.005	-0.013	0.006	-0.022	-0.018	-0.011
	[0.009]	[0.050]	[0.014]	[0.069]	[0.013]	[0.075]
Married	-0.008	-0.016	-0.005	0.016	-0.014	-0.074
	[0.005]	[0.029]	[0.008]	[0.041]	[0.008]+	[0.044]+
High school incomplete [2]	0.015	-0.025	0.020	-0.046	0.008	0.002
	[0.007]*	[0.032]	[0.010]*	[0.041]	[0.011]	[0.051]
High school complete or more	-0.041	-0.224	-0.047	-0.312	-0.035	-0.093
	[0.006]**	[0.034]**	[0.007]**	[0.043]**	[0.009]**	[0.054]+
Employed	0.027	0.073	0.020	0.051	0.032	0.081
	[0.006]**	[0.032]*	[0.009]*	[0.048]	[0.009]**	[0.043]+
Renting	0.000	-0.094	0.004	-0.122	-0.003	-0.052
	[0.006]	[0.036]**	[0.008]	[0.052]*	[0.009]	[0.051]
Constant	0.067	0.594	0.068	0.656	0.080	0.543
	[0.010]**	[0.053]**	[0.012]**	[0.072]**	[0.013]**	[0.068]**
Observations	9204	1394	5460	837	3744	557
R-squared	0.1	0.1	0.11	0.13	0.1	0.08

Notes.- Sample consists of Mexican-born individuals 16 years or older who arrived to the U.S. between 2002 and 2005. Robust standard errors in brackets. Regressions are weighted using surveys' weights. Dependent variable is an indicator variable that takes value 1 if the observation comes from MxFLS and 0 if it comes from the other survey.
+ significant at 10%; * significant at 5%; ** significant at 1%
[1] Omitted 16-24
[2] Omitted primary complete or less