A Review of Methods for Estimating Emigration

Eric B. Jensen
Population Division
U.S. Census Bureau
Working Paper No. 101

September 2013

This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed are those of the author and not necessarily those of the U.S. Census Bureau. Please direct questions to the Net International Migration Branch by phone 301-763-9379 or email pop.international.migration@census.gov.

Acknowledgements

The author wishes to thank Elizabeth Grieco and Melissa Scopilliti, of the U.S. Census Bureau, for their overall guidance on this project. In addition, J. Gregory Robinson, Bashir Ahmed, Howard Hogan, and Victoria Velkoff of the U.S. Census Bureau provided content review.

This report reflects the work of the following individuals who conducted literature reviews in various languages and provided summaries of their findings:

Dalia Ambrozaitiene
Statistics Lithuania

Lara Badre
Central Administration of Statistics (Lebanon) and Medstat III

Víctor Alfredo Bustos y de la Tijera
National Institute of Statistics and Geography (Mexico)

Renuka Bhaskar
U.S. Census Bureau

Giambattista Cantisani
Medstat III

Jeronimo Cortina
University of Houston

Andre Cyr
Statistics Canada

Rima Igoševa
Vilnius University

Gindra Kasnauskiene
Vilnius University

Albert Kraler
International Centre for Migration Policy Development

Monica Perez
Italian National Institute of Statistics

Ignacio Duque Rodriguez
National Institute of Statistics (Spain)

Rafael Lopez Vega
National Population Council (Mexico)

Abstract

International migration statistics measure the movement of people across national borders and often comprise estimates of immigration (migration into a country) and emigration (migration out of a country). Emigration is one of the most difficult components of population change to estimate because the emigrant population is no longer resident in the country and, therefore, cannot be measured directly using censuses or surveys. Researchers and national statistical agencies have used various data, including population registers and surveys, and techniques, including residual methods, data attrition methods, indirect estimation, multiplicity sampling methods, and statistical modeling to estimate emigration. However, the literature describing how these data and techniques are used to estimate emigration is relatively sparse. In addition, foreign language translations are often unavailable, limiting the access of this information to an international audience. One of the projects of the Suitland Working Group is to bring together the published literature on estimating emigration from different languages into one resource document. In this report, we summarize the literature on the methods to estimate emigration, review the strengths and limitations of each method, and provide references for original documents in English, French, German, Italian, and Spanish.

Table of Contents

Introduction

International migration statistics measure the movement of people across national borders and often comprise estimates of immigration (migration into a country) and emigration (migration out of a country). Developing high quality estimates of international migration flows is essential for producing accurate population estimates. Emigration is one of the most difficult components of population change to estimate for several reasons. First, because the emigrant population is no longer resident in the country, it cannot be measured directly using censuses or surveys. In addition, both the native and foreign-born populations are at risk of emigrating. Finally, there are generally few mechanisms for ensuring that administrative data sources are updated following a migration abroad.

Because of the difficulties with estimating emigration, researchers and national statistical agencies have used a variety of data and techniques in an effort to produce timely and accurate statistics. These include population registries, migration surveys, residual methods, data attrition methods, indirect estimation, and multiplicity sampling methods. However, the literature describing the use of these data and techniques for estimating emigration is relatively sparse and difficult to find. In addition, foreign language translations of key research are often unavailable, limiting the access of this information to an international audience.

In 2009, the Suitland Working Group organized a meeting titled *Using Household Surveys to Measure Migration and the Size, Distribution, and Characteristics of Migrant Populations* at the U.S. Census headquarters in Suitland, Maryland, outside of Washington D.C. [1] Several projects were proposed during the conference to improve household survey data on migration specifically and international migration statistics in general. One of the projects was to produce a single source, such as a handbook, describing and synthesizing the methods used to estimate emigration. This report is the outcome of that project.

[1] The *Suitland Working Group* was organized during the joint UNECE/Eurostat Work Session on Migration Statistics held March 3-5, 2008 in Geneva, Switzerland. The *Suitland Working Group* is organized under the framework of the Conference of European Statisticians (CES) Work Plan on Improving International Migration Statistics.

This report presents the results of a literature search of published methods for estimating emigration that was conducted in several languages. Although much of this literature was written from the perspective of more developed countries, the methods reviewed can be broadly applied to all countries.[2] The report describes and synthesizes the methods to determine the strengths and limitations of each method. Finally, the report provides references for original documents in English, French, German, Italian, and Spanish.

The report is divided into sections based on the methodological approach. The first section reviews the use of administrative data for estimating emigration with particular emphasis on population registers. The next section reviews migration surveys including household surveys and port or passenger surveys. Residual methods are then presented followed by a section on methods that use the attrition of people from panel data to estimate emigration. The next section reviews the literature on indirect estimation and multiplicity sampling methods. Finally, statistical models of migration are briefly reviewed.

Population Registers

Population registers provide data that can be used to estimate native and foreign-born emigration (Cantisani and Greco 2006; Poulain et al. 2006; Salvisberg and Heininger 2007). Population registries are official lists or databases maintained by governments to record vital events such as births, deaths, and marriages; demographic characteristics such as age, sex, place of residence, place of birth, and nativity; and in some cases socioeconomic characteristics such as employment status and educational attainment. Registries are most often kept at the local or municipal level and connected at the central level. Moreover, some countries keep and centralize 'mirror versions,' or copies, for statistical purposes. Therefore, periodic reports of individual-level or aggregated data are eventually submitted to the national statistical office (Poulain et al. 2006).

Migration events, both domestic and international, are recorded as new registrations or de-registrations from the local register. In many systems, when a migrant registers in a new

[2] The context of emigration varies by the level of economic development, which could influence how emigration is measured. Migration flows in more developed countries tend to have higher levels of immigration than emigration and the emigrant population consists of both the native-born and foreign-born populations. In developing countries, migration flows are dominated by the emigration of the native-born population.

municipality they are required to list their most recent previous residence. The registry office at the destination will then notify the origin municipality of the move and the person will be de-registered. Many countries require nationals living abroad to register at the consulate of the destination country who then notifies the origin municipality of the move. Some countries, such as Italy, use a system with one registry containing information on the residents currently living in the municipality and a second registry of residents that are currently living abroad (Cantisani and Greco 2006). However, in Scandinavian countries, this notification of change of residence is electronically implemented between pairs of countries (Poulain et al. 2006).

There are several mechanisms that countries use to maintain accurate and up-to-date population registers. Countries often establish time limits whereby a migrant must register in the municipality. These time limits can range from a few days to several months after moving into the destination municipality. Incentives for registering are also used to maintain the quality of the database. For instance, a country may require a person to be registered in order to enroll their children in school or access the health system. Similarly, a country may use incentives such as reduced taxes to encourage nationals living abroad to de-register from the population register in their origin country.

