United States Census Bureau Methods for Estimating Emigration of the Foreign Born

Anthony Knapp
U.S. Census Bureau
UNECE and Eurostat Joint Work Session on Migration Statistics
Geneva, Switzerland
29-31 October 2019

This presentation is released to inform interested parties of research and to encourage discussion. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

Background

- **Emigration Statistics in the U.S.**
  - Emigration statistics published 1908-1957
  - U.S. Customs and Border Protection, Arrival/Departure Form I-94
  - Social Security Administration (various administrative data)

- **Indirect Estimation, 1980-present**
  - Decennial census data on the foreign-born stock
  - Currently using a residual method on annual American Community Survey (ACS) estimates of the foreign-born stock
Residual Method

• A residual method measures cohort change in the foreign-born stock between two years (designated as time 1 and time 2) decomposed by deaths and migration

• The residual is the remaining change, after accounting for deaths, and is assumed to represent net migration flow

• The interval between time 1 and time 2 depends on data availability (e.g., a 5- or 10-year period between censuses)

• The method only considers populations present at time 1 and time 2 and may fail to measure persons who arrived and departed within the time interval

• Census Bureau uses a modified residual method for emigration flow

\[
\begin{align*}
\tilde{M}_{1-2} &= P_2 - (P_1 - D_{1-2}) & [1] \\
\tilde{M}_{1-2} &\text{ residual implied net international migration} \\
P_1 &\text{ foreign-born population measured at time 1} \\
P_2 &\text{ foreign-born population measured at time 2} \\
D_{1-2} &\text{ deaths between time 1 and 2} \\
\tilde{E}_{1-2} &= (P_1 - D_{1-2}) - (P_2 - I_{1-2}) = P_{exp,2} - P_{obs,2} & [2] \\
\tilde{E}_{1-2} &\text{ residual implied emigration} \\
I_{1-2} &\text{ immigration between time 1 and time 2} \\
P_{exp,2} &\text{ expected population at time 2 assuming no migration} \\
P_{obs,2} &\text{ observed population at time 2}
\end{align*}
\]
ACS Questions

Place of Birth

- Used for estimating $P_1$ and $P_2$
- Defines emigrant groups by place of birth

Citizenship

- Limits the sample to the foreign born only (persons who are not U.S. citizens at birth)

Year of Entry

- Used for estimating $I_{1-2}$
- Defines emigrant groups by duration of residence

A Sensitive Method

- Fundamental requirements of the residual method:
  - Populations measured at time 1 and 2 belong to the same universe
  - Coverage in the population is consistent at time 1 and 2
  - Small changes in coverage/estimation of the population either at time 1 or time 2 can result in large changes in emigration estimates
Data Quality and the Residual

(a) Unbiased Residual

- A hypothetical population that experiences high emigration
- In this example, the components for the “observed” and “expected” populations are accurate
- The residual ($\hat{E}_{1-2}$) is unbiased

(b) Underestimated Expected Population

- The same population from example (a), but the “expected” population ($P'_{exp,2}$) is underestimated
- May be caused by under-coverage of the population measured at time 1 ($P_1$) or overstated deaths ($D_{1-2}$)
- The residual ($\hat{E}_{1-2}$) is underestimated
Data Quality and the Residual

(c) Zero Residual

- A hypothetical population that exhibits low migration
- In this example, the residual is zero or near zero
- Outcome may be plausible for relatively closed populations
- However, cases such as this are prone to negative residual estimates (next slide)

(d) Overestimated Observed Population

- The same population from example (c), but the “observed” population is overestimated \( P'_{\text{obs},2} \)
- May be due to over-coverage of the population measured at time 2 \( (P_2) \) or underestimation of immigration \( (I_{1.2}) \)
- This yields negative emigration, a demographic impossibility
ACS-ACS Residual Method

• The current method for the population estimates and projections programs at the Census Bureau

• Uses the annual American Community Survey (ACS) instead of decennial censuses to measure foreign-born populations at time 1 and time 2

