International Migration, Regional Variation, and Human Capital Formation in Mexico

Fernando Riosmena

ABSTRACT

In this paper we estimate net international migration from/to Mexico between 1990 and 2005 by sex, age, and level of education. While there are various estimates using (mainly) US sources, we aim to provide a more detailed set of estimates by (5-year) period, region and level of education. We use data from Mexican censuses, population enumerations, the educational system, and surveys to fit an extension of the Hill Method, an indirect estimation technique akin to demographic analysis. Our formulation extends the method to obtain region- and education-specific estimates. These calculations will provide more detail on the timing, regional scope, and social characteristics of recent immigration from Mexico. In addition, they can be used for decomposing change in human capital stocks, and as a basis to project the size and composition of the labor force as Mexico undergoes the main phase of its demographic 'dividend' in the next few years.

INTRODUCTION

International migration is a demographic force whose imprint has recently expanded from its original regional foci in both the United States and Mexico. Despite its predominantly surreptitious character, we do have a clear idea of the degree of magnitude of the phenomenon and its recent evolution thanks to the work of demographers using a variety of indirect estimation techniques (Bean et al. 2001; Hill and Wong 2005; Passel 2005). In addition, mostly thanks to many community-based studies but also to a few analyses of population-based data, we know that the regional distribution of migrants from Mexican origins (and US destinations) has changed in significant and variegated ways (Durand, Massey, and Charvet 2000; Massey, Durand, and Malone 2002; Zúñiga and Hernández-León 2005). We also know a great deal about the social characteristics of migrants, though it is no settled matter in the literature if Mexican migrant selectivity is (slightly) positive or (slightly) negative (Chiquiar and Hanson 2005; Ibarraran and Lubotski 2005; Orrenius and Zavodny 2005).

Despite these advances, studies on stocks/flows have been relatively disjoint from issues of regional variation[1] and –especially- social characteristics. Yet not necessarily a flaw, there is no insurmountable barrier refraining us from knowing some more about the characteristics of recent flows of Mexicans while estimating their quantity. While we do not claim that this is a better way to analyze the social characteristics of migrants (if anything, as these quantities first represent net migrants and thus two separate processes), it indeed represents a parsimonious, population-based way to study some detail about population movement. Moreover, these estimates can also be used to assess how international migration, along with other demographic factors, have helped change (or preserve) the composition of human capital stocks at the regional

[1] A clear exception is an analysis performed by the National Population Council of Mexico –the government-funded demography think tank and population policy-implementing wing- where indices of migratory intensity were derived at the municipal level (CONAPO INSYR).

level, and serve as a basis for the future projection of the labor force and some of its characteristics. This is extremely relevant from a policy perspective for both Mexico and the US as the first baby boomers start to retires in the latter while the former is about to begin 10-15 years of demographic "bonus" (Jackson 2005).

In this paper we perform an estimation of net international migration from/to Mexico between 1990 and 2005 by sex, age, and level of education as part of a broader analysis of recent changes and future perspectives in said age-structured, socially-differentiated population. While there are various estimates using US sources (and to a lesser extent, Mexican ones) we aim to provide with a more detailed set of estimates by (5-year) period, region and level of education by taking advantage of several data sources and different estimation methods. We hereby describe the general methodology to be used and the data sources thus far available to us. Definitions for educational levels and regions (i.e. conglomeration of states) are finally shown in the Appendix, though we are considering their expansion.

DATA AND METHODS

First and foremost, our method can be regarded as relatively simple extension of the so-called Hill method (see Hill and Wong 2005), an indirect estimation technique in which net migration by age (and sex) is residually obtained from an (age-graded) balancing equation:

$$ {}_nNM_x = {}_nN_x^2 - {}_nN_x^1 - B_x + B_{x+n} + {}_nD_x \qquad (1) $$

In this case, birthdays to age x (B_x) constitute entries into the population while birthdays to age $x+n$ (B_{x+n}) and deaths ($_nD_x$) constitute exits.[2] $_nN_x^1$ and $_nN_x^2$ are age-specific population counts in times 1 and 2 (separated by Δt years). The residual, deemed as an estimate of age-specific net migration out of the geographical unit of analysis, can of course be positive or negative. If said unit of analysis is a smaller aggregation to a country, then the residual naturally combines domestic inter-regional and international flows. Hence, one can further include an estimate of net domestic inter-regional migration (denotes as $_nR_x$, also positive or negative) to obtain international flows:

$$_nNM_x = {}_nN_x^2 - {}_nN_x^1 - B_x + B_{x+n} + {}_nD_x - {}_nR_x \qquad (2)$$

-FIGURE 1 ABOUT HERE-

The model depicted in Equation 1 has been estimated before using both Mexican and US aggregate census data (Hill and Wong 2005). We first propose to use Mexican sources only, but also including those coming from *complete* population enumerations (not used by Hill and Wong), known as *Conteo* and equivalent to a census in coverage though not in information breadth, performed in 1995 and 2005. By doing this, we estimate migration for 5-year (as opposed to 10-year) periods, thus providing a bit more detail about the evolution of migration flows in the past 15 years.[3] In addition, we propose an extension of Equation 1 in which the state-space is not only more specific in terms of geography (like in Equation 2), but where we also include educational transitions:

[2] Birthdays into age x between time 1 and time 2 are calculated from the geometric mean of adjacent age groups: $B_x = \Delta T / n \sqrt{{}_nN_{x-5}^1 \cdot {}_nN_x^2}$.