Strengths of the methodology

Population registries are a useful source of data for estimating emigration. Depending on their level of centralization and linkage with other administrative or statistical sources, population registries provide extensive information about the size and characteristics of the emigrant population. Also, registries are much more current than surveys and censuses which allow for estimates to be produced on an annual or even quarterly basis.

Limitations of the methodology

There are challenges to using data from population registries to estimate emigration. In countries without centralized population registries, there can often be considerable variation in the quality of data that central offices receive from individual municipalities. Coverage of the population register and differential coverage for certain groups, such as migrants leaving without their family and undocumented migrants, can bias estimates produced using registers. It is especially

important for data from population registries to be up to date. If there is a significant time lag between changes in residence and when the register is updated, emigration will be underestimated. The most significant limitations to using population registers for estimating emigration is that emigrants often fail to have themselves de-registered from the population register in the origin municipality. This will also produce an underestimate of emigration.

Migration Surveys

Estimates of emigration can be made using special surveys of migrants. These may include household surveys that ask retrospectively about changes in the country of residence of respondents or household members, port surveys that ask travelers about their intentions to change their country of residence, or even surveys which are sent to nationals living abroad (Rendall, Tomassini and Elliot 2003; Silvestrini and Cariani 2005). In this section, the International Passenger Survey (IPS) which is a port survey conducted in the United Kingdom and the EMIF-North which is a migration-specific household survey conducted in Mexico are reviewed as examples of migration surveys used to estimate emigration. The strengths and limitations of the IPS and EMIF-North surveys, along with those of migration surveys in general, are discussed together.

International Passenger Survey (United Kingdom)

The International Passenger Survey (IPS) is a survey of travelers both entering and leaving the United Kingdom (UK) by air, sea, or the Channel Tunnel. The survey is the main source of migration data for the UK (United Kingdom Office for National Statistics 2012). The IPS is a continuous survey that has been conducted since 1961. The sample size is approximately 250,000 individuals each year. Official estimates of emigration from the UK are produced by starting with emigration counts from the IPS and adding Northern Ireland migration outflows, visitors that originally entered the UK intending to stay for less than a year but ended up staying for longer (visitor switchers), and asylum seeker outflows. Finally, travelers from the UK that declared that they would stay in the destination country for more than a year but left sooner (migrant switchers) are subtracted to calculate the final estimates of emigration.

EMIF-North (Mexico)

The Survey on Migration at the Northern and Southern borders, EMIF (*Encuestas Sobre Migración en las Fronteras Norte y Sur de México*), is a household migration survey conducted along the northern and southern borders of Mexico. Because border crossings by emigrants from Mexico are predominantly at the northern border, the discussion in this section is focused on the EMIF-North survey. The EMIF-North captures four distinct migration flows 1) migrants coming from the south of the border, 2) migrants from the border region, 3) migrants voluntarily returning from the United States, and 4) migrants deported from the United States. The EMIF-North uses a multistage probability spatiotemporal sampling frame in which geographical (region, city, zone, crossing point) and temporal (trimester, day, hour) units are chosen interactively. Prior knowledge about the flow of migrants is essential to the survey design of the EMIF-North because cities, zones, and crossing points are sampled based on the probability proportionate to the flow of migrants to that area (region has a probability of one since all border regions are included). In terms of the temporal units, the EMIF-North selects all trimesters in a year with a probability of one. The days are selected deterministically to facilitate fieldwork, ensure that all the instruments that constitute the EMIF-North and EMIF-South are represented in that particular day, and maintain consistency across survey instruments (EMIF Norte 2010).

Strengths of the methodology

The overall strength of migration surveys is that they provide much more specific data on international migration than general household surveys or censuses. In fact, migration surveys and qualitative data on migrants can potentially provide much more detail about the motivations for migration, the characteristics of migrants, and the overall migration experience than other data sources. Migration surveys have the potential to be more current than censuses and, therefore, may capture new or emerging trends. Because migration surveys often sample respondents throughout the year, they can measure seasonal variations in emigration.

Limitations of the methodology

There are limitations to using migration surveys to measure emigration. Migration surveys can be expensive and time consuming. National statistical agencies may not have the available

resources to regularly conduct migration surveys. Emigration is a rare event, and even within the sample of passenger surveys, the number of people moving abroad is quite small and these estimates are often prone to error (Zaba 1985). Migration surveys are also limited by coverage error, or the under-representation of the full population in the sample.

Residual Method

The residual method is a commonly used technique for estimating emigration. The method is based on the interrelationships between the demographic processes that cause population change, or the population balancing equation. The basic approach of this method is to survive estimates of the population from the first census (along with estimates of immigration during the intercensal time period) forward to the next census and then subtract the enumerated or actual population from the survived or expected population to get a residual estimate of emigration. Because emigration of the native population is usually a rare event, this approach is mainly used to estimate the emigration of the foreign-born population. However, the residual method can also be used to estimate native emigration using data from other countries. This variant of the method—the foreign-census method—is also presented in this section.

Foreign-born Emigration

The residual method begins with the enumerated foreign-born population in a census at time 1 and survives that population forward to the next census at time 2, using life table survival rates.[3] Immigration during the interval is added. Emigration is then calculated as the difference or residual between the *expected* foreign-born population and the *actual* foreign-born population enumerated in the later census (Warren and Peck 1980). This method can be expressed as Equation (1):[4]

$$E_{t1-t2} = P_{t1} - D_{t1-t2} + I_{t1-t2} - P_{t2} \quad (1)$$

[3] The examples of the residual method cited here use decennial census data from the United States. While the residual method for estimating emigration most often uses data from consecutive censuses, other sources of data could be used including survey and administrative data.

[4] In this report, we use upper-case letters to denote population counts, estimates, or levels and lower-case letters to denote rates and probabilities.

where E is the number of emigrants, P_{t1} is the foreign-born population count from the first data source, D are deaths to the foreign-born population during the interval, I is new immigration during the interval, and P_{t2} is the foreign-born population count in the second census. Estimates of emigration can be calculated separately by sex and some specified age groups, depending on the robustness of the data being used.

The model presented in Equation (1) provides a simplified version of the residual method. In actuality, E represents emigration and also the sum of all errors introduced in each of the previous steps (Warren and Peck 1980). Adjustments are needed to reduce the amount of error that ends up in the residual. For instance, the census counts of the foreign-born population *(P_{t1}* and *P_{t2})* could be adjusted for net census coverage if the coverage error is known. Deaths to the foreign-born population during the interval could be another source of error if life tables specific to that population are not available. Other refinements that help reduce the amount of error that is included in the residual include adjusting for persons who died after emigrating, error in the classification of nativity, and the net migration of foreign students.

