Adapting the Census residual method for state-level estimates of emigration

Crystal Yu and Ethan Sharygin

Demographic Research Unit

California Department of Finance

Extended Abstract for PAA

ABSTRACT

To account for components of population change, demographers need to numerically track movements across international borders. There are a variety of methods to count immigration, but emigration remains a challenging area of research. In this paper, we replicate previously published results for the U.S. using publicly available data, and extend the method to estimate emigrants from any U.S. state. We test the sensitivity of the results to different life tables and base populations. We find that duration spent in the U.S. varies across states in its significance as a predictor of emigration. In addition, emigration rates vary according to age and country of origin. Using survival rates for the foreign born population lead to higher modeled rates of emigration. Our findings show that net international migration can be reliably decomposed into inbound and outbound migration, which has useful applications for research and policy.

INTRODUCTION

The California Department of Finance is responsible for producing estimates for the total population for the state of California and its counties. We use the fundamental balancing identity of demography (ending population = starting population + births - deaths + net migration) to estimate population changes since the most recent U.S. census. Of these, net migration is the most difficult to estimate, owing to the lack of a comprehensive system for tracking border crossings. While administrative and survey data allow us to reliably estimate the number of foreign-born immigrants coming to the U.S., there is no direct method to count the corresponding flows outward.

We seek to meet this need by introducing a modified version of the method used by the U.S. Census Bureau to estimate emigration flows of the foreign-born population out of the United States. In this analysis, we define the foreign-born to be anyone who is a naturalized U.S. citizen or non-citizen of the United States. Furthermore, we will count all individuals born in American Samoa as U.S. citizens, even if their actual citizenship status suggests they are foreign-born individuals. We focus our analysis on the foreign-born population as they are the most likely to permanently emigrate from the U.S. back to their home country (Leach and Jensen, 2013). In comparison, U.S. natives are less likely to leave the U.S. for another country (Jensen, 2013); previous research estimates that about 45,000 U.S. natives emigrate from the U.S. annually (Dick et al., 2012).

We test the sensitivity of this method to the arrival cohorts used to estimate rates, and to the choice of mortality schedule applied in the cohort projection. In addition, we are interested in modifying the national level model so it can be used at the U.S. state level. At the state level, the cohorts are subject to increment by in-migration of the foreign born from other U.S. states as well as from foreign countries.

There is also decrement through out-migration of the foreign born from the state being examined – California, in this analysis – to other U.S. states. To this end, we use the American Community Survey (ACS) to count the number of individuals in a specific foreign-born cohort that arrived in the U.S. by a given year. Differentiating foreign-born individuals by entry cohort allows us to not only determine emigration rates and total emigrants for a U.S. state, but also examine how year of arrival, years of residence in the U.S., age, sex, and place of birth each affect emigration rates.

BACKGROUND / RELATED LITERATURE

Compared to the literature on immigration, research on methods for estimating emigration is relatively sparse. Grieco and Rytina (2011) and Jensen (2013) summarize the main data sources and current methods, and provide a useful overview of the issues. One group of methods uses attrition from longitudinal surveys. For example, Van Hook et al. (2006) and Van Hook and Zhang (2011) exploit design features of the Current Population Survey (CPS) to estimate emigration based on attrition between survey waves. However, these calculations result in very high estimates of emigration, since they capture a great number of temporary and circular migrants. Immigration data usually track not *all* entries from another country but only the initial arrival of people who intend to make the U.S. their long-term residence. For this reason, we prefer the residual method, because it defines emigration in a way that better complements immigration as a permanent or long-term move. Another advantage of the residual method is the ready availability of the necessary data.

The residual method, first proposed by Warren and Peck (1980) and improved upon over the years by Ahmed and Robinson (1994), Mulder et al. (2002), and the U.S. Census Bureau Population Estimates Program research group (U.S. Census Bureau; Scopilliti et al., 2011; Leach and Jensen, 2013; Bhaskar et al., 2013a; Bhaskar et al. 2013b), measures the survival of a specially defined cohort of foreign-born immigrants who arrived to the U.S. by a certain year at two points in time. Historically, researchers have used consecutive censuses to estimate emigration over a decade (Ahmed and Robinson, 1994; Mulder et al., 2002; Leach and Jensen, 2013), although with the full implementation of the ACS in 2005, it is now possible to estimate emigration of a defined population of interest over a shorter time period and with greater timeliness (Grieco and Rytina, 2011; Leach and Jensen, 2013). The only modes of decrement from the closed population being examined between time points 1 and 2 are death or emigration. After accounting for the expected number of deaths using life tables, the discrepancy between the anticipated and observed population sizes in a later enumeration of the population is attributed to emigration.