[3] We are also currently exploring the combination of the Hill and cohort survival methods (UN 1983) with one- to three-year age groups and cohorts to approximate the size of the flow in shorter time spans, along other alternatives.

$$ {}_nNM_x^{E_i} = \Delta_n N_x^{E_i} - \Delta B_{x,x+n}^{E_i} + {}_nD_x^{E_i} - {}_nR_x^{E_i} - {}_nE_x^{E_i} + {}_nX_x^{E_i} \quad (3) $$

The main difference between this model and the Hill method is that age groups are education-specific (to group E_i) and, naturally, so are birthdays, deaths, domestic inter-regional migration flows and –more importantly- the residual.[4] Additional entries and exits are of course those of people progressing through the education system (E and X denoting entries and exits respectively), which mostly occurs below age 30. Such formulation does therefore not only require data on education flows/intensities, but also information on the distributions of people by level of education for each census and enumeration point, as well as for deaths and domestic inter-regional migration flows. We describe the sources for each component and their estimation methodology, if any, next.

Population counts

Sex-, age-, region- and education-specific pyramids for 1990 and 2000 were obtained from the XI and XII General Population and Housing Censuses, readily available from INEGI (for 2000 and 2005 figures, see www.inegi.gob.mx, accessed September 20, 2005). The exact same information for 2005 comes from the II Population and Housing Enumeration. Unfortunately, basic tabulations from the first Enumeration, fielded in 1995, did not include as detailed information on educational attainment. We are currently exploring other data sources to calculate these distributions from (e.g. the sample from the Enumeration, or the 1992-1997 National Survey of Population Dynamics), multiply them by 1995 population counts, and thus obtain education-specific counts.

Vital Statistics

[4] We are currently considering ways in which to introduce estimates of measurement error to the analysis in a continuous way (that is, giving an uncertainty range for the error) in order to obtain a reasonable range for net international migration. In terms of our model, this would be basically done by adding an additional term to equation 3 (let us define it as ${}_n\varepsilon_x$) that would have some probability distribution (and could also be a function of other factors in addition to age, sex, region, and education, or a non-linear function of them at the very least).

Data on deaths by sex, age, state, education-level, and year *of registration* were generously provided by INEGI, which compiles data recorded in the National Civil Registry. Births by sex of child were also provided by INEGI. These data need to be adjusted by the fact that a given demographic event might have been registered after the year of occurrence. As INEGI does offer some information on the year of occurrence of events, we will use it to recover vital events occurring in a given year and subtract those occurring in previous years.

Domestic Inter-regional Migration

We estimate domestic inter-regional flows by age, sex, and level of education using the typical 5-year migration question. We obtain the level of regional and education-specific detail necessary for the analysis by using individual-level data from the census and enumeration samples (available from IPUMS at the University of Minnesota and INEGI respectively).

Education Flows

Education flows by state and level are available from Mexican Education Secretariat. However, these figures are not readily available by sex and age for each of the years under study. We have requested these data to the Education Secretariat (no response so far) and have explored some data published by the National Institute for the Evaluation of Education (www.inee.edu.mx). We are currently considering a few interpolation options and the use of additional data sources (e.g. the census/enumeration samples, the National Survey of Population Dynamics) to estimate educational transitions from prevalence figures.

REFERENCES

Bean, F.D., R. Corona, R. Tuiran, K.A. Woodrow-Lafield, 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):411-422.

Chiquiar, D., and G. Hanson. 2005. "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States". *Journal of Political Economy*, 113():239–281

Durand, J., D.S. Massey, and F. Charvet. 2000. "The changing geography of Mexican immigration to the United States: 1910-1996". *Social Science Quarterly* 81(1):1-15.

Hill, K. and R. Wong. 2005. "Mexico-US migration: Views from both sides of the border." *Population and Development Review* 31(1):1-+.

Ibarraran, P., and D. Lubotski. 2005. "Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census". Institute for Research on Poverty Discussion Paper No. 1308-05. University of Wisconsin. www.irp.wisc.edu

Jackson, R. 2005. "Building Human Capital in Aging Mexico". Report of the US-Mexico Binational Council. Center for Strategic and International Studies, and Instituto Tecnológico Autónomo de México. The CSIS Press. Washington, DC.

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

Orrenius, Pia, and Madeline Zavodny. 2005. "Self-Selection among Undocumented Immigrants from Mexico." *Journal of Development Economics* 78(1): 215-240.

Passel, J.S. 2005. "Estimates of the Size and Characteristics of the Undocumented Population." Washington: Pew Hispanic Center. www.pewhispanic.org/reports/report.php?ReportID=44 (accessed September 20, 2006).

United Nations (Population Division) and National Research Council. 1983. Manual X: Indirect Techniques for Demographic Estimation. Department for International Economic and Social Affairs. United Nations: New York. 304 pages.

APPENDIX – DEFINITIONS OF EDUCATION GROUPS AND REGIONS

Education levels

We currently use 5 education levels. Entry into an education level, unless otherwise noted, is granted by an individual having at least one year of education in it.

E1 – No formal education.

E2 – Primary.

E3 – Lower secondary (in Mexico, from 7^{th} to 9^{th} grade, inclusive), and technical studies that only require primary school.

E4 – Higher secondary and technical studies that do not require higher secondary.

E5 – University studies and technical studies requiring higher secondary. education (e.g. Associate's degree equivalent).

Regions

We define 4 regions, or conglomeration of states, in order to represent distinct areas of economic activity and socioeconomic development, as well as distinct international migration behavior in a parsimonious, data-friendly manner.

North – Baja California, Baja California Sur, Chihuahua, Coahuila, Durango, Nuevo León, San Luis Potosí, Sinaloa, Sonora, Tamaulipas.

West-Central – Aguascalientes, Colima, Guanajuato, Jalisco, Michoacán, Nayarit, Zacatecas.

South-Central – Distrito Federal, Hidalgo, Estado de México, Morelos, Querétaro, Puebla, Tlaxcala.

Southeast – Veracruz, Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Yucatán.