Ahmed and Robinson (1994) later modified the census-to-census residual methodology to eliminate the need to estimate immigration during the time interval. They began with an enumeration of the foreign-born population in the United States prior to 1980 and then survived this population forward using age- and sex-specific life tables. The survived foreign-born population was then subtracted from the enumerated foreign-born population in the 1990 Census that entered the United States prior to 1980. The difference between the expected foreign-born population (survived) and the observed (enumerated) foreign-born population becomes the estimate of foreign-born emigration. By focusing on the population that entered the United States prior to 1980, or time 1, this method is much simpler than other residual approaches that include estimates of immigration during the interval. The Ahmed and Robinson census-to-census residual approach can be formally written as Equation (2):

$$P_{t2} = P_{t1} - D_{t1-t2} - E_{t1-t2} \quad (2)$$

Where P_{t2} is the foreign-born population that entered before time 1 and was enumerated in the census at time 2, P_{t1} is the foreign-born population that entered prior to time 1 and were

enumerated in the census at time 1, $D_{t1\text{-}t2}$ are deaths to this population over the interval, and $E_{t1\text{-}t2}$ is the residual or the foreign-born population that emigrated during the interval. This can be seen in Equation (3):

$$E_{t1-t2} = (P_{t1} - D_{t1-t2}) - P_{t2} \quad (3)$$

$$\text{or,} \quad E_{t1-t2} = S_{t2} - P_{t2} \quad (4)$$

where S_{t2} is the expected or survived foreign-born population that entered the United States prior to time 1. Equation (4) gives an estimate of total emigration from the initial stock foreign-born population, or the population who entered the United States prior to time 1, from time 1 to time 2. The total emigration estimate can be used to calculate a 10-year emigration rate by dividing the total estimate by the population observed at time 1 and annual emigration rates can be calculated by multiplying the 10-year estimate by 1/10.

Equation (4) cannot be used to directly estimate emigration of the foreign-born population who entered the United States between time 1 and time 2 because the initial population is not specified in this equation. The initial population can be found using Equation (5):

$$P^*_{t2} = P_{t1-t2} - P_{t1-t2} * e - P_{t1-t2} * (1 - s) \quad (5)$$

where P^*_{t2} is the foreign-born population that came during 1980-1990 and was enumerated in the 1990 Census, $P_{t1\text{-}t2}$ is the foreign-born population that came during 1980-1990 and was at risk of both mortality and emigration, e is an emigration rate derived from Equation (4), and s is the survival rate. Equation (5) can be simplified to:

$$P^*_{t2} = P_{t1-t2} * [(1 - e) - (1 - s)] \quad (6)$$

$$\text{or,} \quad P_{t1-t2} = \frac{P^*_{t2}}{(s-e)} \quad (7)$$

Equation (7) will be true if $s > e$. Also, if P^*_{t2}, s, and e are known then we can estimate P_{t1-t2} by using Equation (7) where the level of emigration from time 1 to time 2 is estimated by multiplying P_{t1-t2} by e. This method assumes that the emigration rate of the more recently

immigrated foreign-born population (those who entered the United States during the interval) is the same as the foreign-born population who immigrated earlier.

Strengths of the methodology

In general, the residual method is fairly straightforward and easy to estimate. Residual methods are often calculated using census data and other sources, which are readily available for most countries. Also, the methodology can be used to produce emigration rates by selected demographic characteristics including sex and broad age groups. In addition, residual methods can be adapted to measure specific period of entry cohorts eliminating the need to estimate immigration during the interval, thus reducing the data requirements and complexity of the model.

Limitations of the methodology

There are two main limitations inherent in the residual method. The first is that the method is especially prone to errors. As noted above, the residual term *(E)* in Equation (1) denotes the estimate of emigration and the sum of all errors in the other terms of the equation (population at time 1, immigration, deaths, and population at time 2). If these errors are not taken into account then the method produces inaccurate estimates of emigration. Differential coverage error of the foreign-born population between the two censuses could contribute to error in the estimates of the population at time 1 and time 2 (Mulder, Guzman and Brittingham 2002). If the estimates are calculated by demographic characteristics, measurement errors in the data used to calculate the residual—age, sex, nativity, and year of entry—can contribute to errors in the emigration estimates. Also, if the life tables or survival rates are not available by nativity, the emigration estimates could be biased if there are significant mortality differentials between the native and foreign-born populations.

The second limitation of that residual method is that the estimates are not current and do not include estimates the emigration of temporary migrants who might immigrate and emigrate between censuses. Censuses are normally conducted every five to ten years, therefore, the time span for estimating emigration with this method can be relatively long. The residual method produces a cumulative emigration rate for the time interval; therefore year-to-year variations in

emigration rates are not detected. Another limitation is that the residual method does not measure emigration of recently arrived immigrants whose migration is temporary or short in duration. Studies using other methodologies have shown that temporary migration, or rates of emigration for recently arrived immigrants, can be significantly higher than those produced using the residual method (Van Hook et al. 2006).

Emigration of the Native-born Population

Emigration of the native-born population is typically a rare event, therefore, the method described above would be extremely sensitive to errors and would not accurately estimate emigration of this population. However, a variant of the residual method can be applied to data from other countries to estimate the *net* migration of natives. This method uses estimates of the stock of natives of a particular country in consecutive censuses, surveys, or population registries of other countries. The stock of natives counted at time 1 is survived forward to time 2 to determine the expected population in time 2. This expected population is then compared to the count or estimate of the population for this same time period. The difference is assumed to be the net migration of natives.

Gibbs et al. (2001) used this approach, supplemented with data from the U.S. State Department, to estimate the net migration of U.S. natives in the 1990s. The analysis used data from foreign census counts and estimates for sixteen countries with observed data for at least two points of time between 1995 and 2000. Specifically, Gibbs et al. used data from Australia, Belgium, Canada, Germany, Great Britain, Greece, Ireland, Israel, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland. For each of these countries, they began with an estimate of the U.S. native-born population at time 1 (which varied depending on data availability). This estimate was survived forward and compared to an estimate or count at time 2. The difference was then divided by the length of time between time 1 and time 2 in order to estimate average annual net migration. Because of the limited data available, U.S. State Department data were used to determine additional net migration flows. These data showed that 58 percent of natives living abroad resided in the countries included in their analysis. This proportion was used to inflate the estimates of net native migration to get an estimate of total net native migration between the United States and all other countries.

More recently, Schachter (2008) employed a similar approach using data from censuses and population registers in 84 countries to produce estimates of the net migration of U.S. natives. For each of the countries in the analysis, a time 1 estimate of U.S. natives was survived forward to a time 2. The resulting estimate of the survived population at time 2 was then compared to the estimate observed in the census or population register. This number was then divided by the number of years between time 1 and time 2 to develop average annual estimates of the net migration of U.S. natives. The definition of U.S. natives varied depending on data availability – where possible, natives were defined as those born in the United States. For some countries where place of birth data were not available, U.S. citizenship was used to define natives. The migration of U.S. natives for each of the 84 countries was summed to produce an estimate of the total net international migration of U.S. natives.