• The interval between time 1 and time 2 is reduced from 10 years to include very recent emigration patterns, especially from temporary and circular migrants

• Uses ACS year of entry estimate to subtract immigrants ($I_{1-2}$) who entered the population after time 1

• Applies survival rates to the population at time 1 to account for deaths ($D_{1-2}$) and to calculate the “expected” population ($P_{exp,2}$)

Residual Estimates by Foreign-Born Groups

• Characteristics of the foreign born are heterogeneous

• Residuals are estimated separately for certain groups to better reflect different emigration propensities

• Currently use (7) emigrant groups defined by place of birth, duration of residence, and sex

• Groups with similar characteristics were combined

<table>
<thead>
<tr>
<th>Place of Birth</th>
<th># of Years in U.S.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recent</td>
<td>Non Recent</td>
</tr>
<tr>
<td>Mexico (male)</td>
<td>≤ 10</td>
<td>-</td>
</tr>
<tr>
<td>Mexico (female)</td>
<td>≤ 10</td>
<td>-</td>
</tr>
<tr>
<td>Mexico</td>
<td>-</td>
<td>&gt; 10</td>
</tr>
<tr>
<td>Canada and Europe</td>
<td>≤ 10</td>
<td>-</td>
</tr>
<tr>
<td>Asia</td>
<td>≤ 5</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>≤ 10</td>
<td>&gt; 5 Asian</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and &gt; 10 Non-Asian</td>
</tr>
</tbody>
</table>
ACS File Selection and Variables

- To reduce noise, six residuals are calculated from five consecutive years of survey data for each emigrant group
- The graph shows **residual periods** from the 2006 to 2010 ACS for calculating emigration rates for the year 2010 estimate

Rate Calculation

- Data Sources:
  - \((P_1, P_2, \text{ and } I_{1-2})\) 1-year ACS files, 2006-2017
  - \((D_{1,2})\) National Center for Health Statistics life tables
- Geographic Level: Nation
- Steps:
  1. Calculate six residuals
  2. Divide by person-years \((PY_{1-2})\) to convert residuals into annualized emigration rates \((r)\)
  3. Calculate the **average rate** (if negative, set to zero)
  4. Apply the average rate to the **at-risk population** to derive annual emigration flow of the foreign born
Residual (in Thousands) and Emigration Rate Calculation

**Example: Recent Canadian and European**

<table>
<thead>
<tr>
<th>Survey File time 1, time 2</th>
<th>( P_1 )</th>
<th>( D_{1,2} )</th>
<th>( P_{exp,2} )</th>
<th>( P_2 )</th>
<th>( l_{1,2} )</th>
<th>( P_{obs,2} )</th>
<th>( P_{Y,1,2} )</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
<td>b</td>
<td>c = a - b</td>
<td>d</td>
<td>e</td>
<td>f = d - e</td>
<td>g = c - f</td>
<td>h</td>
</tr>
<tr>
<td>2006, 2008</td>
<td>1,597</td>
<td>9</td>
<td>1,588</td>
<td>1,890</td>
<td>403</td>
<td>1,488</td>
<td>100</td>
<td>3,134</td>
</tr>
<tr>
<td>2007, 2009</td>
<td>1,534</td>
<td>8</td>
<td>1,526</td>
<td>1,759</td>
<td>355</td>
<td>1,404</td>
<td>122</td>
<td>2,999</td>
</tr>
<tr>
<td>2008, 2010</td>
<td>1,479</td>
<td>8</td>
<td>1,471</td>
<td>1,726</td>
<td>348</td>
<td>1,378</td>
<td>93</td>
<td>2,904</td>
</tr>
<tr>
<td>2006, 2009</td>
<td>1,597</td>
<td>14</td>
<td>1,583</td>
<td>1,891</td>
<td>474</td>
<td>1,417</td>
<td>166</td>
<td>4,687</td>
</tr>
<tr>
<td>2007, 2010</td>
<td>1,534</td>
<td>12</td>
<td>1,522</td>
<td>1,873</td>
<td>459</td>
<td>1,414</td>
<td>108</td>
<td>4,530</td>
</tr>
<tr>
<td>2006, 2010</td>
<td>1,597</td>
<td>19</td>
<td>1,578</td>
<td>2,007</td>
<td>582</td>
<td>1,425</td>
<td>153</td>
<td>6,274</td>
</tr>
</tbody>
</table>