DATA AND METHODS

The residual method of estimating emigration rates relies upon two or more years of age-structured population census data and a corresponding age-specific mortality schedule for that population. To replicate the results presented by Bhaskar et al. (2013a, 2013b), we use the 2000 U.S. total population life table (Arias et al., 2010) and the public use microdata samples (PUMS) files of Census 2000 and the one-year 2001-2010 American Community Surveys. We tabulate two age distributions of the foreign-born population (excluding those residing in group quarters): one for the cohort who entered the U.S. more than 10 years ago (before 1990), and another for the cohort who entered the U.S. within the subsequent 10 years (during 1990-1999).

Using the 2000 U.S. total population life table published by the National Center for Health Statistics (NCHS) (Arias et al., 2010), specifically, the age-specific probabilities of dying, we calculate the

expected population size of our foreign-born cohorts each year between 2001 and 2010[1]. Then, we compare their resulting age distributions to the observed age distributions from the corresponding ACS, implicitly assuming that the probability of dying is the same for the foreign-born population as it is for the U.S. population overall. We also assume that the probability of dying did not change during the decade. These assumptions may be problematic, as previous research suggests that the foreign-born population may be positively selected on the basis of health and thus have lower mortality rates than the native born population (Palloni and Arias, 2004; Goldman et al., 2006; Dupre et al., 2012). Moreover, mortality rates are known to have changed during the decade (Miniño and Murphy, 2012). We test the sensitivity of the results under alternative assumptions, for example, using life tables calculated exclusively for the foreign-born population, and using annual life tables generated by interpolating between decennial census years of 2000 and 2010.

To calculate state-level emigration rates for the same time period, we use California as an example. We use the same population data sources, but in lieu of the U.S. total population life table, we use the NCHS California state life table (National Center for Health Statistics, 2012). Moreover, for each of the two foreign-born cohorts, we tabulate the age distribution of domestic in- and out-migrants to California using the ACS. Because the ACS does not specify the timing of migration relative to age, we assume that moves occur at a uniform rate throughout the year, and that half of the population moved while one year younger than the age they reported in the ACS. In other words, we use data from two consecutive one-year ACS files to estimate the total number of foreign-born individuals moving to California from another U.S. state and the number of foreign born individuals moving from California to another U.S. state for any given calendar year. Incorporating these numbers allows us to track the size of our two foreign-born cohorts from 2000 to 2010.

RESULTS AND DISCUSSION

Using the observed and expected population size estimates, we can compute the residual population size between 2001 and 2010. After annualizing the residuals, we compute an estimated emigration rate by calculating the proportion of foreign-born individuals presumed to have left the U.S. out of the total cohort of foreign born present in 2000. Following Bhaskar et al. (2013a, 2013b), we disregard rates calculated from the 2001-2004 ACS due to small sample sizes. For the remaining years, we calculate each annual rate as an average of the rates from the same year and the two preceding years. That is, the emigration rate calculated for 2007 is based on an average of the rates for 2005, 2006, and 2007, while the 2008 rate is an average of the rates from 2006, 2007, and 2008, and so on. We then average these three average rates to obtain an annual emigration rate per 1000 population (Table 1).

Table 1: Estimated annual emigration rate per 1000 population from U.S., 2001 to 2010[2]

---

[1] A small proportion of the foreign born population age 65 and over have inconsistent values between their stated age and the year which they entered the United States (for example, records in the Census 2000 PUMS which report an age of 70, but 1925 for the year of entry to the U.S.). We drop 1268 cases with inconsistent reporting, which may be due to respondent error, data entry errors, or the results of imputation or nondisclosure procedures (Alexander et al., 2010).

[2] This slight difference in time period necessarily arises from how and when the ACS is administered. The ACS is an ongoing survey that is answered by respondents throughout the calendar year; some may respond in January

	Estimated average emigration rate of U.S. foreign-born population
Entered <= 10 years ago (arrived 1990-1999)	13.7
Entered > 10 years ago (arrived before 1990)	2.5

Source: Author's calculations. These rates are derived using Census 2000 PUMS and ACS 2001 to 2010 PUMS data, as well as the NCHS 2000 U.S. total population life table (Arias et al., 2010).