Strengths of the methodology

In countries where emigration of the native population is a rare event, there tends to be limited data and methods for estimating the outmigration of this population. Few censuses or surveys collect information on the number of people who have left a particular country and the data that are available may be limited by recall bias and small sample size. However, countries can estimate the number of native emigrants using data from other countries that identify their binationals. A particular strength of this method is that most countries collect data on the country of birth or citizenship of their resident population. Also, once the data from other countries have been compiled, the methodology to estimate emigration becomes fairly simple.

Limitations of the methodology

There are several important limitations to using a residual approach with foreign census data to estimate the net migration of natives. First, the quality of the data varies between countries and the coverage of immigrants (in the example above, the U.S. native-born population living in other countries) in foreign censuses is unknown. Potential undercoverage of the native-born population living in other countries in the various data sources could result in error in the estimate of net migration. Additionally, the impact of coverage errors in the overall estimates of net migration will be exacerbated if there is differential coverage error between the two censuses.

Therefore, change in the coverage of U.S. natives from one census to the next would affect the estimate of native migration using this methodology.

An additional limitation of the residual method involves identifying the population being measured. In the Schachter (2008) analysis, the definition of the U.S. native-born population varied between countries due to data availability. For countries that did not collect data on country of birth, information on country of citizenship was used to classify the U.S. native born. This results in inconsistencies in the measurement universe and potential overlap with estimates of migration of naturalized U.S. citizens or undercoverage of migration of U.S. nationals that were born abroad of U.S. citizen parents.

There are other limitations to using data from other countries to estimate emigration. First, differences in residency rules across countries and data collection time periods can produce error in the estimates. Also, the assumption that death rates of U.S native born living in the United States are similar to those living abroad might not be accurate. Finally, the assumption that migration within the period is constant and can be annualized by dividing the estimate by the number of years in the time interval could be potentially problematic if there is temporal variation in emigration.

Panel Data Attrition Methods

Panel data attrition methods may also be used to estimate emigration. These methods use the attrition of the foreign-born population from administrative data or household surveys as an indicator of return migration. The basic approach of these methods is to decompose the missing data, or attrition of the foreign-born population, from longitudinal panel data into its component parts, including emigration. In general, these methods are useful because they use administrative or existing survey data, which reduces the time and resources needed to create estimates of emigration. Also, data attrition methods can provide estimates of migration for relatively short durations, which would not be measured by the census-to-census residual method. Similarly, these methods produce timely estimates, which can be used to estimate annual flows of net migration.

There are limitations to data attrition methods for estimating emigration. Administrative data systems and panel surveys are typically not designed for measuring migration, therefore, the data sources might not contain all the data necessary to estimate emigration without making substantial assumptions about the data.

Two panel data attrition methods are reviewed here. The first uses linked administrative data from the former Immigration and Naturalization Service, now part of the U.S. Department of Homeland Security (DHS), on a cohort of legal immigrants to the United States to estimate cumulative emigration rates for that cohort.[5] The second attrition method uses data from matched Current Population Survey (CPS) files to estimate emigration rates for the total foreign-born population residing in the United States.

Linked Administrative Data

Jasso and Rosenzweig (1982) developed estimates of emigration rates for a cohort of legal immigrants to the United States that were admitted to permanent resident status in 1971. The method produces upper- and lower-bound estimates of cumulative emigration rates for the 1971-1979 period. The main source of data for the study was a probability sample (n=3,758) drawn from administrative data on the cohort of immigrants who were admitted for permanent resident status in 1971 (N=370,478). This cohort included both newly arrived immigrants and immigrants that were residing in the United States but had been admitted to permanent resident status that same year. The data were then matched annually to naturalization records and data from the Annual Alien Address Program (Alien Address file).[6] Life tables for the U.S. population were used to estimate deaths to individuals in the cohort. Finally, data on country of birth and year of entry from the Survey of Income and Education (SIE) conducted by the U.S. Census Bureau were used to develop the lower-bound estimates of cumulative emigration from 1971-1979.

[5] The Department of Homeland Security (DHS) of the U.S. Government was created in 2003 by merging several government agencies including the Immigration and Naturalization Service (INS). The data collection and dissemination duties of INS were relocated to the Office of Immigration Statistics (OIS) at that time.

[6] The Alien Registration Act of 1940 required all aliens entering or residing in the United States to provide address information. INS, through the Annual Alien Address Program, collected annual registration data from self-reports of persons applying for legal permanent residence. The Alien Address Program lasted from 1951 to 1981 when it was discontinued following an amendment to the Alien Registration Act.

The methodology used to develop the upper-bound cumulative emigration rates begins with Equation (8),

$$P = I - D - E \tag{8}$$

where *P* is the cohort admitted in 1971 that are still present, *I* represents the number of immigrants admitted for permanent resident status in 1971, *D* is the number of deaths to the 1971 cohort, and *E* represents emigration of that cohort. The immigrants' records may transition from the Alien Address file to the Naturalization file as their immigration status changes. This is accounted for in Equation (9),

$$P = C + R + N \tag{9}$$

where *C* is the number of cohort members who became naturalized citizens by the end of 1978, *R* is the number of cohort members who reported their address in the Naturalization and Alien Address files by January 1979, and *N* is the number of cohort members who were alive and resident in the United States in January 1979 but not registered with the DHS. Substituting (9) into (8) and solving for emigration *(E)* results in Equation (10).

$$E = I - D - C - R - N \tag{10}$$

Assuming that nonreports *(N)* to the 1971 cohort are zero, the above equation can be rewritten to express the upper bound of the emigration estimate where the estimate is upwardly biased by the total number of nonreports *(N)*.

$$E \leq (I - C - R - D) \tag{11}$$

Emigration rates *(e)* are then calculated by taking the number of emigrants estimated using Equation (11) and dividing by the number of immigrants in the 1978 cohort that would have survived to the end of the time interval (1979). The emigration rates can be expressed as Equation (12).

$$\frac{e}{I-D} = \frac{I-D-C-R}{I-D} = 1 - \frac{C+R}{I-D} \tag{12}$$

There are two methods for estimating the lower-bound emigration rates. The first approach uses data from the 1976 SIE to create estimates of the foreign-born immigrant population by country of birth and year of entry group. These estimates are weighted by the probability of selection into the sample. The population estimates for the groups reporting entry to the United States in 1970 and 1971 are averaged to account for fiscal year reporting in the INS data and calendar year reporting in the SIE. This was done because the INS data in 1971 was reported for the fiscal year, which at that time was from July 1, 1970 to June 30, 1971, while the SIE data covered the January 1, to December 31, 1971 time period. The averaged cohort is then compared to the 1971 cohort. The lower-bound emigration rate, e_L, is estimated as Equation (13),

$$e_L = \frac{I-Q}{I} = 1 - \frac{Q}{I}, \text{ where } Q < I \qquad (13)$$

where *I* is the size of the original 1971 immigrant cohort and *Q* is the estimate of the 1971 immigrant cohort derived from the SIE using data on year of entry.