**NOTES:** Totals are in thousands. Calculations may not be exact due to rounding.

**SOURCE:** U.S. Census Bureau, Population Division. Simulation of 2006-2010 ACS PUMS data.

**Average Rate =** 0.031

---

Annual Emigration (in Thousands) Example: Recent Canadian and European, 2010-2018

<table>
<thead>
<tr>
<th>Estimates Year</th>
<th>Estimates Period</th>
<th>6-Rate Average</th>
<th>At-Risk Population</th>
<th>Emigration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative</td>
<td>1 April 2010 – 30 June 2018</td>
<td>-</td>
<td>-</td>
<td>442</td>
</tr>
<tr>
<td>2010</td>
<td>1 Apr 2010 – 30 June 2010</td>
<td>0.031</td>
<td>1,454</td>
<td>11*</td>
</tr>
<tr>
<td>2011</td>
<td>1 July 2010 – 30 June 2011</td>
<td>0.033</td>
<td>1,389</td>
<td>46</td>
</tr>
<tr>
<td>2012</td>
<td>1 July 2011 – 30 June 2012</td>
<td>0.036</td>
<td>1,263</td>
<td>45</td>
</tr>
<tr>
<td>2013</td>
<td>1 July 2012 – 30 June 2013</td>
<td>0.033</td>
<td>1,208</td>
<td>40</td>
</tr>
<tr>
<td>2014</td>
<td>1 July 2013 – 30 June 2014</td>
<td>0.037</td>
<td>1,211</td>
<td>45</td>
</tr>
<tr>
<td>2015</td>
<td>1 July 2014 – 30 June 2015</td>
<td>0.040</td>
<td>1,175</td>
<td>47</td>
</tr>
<tr>
<td>2016</td>
<td>1 July 2015 – 30 June 2016</td>
<td>0.053</td>
<td>1,234</td>
<td>65</td>
</tr>
<tr>
<td>2017</td>
<td>1 July 2016 – 30 June 2017</td>
<td>0.060</td>
<td>1,184</td>
<td>71</td>
</tr>
<tr>
<td>2018</td>
<td>1 July 2017 – 30 June 2018</td>
<td>0.060</td>
<td>1,201</td>
<td>72</td>
</tr>
</tbody>
</table>

**NOTES:** Estimates are for demonstration purposes only and are not official. Totals are in thousands. *2010 reflects ¼ year.

Subnational Distribution

• Distribute national emigration totals for the (7) emigrant groups by:
  • Demographic composition (age, sex, race/Hispanic origin), and
  • Geographic distribution (51 states and 3,142 counties)

• Subnational distributions determined by using a “proxy” universe based on the recent stock of the foreign-born population (e.g. geographic distribution and demographic composition of the Canadian- and European-born population living in the U.S. for 10 years or less)

• Estimates are controlled such that counties sum to the states and states sum to the nation

Evaluation and Lessons Learned

• The ACS-ACS residual method tends to be stable for groups that exhibit high levels of emigration but less so for groups that exhibit little emigration (e.g. non-recent immigrants)

• Results are highly sensitive to coverage in the surveys as well as periodic changes in the survey methods and field operations

• Due to the file selection procedure, some survey years will have disproportionately larger influence on average rate

• Large at-risk populations are more sensitive to changes in emigration rates
  • In a single year, the rate for the Non-Recent Mexico group increased by only 0.4 per 1,000 but resulted in a 153,000 increase in emigration levels
Contact Information

Thank you!

Anthony Knapp  
U.S. Census Bureau, Population Division  
Net International Migration Branch

E-mail: anthony.knapp@census.gov  
Branch Phone: 1-301-763-9379