Data Appendix

The data appendix describes the construction of the main variables. See http://www.ennvih-mxfls.org/ for more details regarding the sampling and data collection methods:

Migration status: The second round of the Mexican Family Life Survey (MxFLS-2) contains an indicator for individuals who are currently residing in the United States. This is supplemented by information from the migration modules in the second round of the survey on individuals who moved to the United States but who returned to Mexico by time of the second round. This is based on affirmative answers to the questions “Change of residence was to the United States?” for those who indicate migrations of more than 12 months since the first round and “Was the trip to the United States?” for those who indicate moves of less than 12 months since the first round. Return migrants are not excluded because the decision to return to Mexico is likely based on ex-post realizations of their experiences in the United States. Indicators for whether individuals have visited the U.S. in the past is based on affirmative answers to the questions from the first MxFLS round in 2002 “Change of residence was to the United States?” for those who indicate migrations of more than 12 months and “Was the trip to the United States?” for those who indicate moves of less than 12 months.

Education: Years of completed schooling is constructed on the basis of three different questions reported directly by individuals in book 3A: (1) “Which is/was the last level of schooling you attended?” with responses including “Without Instruction”, “Preschool or Kinder”, “Elementary”, “Secondary”, “Open Secondary”, “High school”, “Open High school”, “Normal basic”, “College”, or “Graduate”; (2) “Which is the last schooling grade you completed?” ranging from “Did not complete the first grade” to “Seventh grade or more”; and (3) “Did you obtain the degree that certifies you are a graduate at that level?” with responses including “Yes, you graduated/have a degree”, “Not yet graduated”, and “Have not finished all courses”. Question (2) was only asked for those who last attended “Elementary”, “Secondary”, and “High school” while question (3) was only asked for those who last attended “Open Secondary”, “Open High school”, “Normal basic”, “College”, and “Graduate”. In cases where this information is not reported, an analogous set of question is asked of a knowledge member of the household through proxy reports.

Years of completed schooling is taken as 0 for individuals who report “Without Instruction” or “Preschool or Kinder”. For individuals who report “Elementary”, years of completed schooling ranges from 0 to 6 based on the number of grades completed (there are no responses for more than “sixth grade”). For individuals who report “Secondary”, years of completed schooling ranges from 6 to 9 based on the number of grades completed (coding the less than 1% who report more than “third grade” as 9). For individuals who report “High school”, years of completed schooling ranges from 9 to 12 based on the number of grades completed (coding the less than 2.5% who report more than “third grade” as 12). For individuals who report “Open Secondary”, years of completed schooling is taken as 9 for those who graduated or have a degree, and 7 otherwise. Similarly, for individuals who report “Open High school” or “Normal basic”, years of completed schooling is taken as 12 for those who graduated or have a degree, and 10 otherwise. For individuals who report “College”, years of completed schooling is taken as 16 for those who graduated or have a degree, and 14 otherwise. For individuals who report “Graduate”, years of completed schooling is taken as 18 for those who graduated or have a degree, and 17 otherwise.

Years of completed schooling is also constructed on the basis of the roster reported on individuals by the head of household (or whoever responds to the household interview). However, only questions (1) and (2) were asked through the roster. Consequently, it is not possible to distinguish graduates from non-graduates for those who last attended “Open Secondary”, “Open High school”, “Normal basic”, “College”, and “Graduate”. Instead, the proportion of graduates derived from the individual reports is used to impute the average years of completed schooling for those reporting these categories, and these are rounded to the nearest integer. The main variable in the analysis uses information based on individual reports and supplements missing observations with information from proxy and roster reports.
Employment: Information about whether individuals have worked during the past year is based on the following question from book 3A: “In the last 12 months, have you worked (or developed any activity to help household expenditures)?”. In cases where this information is not reported, an analogous set of question is asked of a knowledge member of the household through proxy reports. Employment information is also constructed on the basis of the roster reported on individuals by the head of household (or whoever responds to the household interview). The question is “During the last 12 months did (...) work or develop any activity to help household expenditure” and those who are reported in the affirmative are coded as employed. The main variable in the analysis uses information based on individual reports and supplements missing observations with information from proxy and roster reports.

Earnings: Earnings in the primary and secondary jobs are reported by individuals for the past year in book 3A. For employees, rural laborers, agricultural and non-agricultural workers, earnings are reported on the basis of the following question: “How much did you earn in the last 12 months, since [date 12 months ago] until today, for working as [main job]?”. For annual earnings in a primary job, individuals were asked to provide the detailed amount in the following categories: wages or salary (after taxes), piecework, commissions and tips, extra hours, Christmas bonus, additional bonus, vacation premium, profit distribution, meals, housing, transportation, medical benefits, and others. For earnings from secondary jobs, individuals were asked to report the total amount. And if individuals did not provide a detailed breakdown for their primary job, they were also simply asked to report the total amount. Annual earnings are taken as the sum of earnings from the primary and secondary job, based on the detailed earnings or the total earnings depending on which was reported. In cases where this information is not reported, an analogous set of question is asked of a knowledge member of the household through proxy reports.

For business proprietors, employers, self-employed workers, and peasants who work on their own plot, earnings are reported on the basis of the following question: “How much money did you earn from working as [main job] during the past 12 months, since [date 12 months ago] until today?” Individuals report gross income/profits and net income/profits. Earnings for business proprietors, employers, self-employed workers, and peasants who work on their own plot, are taken to be equal to net income/profits, unless missing, in which case gross income/profits are used. Again, in cases where this information is not reported, an analogous set of question is asked of a knowledge member of the household through proxy reports.

We combine earnings information for workers and gross profits (or net profits if gross profits are missing) for business proprietors, employers, self-employed workers, and peasants. The main variable in the analysis uses information based on individual reports and supplements missing observations with information from proxy and roster reports. However, for individuals with missing self-reported earnings information, we also consider an alternative specification where we imputed earnings based on age, years schooling, marital status, urban status, state-level fixed effects and earnings information as reported by the head of household. We drop outliers that appear to be data errors.

Other variables: Information on marital status and age is available from the roster file. Household assets are taken as the total value of all household asset categories by the head of household, which include dwellings and land, bicycles, motorcycles and automobiles, electronic and kitchen appliances, savings and financial assets, farming equipment, and livestock. Credit and borrowing information is available from individual reports. An indicator for having the ability to borrow is based on an affirmative answer to the question “Do you know any person or place where you can borrow or ask for a credit?” Self-reported health is elicited in response to the question “Currently, could you say that your health…” on a scale of 1 (very good) to 5 (very bad). Cognitive ability is based on the number of correct responses to a Raven’s Progressive Matrices test with 12 multiple-choice questions. An indicator for having a relative in the U.S. is based on an affirmative answer to the question “Do you have any relative living in the US?”

Proxy information: A proxy book was fielded to collect information about household members that could not be interviewed in person, mainly because they were absent during the time of the fieldwork. Information in the proxy is based on a more limited set of identical questions to those presented in the individual one-on-one interviews, and provided by a household member knowledgeable about the absent member. Proxy reports account for approximately 10-15 percent of observations for most variables.