The second approach for developing lower-bound estimates of the emigration rates is to vary the response rates for the Alien Address file. Three response rates—50, 75, and 90 percent—were used to estimate the lower-bound estimate of emigration. An advantage of this method is that lower-bound emigration estimates can be produced with more precision for groups with higher rates of naturalization. The rates published in the Jasso and Rosenzweig (1982) article are cumulative rates for the time period 1971-1979. To develop annual estimates of emigration, these rates may be annualized and then applied to the population at risk of emigrating.

Strengths of the methodology

There are several strengths to Jasso and Rosenzweig's (1982) methodology for measuring emigration. The method uses administrative data to estimate emigration which can be more cost effective and less resource intensive than other methods such as carrying out special migration surveys. Also, this method follows a cohort of immigrants over a specific time period, eliminating many of the problems associated with measuring change using cross-sectional data (Borjas 1985). Emigration rates can be estimated by demographic characteristics including place-

of-birth groups. Finally, by developing both upper- and lower-bounded estimates of emigration rates, this methodology produces a range of estimates.

Limitations of the methodology

The approach developed by Jasso and Rosenzweig (1982) to estimate emigration has limitations that may prevent other researchers from using the method. The data required for this method are specific and may not be available in many countries. When using the method of bounds, which the authors use to develop a range of estimates, there is the problem that the bounds might represent an over- or underestimate of the population. Therefore, a more data-driven method for estimating the range of estimates may be desirable.

CPS Matching Method

Van Hook et al. (2006) developed a method to estimate emigration of the foreign-born population in the United States using the attrition in matched Current Population Survey (CPS) files. The CPS is a monthly survey of about 60,000 households in the United States. The CPS has a quasi-longitudinal design in which the same household is included in the survey for four consecutive months and then rotates out of the survey for eight months; they are then brought back into the sample for the same four months the following year (U.S. Department of Labor 2012). The sampling frame for the CPS is made up of addresses, not individuals, therefore, if respondents move to a new address they drop out of the CPS sample.

The CPS matching method focuses on the probability that a foreign-born respondent will not be followed-up in subsequent waves of the survey to estimate emigration (u^f). This can be represented in Equation (14) for which most of the terms can be estimated directly using data from the CPS.

$$u^f = m^f + d^f + e^f + r^f \quad (14)$$

The proportion of internal migrants *(m^f)* is estimated using CPS data on the place of residence one year ago. The probability of death *(d^f)* for the foreign born is estimated using data from the National Health Interview Survey (NHIS). The final two terms are the probability of emigrating

(e^f) and the residual non-follow-up probability *(r^f)*. These terms are not estimated directly, but by making some assumptions for r^f, the authors are able to solve for e^f.

The first assumption that allows the authors to solve for e^f is that the residual non-follow-up probability of foreign-born adults *(r^f)* is the same as that of second-generation adults *(r^s)*. Second-generation adults are the U.S. born children of immigrants *(s)*.

$$r^f = r^s = u^s - m^s - d^s - e^s \quad (15)$$

The second assumption is that emigration of second-generation adults *(e^s)* is negligible or zero, therefore the term drops out leaving an expression that can be estimated using existing data. The probability of foreign-born emigration is then:

$$e^f = u^f - m^f - d^f - (u^s - m^s - d^s) \quad (16)$$

where the predicted probability of non-response for the foreign-born and second-generation populations *(u^f* and *u^s)* were estimated with logistic regression on the CPS data, the probability of internal migration of the foreign-born and second-generations *(m^f* and m^s*)* are estimated using CPS data on the residence of respondents one year ago, and the probability of dying *(d^f* and d^s*)* was estimated using data from the National Health Interview Survey-National Death Index (NHIS-NDI).[7] The emigration rates can then be applied to stock estimates of the foreign-born population to get estimates of the level of emigration.

Strengths of the methodology

The Van Hook et al. (2006) method for estimating emigration of the foreign-born population has several strengths. The CPS matching method uses data from a household survey to estimate emigration, making the method possible to replicate in other countries that have longitudinal household or labor force surveys. The method provides estimates of foreign-born emigration that are more current than estimates from residual methods, which often have an estimate time span of five to ten years. The method is able to capture recent emigration and temporary migration,

[7] The NHIS-NDI data link death records with data from the National Health Interview Survey, which is the primary source of health information for the United States (Centers for Disease Control and Prevention 2012).

which may occur in the interim period between censuses. For this reason, the results using the CPS matching method are higher than estimates calculated using a census-to-census residual method. The CPS matching method can be used to estimate emigration by very specific demographic and social characteristics—education level, health status, language ability—which could potentially inform immigration policy makers.

Limitations of the methodology

A limitation of the CPS matching method is that it relies on several key assumptions. First, the assumption that the residual probabilities of non-response for the foreign-born and second-generation adult populations are equal could be overestimating emigration rates, especially if there are compositional differences caused by immigrant selectivity between these two populations. Second, this method assumes that the emigration of second-generation adults is negligible or zero. While the level of native emigration in the United States is relatively low (Gibbs et al. 2003, Schachter 2008), assuming it to be zero would bias the results of the CPS matching method by overestimating emigration. Furthermore, emigration of second-generation adults is probably higher than emigration of other natives because second-generation adults often have the language ability, cultural capital, and transnational social networks needed to facilitate emigration.

Indirect Estimation and Multiplicity Sampling Methods

Indirect estimation and multiplicity methods use data on the residence of relatives living abroad from household surveys to estimate emigration. Indirect estimation methods for measuring emigration were largely adapted from the literature on indirect techniques for estimating mortality in developing countries using household surveys (Zaba 1985). Special survey questions about the residence of household members or relatives are used to identify the population of interest. There are two main approaches to the indirect estimation methods that have been used in the literature—the residence of children and the residence of siblings—which are discussed in more detail below. Multiplicity sampling methods use similar survey questions to estimate emigration as indirect estimation methods, but also adjust for multiplicity, the probability that

more than one household may identify the same emigrant, between the sample and population (Woodrow-Lafield 1996).

Indirect Estimation

Indirect estimation methods estimate emigration using responses to questions on household surveys about the residence of household members or relatives (Bonaguidi 1990; Hill 1979; Somoza 1980; Somoza 1981; Zaba 1985; Zlotnik 1987). Indirect estimation methods using household surveys were originally developed to measure mortality in developing countries where vital record systems were incomplete (Zaba 1987). There are two main approaches to the indirect estimation method: the residence of children and the residence of siblings. The residence of children approach estimates emigration using responses to survey questions asked to mothers about the current residence of all children ever born to them and makes adjustments for orphaned children or children whose mother has also emigrated (Somoza 1980; Somoza 1981; Zaba 1985; Zaba 1987). The residence of siblings approach uses responses to survey questions asked of all respondents about the current residence of their siblings and adjusts for multiple reporting or when the emigrant's siblings have also emigrated (Hill 1979; Hill 1983).