Our national level emigration estimates agree closely with the published rates used by the Census Bureau during this period. They found that for the April 1, 2000 to March 31, 2010 period, approximately 14 per 1000 foreign-born persons that had arrived in the U.S. within the past 10 years emigrate per year, while 3 per 1000 foreign-born individuals that entered the U.S. more than 10 years ago leave annually (Bhaskar et al., 2013a; Bhaskar et al., 2013b). We find similar results in table 1. Applying these rates to the population at risk of emigration in the U.S. – namely, the foreign-born that have resided in the U.S. for at least one year – we calculate that approximately 2.3 million foreign-born persons emigrated out of the U.S. during 2001-2010.

Next, we present our estimated emigration rates for California, derived from different mortality schedules. The first set of rates uses the California statewide life table published by NCHS, while the next four sets use custom life tables using data on number of deaths collected by the California Department of Public Health (CDPH). We calculate four different mortality schedules: the first for the total population in California in 2000, the second for the foreign-born population residing in California in 2000, the third for the total population in California in 2010, and the fourth for the foreign-born population in California in 2010. We then subject the foreign-born entry cohorts that arrived to the U.S. by 1990 and between 1990 and 1999 to these various rates to calculate their rate of emigration from California (see table 2).

Table 2: Estimated average CA emigration rates per 1000 population, using different probabilities of dying based on different populations (shown in table)

	Mortality schedule derived from data collected by				
	National Center for Health Statistics (NCHS)	California Department of Public Health (CDPH)			
	2000	2000		2010	
	Total CA population (Baseline calculation)	Total CA population	CA foreign-born population only	Total CA population	CA foreign-born population only
Entered U.S. <= 10 years ago (arrived 1990-	16.3	16.3	16.4	16.8	17.3

while others respond in December. This means that the collected data spans one calendar year. In using ACS data to estimate the population at risk of emigration, we then capture the at-risk population of a given calendar year.

1999)					
Entered U.S. > 10 years ago (arrived before 1990)	8.0	7.8	8.1	9.0	10.3

Source: Population from Census 2000 PUMS and ACS 2010 1-year PUMS; baseline mortality ($_nq_x$) from National Center for Health Statistics CA Life Tables 1999-2001); deaths from California Department of Public Health used to calculated California total and foreign-born only mortality rates.

The results in table 2 suggest that emigration patterns from California are quite different from those in the U.S. overall. While in all cases the rate of emigration is lower for those who have been resident in the U.S. for longer than ten years, the rate of emigration among those who arrived prior to 1990 is much higher than the national rate. This could be due to a number of factors: the age structure of the foreign born population, the mixture of countries of origin, state and national policies, citizenship status, and differences in the integration and kinship networks of immigrants.

Moreover, table 3 demonstrates that the use of disparate mortality schedules yields different estimates for emigration rates from California. Life tables calculated for the foreign born population only tend to reflect lower mortality, which result in a larger residual and consequently, a higher rate of emigration (due to fewer projected deaths for the entry cohort). This affirms that the choice of mortality schedule does, indeed, make a difference in the estimated rate of emigration for the foreign-born population. Additionally, this suggests that the foreign-born population may be emigrating at a higher rate than previously suggested.

Applying the baseline emigration rates for California to the population at risk (namely, foreign-born individuals that have resided in the U.S. for at least one year) in California from 2001 to 2010 results in an estimated 980,000 emigrants from California during 2001-2010. We intend to repeat this calculation using additional mortality schedules and different entry cohorts (for example, the residuals during 2005-2014 for cohorts that entered the U.S. during 1995-2004).

When we limit our analysis to foreign-born individuals born in Mexico, we find that the Mexican-born population that have been in the U.S. for fewer than 10 years emigrate from the U.S. at a rate similar to other foreign-born emigrants with the same years of residency (Table 3). However, the Mexican-born population that have resided in the U.S. for more than 10 years are more likely than other foreign-born individuals born elsewhere in the world and have also lived in the U.S. over 10 years to leave the U.S.

Table 3: Estimated annual emigration rate per 1000 population from U.S. among the Mexican-born population, 2001-2010

	Estimated average emigration rate among U.S. Mexican-born population
Entered <= 10 years ago (arrived 1990-1999)	13.0
Entered > 10 years ago (arrived before 1990)	6.4

Source: Author's calculations. These rates are derived using Census 2000 and ACS 2001 to 2010 data, as well as the 2000 U.S. total population life table published by NCHS (Arias et al., 2010).