There are two adjustments made for children that do not have a resident mother because they are orphans or their mother has also migrated. A relatively simple adjustment for orphaned children can be made as long as the survey instrument also collects data on the survival of mothers. The total number of emigrants is calculated by dividing the number of emigrants with a surviving mother by the proportion not orphaned in each age group (Zaba 1985). This method assumes that the likelihood of emigration and being an orphan is independent. Similarly, the emigration of children whose mother has also emigrated can be estimated if the survey collects information on the residence of the mother. In this case, the total number of emigrants is calculated by dividing the number of emigrants with a resident mother by the proportion of respondents in that age group with a resident mother. This method assumes that the likelihood of the child emigrating is independent of the likelihood of the mother emigrating. This assumption may be problematic, especially for young emigrants that are likely to migrate with their family, and will lead to an underestimate of emigration.

The residence of children and residence of siblings methods have also been used to estimate the demographic characteristics of emigrants. The sex of the emigrant can be measured directly if the survey instrument is structured to collect information by sex. For example, Zaba (1985) recommends that for the residence of children approach, mothers are asked the following question:

> "Of all the children born alive by this woman, how many –
>
> (i) sons are living in this country?
> (ii) daughters are living in this country?
> (iii) sons are living abroad?
> (iv) daughters are living abroad?
> (v) sons have died?
> (vi) daughters have died?"

For the residence of siblings approach, the following question is recommended:

> "Of this person's brothers and sisters (by the same mother and including the respondent), how many –
>
> (i) brothers are living in this country?
> (ii) sisters are living in this country?
> (iii) brothers are living abroad?
> (iv) sisters are living abroad?
> (v) brothers have died?
> (vi) sisters have died?"

Estimating the age distributions of emigrants is a much more complicated process. For the residence of children approach, Zaba (1985) uses the age of the mother, the survival rates for the child, and observed fertility rates for the population to model the approximate age of children that have emigrated. The age of emigrants can also be estimated using modeled age distributions of children from the population based on the mother's age (Hill 1981).

Multiplicity Sampling Method

The multiplicity method uses data from special surveys to estimate emigration. These surveys use a probability sampling technique called network sampling to identify rare populations

(Sirken 1970). In network sampling, respondents to the survey are asked to identify additional observations or elements to the sample that have a given characteristic. The elements in the original sample are referred to as selection units and the elements with the specific characteristic are referred to as observation units. The counting rule, or the condition for linking observation units to selection units, for these surveys must account for the fact that observation units can be linked to more than one selection unit (i.e., a person living abroad could be identified by more than one survey respondent).

Multiplicity counting rules establish linkages between observation units and selection units to ensure that there is full coverage of the population of interest. Individuals are linked to households that have enough information to report on their status with regards to the characteristic of interest (e.g., a rare medical condition or member of a hard-to-count population), and the distribution of the multiplicities tend to have a large mean and relatively small variance. The multiplicity counting rule informs the multiplicity adjustment which is used to adjust for multiplicity among the observation units. In general, network sampling enhances the response rates or enumeration of rare populations and/or populations that would be excluded from the survey because of their place of residence, both of which apply to emigrants (Sudman, Sirken and Cowan 1988).

Woodrow-Lafield (1996) used data from surveys with network sampling and a multiplicity adjustment to estimate emigration from the United States. The study applied the multiplicity method to data from the July 1987, June 1988, and November 1989 emigration supplements of the Current Population Survey (CPS). The CPS is a monthly nationally-representative labor force survey which samples about 60,000 households each month. These specific CPS surveys included questions on immediate relatives who had previously lived in the United States but were now living abroad. In addition, questions on the age, sex, date of departure, major activity abroad, country of birth, and citizenship of respondent's relatives living abroad were included in the survey instrument.

The number of emigrants was identified using a consanguineal counting rule whereby survey respondents were asked to indicate if any immediate relatives (children, siblings, parents, half-brothers, and half-sisters) were living abroad. The next step used the base survey weights from

the CPS to develop a multiplicity probability, which accounted for each emigrant's number of relatives and the probability of selection for those relatives. Specifically, the method sums the probability of selection for all possible relatives for each emigrant and then sums the probability of selection for the emigrant in the population (the population multiplicity). The population multiplicity adjusts for emigrants who have resident relatives that were not included in the CPS. The method, however, is not able to estimate emigrants who do not have relatives in the United States, which is especially common for foreign-born emigrants whose relatives remained in the origin country or who emigrated as whole families.

The proper specification of the multiplicity adjustment is vital to this method because it has a relatively large impact on the final estimates. Woodrow-Lafield (1996) estimated that there were 1,724,000 immediate relatives of native emigrants resident in the United States in the November 1989 CPS. After applying the multiplicity adjustment, the final estimate of native emigrants was 353,000, a reduction of nearly 80 percent. Furthermore, her study finds that the multiplicity-adjusted estimates include significantly more recent emigrants than those that emigrated during earlier time periods. This could indicate potential error in the recall of later events, that emigrants from earlier time periods have returned to the United States, or that emigrants from later time periods no longer have relatives who are still resident in the United States. All of these factors may lead to an underestimate of emigration during later time periods.

Strengths of the methodology

The indirect estimation and multiplicity sampling methods have some clear advantages over other approaches to estimating emigration. First, these methods can be used to estimate the recently emigrated population. Second, the special questions needed for these methods can be added to existing household surveys, therefore, data collection is much less expensive than migration surveys. Finally, these methods can be used to estimate emigration of both the native and foreign-born populations; however, estimating emigration of the foreign-born population can be problematic because the likelihood of a whole family move is greater than for the native population.

Limitations of the methodology

These methods also have limitations that may bias the estimates. Given that both the indirect estimation and multiplicity sampling methods rely on data from household surveys, which sample the relatives of emigrants who are resident in the United States, the estimates do not include whole family migrations or emigrants who do not have relatives in the resident population. Another limitation of the methods is that unless the period of departure is included in the survey instrument, the emigration rates effectively cover long time periods and are therefore difficult to use in estimating annual or recent emigration. Survey errors including coverage and recall bias could also bias the estimates of emigration.

Statistical Modeling

Statistical modeling can be used to produce migration estimates, especially when data are missing or incomplete (Willekens 1999). Statistical models are abstract representations of real world phenomena that can incorporate both empirical observations and theoretical constructs. Statistical models have helped researchers develop general theories about migration. For instance, Rogers and Castro (1981) used statistical models to identify patterns in migration schedules, or age-specific migration rates, which have successfully been generalized to numerous populations. Statistical models can produce estimates at the individual level, such as that probability or the likelihood that an individual will migrate, or at the aggregated level, such as the probability that a migration stream originated in a particular region or country. In this section, we provide a brief overview of statistical models and their application to estimating emigration.

Individual-level statistical models use micro data to calculate the probability that individuals in the population—given the age, covariate, and spatial structures of the population—will migrate (Willekens 2008). The age structure of the migrant population is of special importance because the most common approach to modeling migration at the individual level uses multivariate analysis of time-to-event data to model migration using age as the duration variable (Rogers and Castro 1981; Willekens 2008). The resulting multistate survival models can incorporate both

continuous time-to-event data like what would be collected from a population register, and categorical time-to-event data like what would be collected from a census or survey.

Statistical modeling is also used to estimate aggregate migration flows between regions or countries. Models of place-to-place migration start with double-entry contingency tables where the flows to and from countries are represented in a single matrix. Because these tables often have missing or incomplete data, statistical models are used to estimate the missing cells. These models include the demographic fixed-rate model, gravity model, entropy maximization model, and log-linear model (Plane 1982; Raymer 2007). The gravity and entropy maximization models are spatial-interaction models that weight countries of origin or destination based on their geographic proximity. The log-linear model can also incorporate spatial dependence as well as other covariates such as the age structures, economic dependence, and cultural similarities (common language) between origin and destination countries (Raymer and Rogers 2007). The log-linear model is especially useful when estimating migration flows where data in the origin-destination matrix are sparse or missing (Raymer 2007; Raymer and Rogers 2007).

Strengths of the methodology

Statistical models of migration flows have benefits over other methodologies. In multivariate statistical models, different dimensions of population structure can be used to develop robust indicators of migration behavior and flows. Statistical models can highlight patterns in migration that can be generalized to other populations. For instance, the model migration schedules developed by Rogers and Castro (1981) have shown that migration has a very strong age structure similar to fertility (Raymer and Rogers 2008). Statistical modeling helps to overcome issues of missing or sparse data, which is especially important given that migration data tend to have more uncertainty than other demographic data. Finally, statistical models express the probability that an event will occur, which can be helpful when assessing the validity of an estimate.

Limitations of the methodology

Statistical models of migration have limitations for estimating emigration. The individual-level statistical models require specific types of data—migration histories—that are often unavailable

for emigrants. There are also data restrictions for aggregate-level models. In order to model country-to country migration flows, the data from the various countries must be harmonized so that time periods, definitions of who is and who is not a migrant, and other data issues are consistent (Poulain et al. 2006). A limitation of the log-linear model is that the marginal totals for the country-to-country flow matrix must be either known or estimated requiring an additional step in the process of developing emigration estimates (Raymer 2007). Finally, statistical models tend to be technical to calculate and the results might be difficult to translate into a single set of estimates that will be used by national statistical agencies or policy makers.

Summary and Conclusions

Developing accurate and timely estimates of emigration is essential for producing valid population estimates. However, emigration has proven to be one of the most difficult components of population change to measure. This report has provided a review of the methods and techniques used by researchers and national statistical agencies to estimate emigration. The review has also provided an assessment of the strengths and limitations of each approach. In this section, we summarize the challenges that are commonly shared by the methods for estimating emigration focusing specifically on data and methodological issues.

Data Challenges

The availability of data on emigrants is the principal challenge to estimating emigration for most countries. Emigrants are no longer resident in the country of estimation; therefore they cannot be directly measured using censuses or household surveys. Even population registries, which typically require emigrants to de-register from the registry, are plagued with incomplete data on emigrants. Because direct data on emigrants are unavailable, emigration is often estimated indirectly using data on the change in the stock population or information collected from the relatives of emigrants. There is also variation in the availability of data used to indirectly estimate emigration. Data from population registries and censuses are available for most countries, however, these data often lack the specificity needed to produce unbiased estimates of emigration. While specialized migration surveys may contain detailed information on emigration, few countries regularly conduct these surveys.

Table 1. Typology of Emigration Methodologies by Timeliness and Availability of Data

Method	Timeliness of data	Data commonly available
Population registers	Yes	No
Migration surveys	Yes	No
Residual methods	No	Yes
Panel data attrition method	Yes	No
Indirect estimation method	Yes	No
Multiplicity sampling method	Yes	No
Statistical models	Yes	No

The timeliness of data used to estimate emigration is also a challenge for producing estimates. For instance, a country using a population registry that is accurate and up-to-date could potentially produce monthly or quarterly emigration estimates while a country that uses census data with the residual method would produce estimates that could not be updated for five to ten years depending on the frequency of the census. Indirect estimation and multiplicity sampling methods produce emigration rates that are most accurate for more recent emigrants, but unless period of departure is included in the survey estimates, the emigration rates effectively cover long time periods. The frequency with which data used to estimate emigration is collected or produced creates a significant challenge.

In producing estimates of emigration, there is often a trade-off between the availability and timeliness of data. A typology of emigration estimation methodologies by the availability and timeliness of data is presented in Table 1. Data from population registries are available for some countries and can be used to produce current estimates of emigration. Data from population censuses are also available for most countries but emigration estimates produced using population censuses are often outdated because of the long duration between censuses. Specialized migration surveys are less common than other sources of data, however, these surveys may produce current estimates of emigration. Similarly, special questions on household surveys that are used with indirect estimation and multiplicity sampling methods are uncommon but can produce estimates of recent emigrants. Panel data attrition methods may also produce current estimates, but the specific data requirements for these methods may not allow them to be replicated in other countries. While statistical modeling has been used to compensate for

incomplete and missing data, these methods also have data requirements (i.e., harmonized migration flow data) that may be unavailable for many countries.

Methodological Challenges

In addition to data challenges, there are also methodological challenges to estimating emigration. While both the native and foreign-born populations are at risk of emigrating, the substantially higher emigration rates for the foreign-born population often presents a challenge to estimating emigration. Many national statistical agencies estimate emigration of the native and foreign-born populations together which does not allow for varying rates across these populations. This produces its own challenges because data on the foreign-born population may have greater coverage or measurement error than data on the native population, which would bias the emigration estimates. Emigration of the native-born population is usually a relatively rare event making estimation difficult.

Also, direct data on emigrants is often unavailable, therefore, most methods for estimating emigration use indirect data which may create errors in the estimates. While all estimation contains some error, methods that use indirect data often require many assumptions about the data that may produce more error than other methods of estimating populations. Estimating emigration is also methodologically challenging because it is not a final process such as fertility or mortality. International migration may be temporary or even circular and these repeated migration events may not be measured very well using many emigration methodologies. The CPS matching method measures much higher emigration rates than the residual method using census data because the CPS matching method captures temporary and circular migration between censuses.

Conclusion

International migration statistics often include estimates of immigration and emigration. Emigration is one of the most difficult components of population change to estimate because of both data and methodological challenges. Despite significant development of methods for estimating emigration over the last three decades, estimating emigration remains a challenge. Furthermore, at this time, there is not a universal methodology implemented across countries.

For this report, a review of the published literature on methods for estimating emigration in several different languages was conducted. Next, a detailed description of the most common methods including population registries, migration surveys, residual methods, data attrition methods, indirect estimation, the multiplicity sampling method, and statistical modeling was provided. In addition to detailed descriptions, the strengths and limitations of each methodology were discussed. The overall goal of this project was to prepare a document that will aid researchers and national statistical agencies in not only producing estimates of emigration but also developing new methods for estimating emigration that can overcome the challenges presented above.