The rates presented in table 3 agree closely with the ones calculated by Scopilliti et al. (2011), who examined migration patterns between the U.S. and Mexico for the time period April 1, 2000 to March 31, 2010. They found that Mexicans that had arrived between 1990 and 1999 emigrate from the U.S. at a rate of approximately 13 to 14 per 1000 individuals per year, while those that entered the U.S. before 1990 left at an annual rate of 7 per 1000.

Applying the rates in table 2 to the population at risk of emigration – here, the foreign born population born in Mexico that have resided in the U.S. for more than one year, we find that approximately 950,000 individuals born in Mexico emigrated from the U.S. between 2001 to 2010. Note that we are unable to determine whether these Mexican-born emigrants returned directly to Mexico, or if they relocated to another country after leaving the U.S. Regardless, our estimates show that almost 2 out of every 5 foreign-born emigrant from the U.S. were born in Mexico, a considerable number. We repeat this calculation for California (see table 4), where a large number of Mexican-born immigrants reside, and intend to examine rates for immigrants from other countries in Asia and Latin America.

When we examine the emigration patterns of Mexican-born individuals residing in California, we observe the results in table 4. We find that Mexican-born individuals living in California that have been resident in the U.S. for fewer than 10 years leave at a rate comparable to that of foreign-born individuals born elsewhere in the world. However, those that have been in the U.S. for longer than 10 years are leaving at a much higher rate from California than from the U.S. as a whole. We postulate that this divergence in emigration rates for the Mexican-born population from California relative to the national level findings may arise from the age structure of the Mexican-born population in California relative to the U.S., the citizenship status and kinship networks of the Mexican-born in California, differences in policy relating to undocumented immigrants, or some combination of factors. The median duration of residence in the U.S. for the Mexican-born population is high for California, so the consequences of using different rates of emigration for long-term residents are particularly large in estimating or projecting the population.

Table 4: Estimated annual emigration rate per 1000 population from California among the Mexican-born population, 2001-2010

	Estimated average emigration rate among U.S. Mexican-born population
Entered <= 10 years ago (arrived 1990-1999)	12.2
Entered > 10 years ago (arrived before 1990)	9.3

Source: Author's calculations. These rates are derived using Census 2000 and ACS 2001 to 2010 data, as well as the 1999-2001 CA total population life table published by NCHS (National Center for Health Statistics, 2012).

Additionally, as our emigration rates for California are based primarily on the observed rates from 2005 onwards, it is likely that changing patterns in Mexico-U.S. migration due to policy and economic conditions are reflected in the rates presented here. Past research has found that there has been a reversal in the flow of U.S.-Mexico migration during the 2000s, with more Mexican-born returning to Mexico from the U.S. as well as fewer Mexicans migrating to the U.S. (Van Hook and Zhang, 2011).

When we apply these emigration rates to the Mexican-born population in California at risk of emigration, we estimate that approximately 420,000 left California for another country between 2001 and 2010. We

are unable to determine whether these Mexican-born emigrants returned directly to Mexico, or if they relocated to another country after leaving the U.S. Regardless, we find that approximately 2 out of every 5 Mexican-born emigrants that left the U.S. did so from California.

Our next steps are to develop and test hypotheses by expanding our analysis to additional states and time periods. We intend to disaggregate the foreign born cohorts we have constructed according to citizenship status, country or world region of birth, and to analyze the age and sex profile of our calculated emigration rates. In order to assess the contribution of period and cohort effects, we also intend to test a variety of definitions of foreign born cohorts according to dates of arrival in the U.S.

Better estimates of the rate of emigration may enable more accurate post-censal population estimates and projections, which have significant fiscal impact in states like California where budgets are explicitly related to population growth.

REFERENCES

Ahmed, B. and J.G. Robinson (1994). Estimates of Emigration of the Foreign-born Population: 1980-1990. U.S. Census Bureau Population Division Working Paper No. 9. Washington D.C.: U.S. Census Bureau. Retrieved from http://www.census.gov/population/www/documentation/twps0009/twps0009.html

Alexander, J.T., M. Davern and B. Stevenson (2010). Inaccurate Age and Sex Data in the Census PUMS Files: Evidence and Implications. National Bureau of Economic Research Working Paper 15703. Retrieved from http://www.nber.org/papers/w15703

Arias, E., B.L. Rostron and B. Tejada-Vera (2010). United States Life Tables, 2005; Volume 58, Number 10. Hyattsville, MD: National Center for Health Statistics. Retrieved from http://www.cdc.gov/nchs/data/nvsr/nvsr58/nvsr58_10.pdf

Bhaskar R., R.Cortés, M. Scopilliti, E. Jensen, C. Dick, D. Armstrong and B. Arenas-Germosén (2013a). Estimating Net International Migration for 2010 Demographic Analysis: An Overview of Methods and Results. U.S. Census Bureau Population Division Working Paper No. 97. Washington, D.C.: U.S. Census Bureau. Retrieved from http://www.census.gov/people/publications/popworkingpapers.html

Bhaskar R., B. Arenas-Germosén and C. Dick (2013b). Demographic Analysis 2010: Sensitivity Analysis of the Foreign-Born Migration Component. U.S. Census Bureau Population Division Working Paper No. 98. Washington D.C.: U.S. Census Bureau. Retrieved from http://www.census.gov/people/publications/popworkingpapers.html

Dick, C., E.B. Jensen and D.M. Armstrong (2012). Estimates of International Migration for United States Natives. Paper presented at the 2012 Federal Committee on Statistical Methodology Research Conference.

Dupre, M.E., D. Gu and J.W. Vaupel (2012). Survival Differences among Native-Born and Foreign-Born Older Adults in the United States. PLoS ONE 7(5): e37177. http://doi.org/10.1371/journal.pone.0037177

Goldman, N., R.T. Kimbro, C.M. Turra, and A.R. Pebley (2006). Socioeconomic Gradients in Health for White and Mexican-origin Populations. *American Journal of Public Health* 96:2186-2193.

Grieco, E.M. and N.F. Rytina (2011). U.S. Data Sources on the Foreign Born and Immigration. *International Migration Review* 45.4: 1001-1016.

Jensen, E.B. (2013). A Review of Methods for Estimating Emigration. U.S. Census Bureau Population Division Working Paper No. 101. Retrieved from http://www.census.gov/people/publications/popworkingpapers.html

Leach, M.A. and E.B. Jensen (2013). Estimating Foreign-Born Emigration from the United States Using Data from the American Community Survey. Paper presented at the 2013 Federal Committee on Statistical Methodology Research Conference.

Miniño, A.M. and S.L. Murphy (2012). Death in the United States, 2010. National Center for Health Statistics Data Brief No. 99. Hyattsville, MD: National Center for Health Statistics.

Mulder, T.J., B. Guzmán and A. Brittingham (2002). Evaluating Components of International Migration: Foreign-Born Emigrants. U.S. Census Bureau Population Division Working Paper No. 62. Washington, D.C.: U.S. Census Bureau. Retrieved from http://www.census.gov/population/www/documentation/twps0062/twps0062.html

National Center for Health Statistics (2012). National Vital Statistics Report (California); Volume 60, Number 9. Hyattsville, MD. Retrieved from http://www.cdc.gov/nchs/data/dvs/lewk4_california.pdf

Palloni, A. and E. Arias (2004). Paradox Lost: Explaining the Hispanic Adult Mortality Advantage. *Demography* 41:385-415.

Ruggles, S., K. Genadek, R. Goeken, J. Grover and M. Sobek. Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

Scopilliti, M., R. Bhaskar, E.B. Jensen and V.A. Velkoff (2011). Using Census and Survey Data to Estimate Migration Between the United States and Mexico. Paper presented at 2011 Population Association of America Annual Meeting, Estimates of Mexico-U.S. Migration from the Perspective of Both Countries: 2010 Censuses and Other Sources.

U.S. Census Bureau (2015). Population Estimates Methodology. Retrieved from http://www.census.gov/popest/methodology/index.html

Van Hook, J., W. Zhang, F.D. Bean and J.S. Passel (2006). Foreign-Born Emigration: A New Approach and Estimates Based on Matched CPS Files. *Demography* 43.2: 361-382.

Van Hook, J. and W. Zhang (2011). Who Stays? Who Goes? Selective Emigration Among the Foreign-Born. *Population Research and Policy Review* 30: 1-24.

Warren, R. and J.M. Peck (1980). Foreign-Born Emigration from the United States: 1960 to 1970. *Demography* 17.1: 71-84.