References

Ahmed, Bashir and J. Gregory Robinson. 1994. "Estimates of Emigration of the Foreign-Born Population: 1980-1990." *Population Division Working Paper No. 9*. U.S. Census Bureau. (English)

Bonaguidi, Alberto. 1990. "Measurement of Emigration Using Indirect Techniques." *European Journal of Population-Revue Europeenne De Demographie* 6:113-116. (French)

Borjas, George J. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *Journal of Labor Economics* 3:463-489. (English)

Cantisani, Giambattista and Valeria Greco. 2006. "Country Report: Italy." Pp. 477-489 in *THESIM: Towards Harmonised European Statistics on International Migration*, edited by M. Poulain, N. Perrin, and A. Singleton. Louvain-la-Neuve: Presses universitaires de Louvain. (English)

EMIF Norte. 2010. "Migration Survey on the Northern Border of Mexico, 2010." http://www.colef.net/emif/publicacionesnte.php. (Spanish)

Gibbs, James C. , Gregory S. Harper, Marc J. Rubin, and Hyon B. Shin. 2001. "Evaluating Components of International Migration: Native Emigrants." *Population Division Working Paper No. 63*. U.S. Census Bureau. (English)

Hill, Kenneth. 1979. "The Use of Information on Residence of Siblings to Estimate Emigration by Age." *Notas de Poblacíon* 7:71-89. (Spanish)

—. 1981. "A Note of Estimating the Age Distribution of Surviving Emigrant Children." *Notas de Poblacíon* 27.(Spanish)

—. 1983. "A Note on the Use of Information Concerning the Residence of Surviving Siblings to Estimate Emigration." *Notas de Poblacíon* 11:33-58. (Spanish)

Jasso, Guillermina and Mark R. Rosenzweig. 1982. "Estimating the Emigration Rates of Legal Immigrants Using Administrative and Survey Data: The 1971 Cohort of Immigrants to the United States." *Demography* 19:279-290. (English)

Mulder, Tammany J., Betsy Guzman, and Angela Brittingham. 2002. "Evaluating Components of International Migration: Foreign-Born Emigrants." *Population Division Working Paper No. 62*. U.S. Census Bureau. (English)

Plane, David A. 1982. "An Information Theoretic Approach to the Estimation of Migration Flows." *Journal of Regional Science* 22:441. (English)

Poulain, Michel, Nicolas Perrin, Ann Singleton, and THESIM Project. 2006. *THESIM: Towards Harmonised European Statistics on International Migration*. Louvain-la-Neuve: Presses universitaires de Louvain. (English)

Raymer, James. 2007. "The Estimation of International Migration Flows: A General Technique Focused on the Origin-Destination Association Structure." *Environment and Planning A* 39:985-995. (English)

Raymer, James and Andrei Rogers. 2007. "Using Age and Spatial Flow Structures in the Indirect Estimation of Migration Streams." *Demography* 44:199-223. (English)

—. 2008. "Applying Model Migration Schedules to Represent Age-Specific Migration Flows." Pp. 173-192 in *International Migration in Europe: Data, Models and Estimates*, edited by J. Raymer, F. Willekens, and the University of Southampton. Division of Social Statistics. Hoboken, N.J.: Wiley. (English)

Rendall, Michael, Cecilia Tomassini, and David J Elliot. 2003. "Estimation of Annual International Migration from the Labor Force Surveys of the United Kingdom and the Continental European Union." *Statistical Journal of the United Nations ECE* 20:219-234. (English)

Rogers, Andrei. and Luis J. Castro. 1981. "Model Migration Schedules." Report RR-81-30. Laxenburg, Austria: International Institute for Applied Systems Analysis. (English)

Salvisberg, Esther and Marcel Heininger. 2007. "Im- and Emigration of the Permanent Foreign Population: From Register Counts to Statistics." Federal Statistical Office of Switzerland. (German)

Schachter, Jason. 2008. "Estimating Native Emigration from the United States." Memorandum dated December 24, delivered to the U.S. Census Bureau. (English)

Silvestrini, Angela and Giovanni Cariani. 2005. "Survey on the Italians Abroad on 21st March 2003: Demographic Characteristics." General Bureau of Italians Abroad and Policy Migration and Italian National Institute for Statistics. (Italian)

Sirken, Monroe G. 1970. "Household Surveys with Multiplicity." *Journal of the American Statistical Association* 65:257-266. (English)

Somoza, Jorge L. 1980. "Indirect Estimates of Emigration: Two Methods Based on Information on Place of Residence of Children and Siblings." *Notas de Poblacíon* 8:93-122. (Spanish)

—. 1981. "Indirect Estimates of Emigration: Applications of Two Procedures Using Information on Residence of Children and Siblings." *International Union for the Scientific Study of Population (IUSSP) Paper* 5:35-60. (English)

Sudman, Seymour, Monroe G. Sirken, and Charles D. Cowan. 1988. "Sampling Rare and Elusive Populations." *Science* 240:991-996. (English)

U.S. Department of Labor. 2012. "Labor Force Statistics from the Current Population Survey: Technical Documentation." September 9, 2012. http://www.bls.gov/cps/documentation.htm. (English)

United Kingdom Office for National Statistics. 2012. "International Passenger Survey (IPS) Methodology." September 8, 2012. http://www.ons.gov.uk/ons/guide-method/method-quality/specific/travel-and-transport-methodology/international-passenger-survey/index.html. (English)

Van Hook, Jennifer, Weiwei Zhang, Frank D. Bean, and Jeffrey S. Passel. 2006. "Foreign-Born Emigration: A New Approach and Estimates Based on Matched CPS Files." *Demography* 43:361-382. (English)

Warren, Robert and Jennifer Marks Peck. 1980. "Foreign-Born Emigration from the United States: 1960 to 1970." *Demography* 17:71-84. (English)

Willekens, Frans. 1999. "Modeling Approaches to the Indirect Estimation of Migration Flows: From Entropy to EM." *Mathematical Population Studies* 7:239-78. (English)

—. 2008. "Models of Migration: Observations and Judgements." Pp. 117-147 in *International Migration in Europe : Data, Models and Estimates*, edited by J. Raymer, F. Willekens, and University of Southampton. Division of Social Statistics. Hoboken, N.J.: Wiley. (English)

Woodrow-Lafield, Karen. 1996. "Emigration from the USA: Multiplicity Survey Evidence." *Population Research and Policy Review* 15:171-199. (English)

Zaba, Basia. 1985. *Measurement of Emigration using Indirect Techniques: Manual for the Collection and Analysis of Data on Residence of Relatives*. Liege, Belgium: Ordina Editions. (French / English)

—. 1987. "The Indirect Estimation of Migration: A Critical Review." *International Migration Review* 21:1395-445. (English)

Zlotnik, Hania. 1987. "The Use of Information on Residence of Relatives to Measure International Migration." *Notas de Poblacíon* 15:25-65. (Spanish)