Potential Improvements to DHS Illegal Alien Population Estimates: Collection and Use of Data

March 5, 2019
Fiscal Year 2018 Report to Congress

Homeland Security

Deputy Under Secretary for Management

Message from the Deputy Under Secretary for Management

March 5, 2019

The following report, "Potential Improvements to DHS Illegal Alien Population Estimates: Collection and Use of Data," has been prepared by the Office Strategy, Policy, and Plans.

This document has been compiled pursuant to a requirement in the Joint Explanatory Statement, which accompanies the Fiscal Year 2018 Department of Homeland Security (DHS) Appropriations Act (P.L. 115-141).

Pursuant to Congressional requirements, this report is being provided to the following Members of Congress:

The Honorable Lucille Roybal-Allard
Chairwoman, House Appropriations Subcommittee on Homeland Security

The Honorable Chuck Fleischmann
Ranking Member, House Appropriations Subcommittee on Homeland Security

The Honorable Shelley Moore Capito
Chairman, Senate Appropriations Subcommittee on Homeland Security

The Honorable Jon Tester
Ranking Member, Senate Appropriations Subcommittee on Homeland Security

Inquiries relating to this report may be directed to me at (202) 447-3400.

Sincerely,

Chip Fulghum
Deputy Under Secretary for Management

Executive Summary

Since 2005, pursuant to the Immigration and Nationality Act, DHS has released annual estimates of the illegal alien population residing in the United States. DHS generates this estimate by subtracting the estimated legally resident foreign-born population from the estimated total foreign-born population—an approach known as the "residual method."

This report describes potential steps that DHS and others could take to improve the collection and use of datasets to strengthen DHS population estimates and improve overall immigration reporting and analysis. These potential steps include:
- Assess the size of the undercount of foreign-born populations in the U.S. Census American Community Survey;
- Develop nativity-specific mortality estimates;
- Collect additional information about foreign-born departures;
- Implement the DHS Immigration Data Integration Initiative;
- Assign unique identifiers to nonimmigrants;
- Capture additional immigration benefits data in electronic data systems;
- Launch a new version of the New Immigrant Survey; and
- Include questions about adjustments to lawful permanent resident status in the Census Survey of Income and Program Participation.

Some of these potential improvements would be costly or difficult to implement; however, the report does not address these challenges or weigh these costs against potential improvements to DHS population estimates.

Potential Improvements to DHS Illegal Alien Population Estimates: Collection and Use of Data

Table of Contents

I.   Legislative Language

The Joint Explanatory Statement, which accompanies the Fiscal Year (FY) 2018 Department of Homeland Security (DHS) Appropriations Act (P.L. 115-141), includes the following requirement:

> The Department shall also report within 180 days of the date of enactment of this Act on how to improve its collection and use of data sets necessary to develop an estimate for the entire population of those illegally present in the United States.

II. Background

The Immigration and Nationality Act directs DHS to publish an annual estimate of the number of aliens unlawfully present in the United States each fiscal year, among other data.[1] DHS's Office of Immigration Statistics (OIS) has responded to this direction since 2005 by releasing annual estimates of the illegal alien population residing in the United States.[2]

As described in detail in DHS's published estimates,[3] OIS generates its estimate by subtracting the estimated legally resident[4] foreign-born population (i.e., foreign-born citizens, lawful permanent residents (LPR), asylees, refugees, and resident nonimmigrants) from the estimated total foreign-born population. This approach is known as the "residual method." OIS estimates the legally resident foreign-born population primarily based on DHS administrative records of lawful admissions, controlling for estimated, deaths, and departures; and it estimates the total foreign-born population based on U.S. Census Bureau data from the American Community Survey (ACS), controlling for the estimated under-count of the foreign-born population in Census surveys. Two prominent private-sector research organizations also produce similar findings using the residual method to estimate the illegal alien population.[5]

---

[1] Immigration and Nationality Act § 103(d), (e), 8 U.S.C. 1103(d), (e).

[2] Reports may be accessed at https://www.dhs.gov/immigration-statistics/population-estimates/unauthorized-resident. Prior to the establishment of DHS, the legacy Immigration and Naturalization Service published population estimates in 1996 and 2000. DHS failed to release a population estimate for 2013, but included a 2013 estimate as part of its 2014 release.

[3] Ibid.

[4] OIS' methodology exclusively focuses on the resident foreign-born population, i.e., excluding lawfully present short-term visitors, such as certain tourists and business travelers, who neither temporarily nor permanently are considered to reside in the United States.

[5] Jeffrey S. Passel and D'Vera Cohn, "As Mexican Share Declined, U.S. Unauthorized Immigrant Population Fell in 2015 Below Recession Level," Pew Research Center, April 25, 2017, http://www.pewresearch.org/fact-tank/2017/04/25/as-mexican-share-declined-u-s-unauthorized-immigrant-population-fell-in-2015-below-recession-level/; Center for Migration Studies (CMS), "State Level Unauthorized Population Estimates," http://data.cmsny.org/. Note that the CMS methodology is based on a one-time residual estimate (for the 2010 illegal alien population), with annual updates derived by applying population controls from the 2010 estimates to more recent U.S. Census cohorts.

III. Results

DHS's ability to describe the illegal alien population depends on its ability to describe the different population groups included in the residual methodology: the total foreign-born population and the subgroups that comprise the legally resident foreign-born population. Data limitations mean that neither of these populations can be described with precision:

- The ACS provides a direct estimate of the foreign-born population, but the ACS is widely believed to undercount the foreign-born population—and particularly the unauthorized immigrant population—at higher rates than the native-born population. The exact degree of the undercount is unknown.[6]
- DHS uses administrative data on LPR, refugee, asylee, and nonimmigrant admissions along with naturalizations to build its estimate of the legally resident foreign-born population, but DHS does not know how many lawfully admitted aliens have deceased or departed the United States.

OIS has identified 8 potential steps that DHS, other government agencies, and private-sector research bodies could take to improve the collection and use of data to address these limitations. Additionally, OIS has identified broader improvements to DHS data systems that would strengthen DHS's population estimates and its overall immigration reporting and analysis. Some of these potential improvements would be costly or difficult to implement; however, this report does not address these challenges or weigh these costs against potential improvements to DHS population estimates.

1) Assess the size of the undercount of foreign-born populations in the ACS

DHS uses the ACS to estimate the size of the total foreign-born population. However, little direct research exists on the size of the ACS undercount or how it varies by nativity and legal status, causing uncertainty in the degree of undercount for different groups of foreign-born individuals.

The most direct way to mitigate this uncertainty and improve this aspect of DHS's population estimate would be for the U.S. Census Bureau or a private-sector research body to assess the ACS undercount for relevant population groups systematically.

2) Develop nativity-specific mortality estimates

DHS's estimate of the legally resident foreign-born population is based on administrative arrival records and must control for deaths among those known to have legally entered. Mortality rates can be estimated from standard demographic tables, but the standard tables are known to

---

[6] See Kaneshiro, Matheu. "Missing Minorities? The Phases of IRCA Legalization and Relative Net Undercounts of the 1990 vis-à-vis 2000 Decennial Census for Foreign-born Cohorts." *Demography* 50, 5 (2013): 1897-1919; Van Hook, Jennifer, Frank D. Bean, James D. Bachmeier, and Catherine Tucker. "Recent Trends in Coverage of the Mexican-Born Population of the United States: Results From Applying Multiple Methods Across Time." *Demography* 51, 2 (2014): 699-726.

overestimate mortality for foreign-born populations.[7] The absence of reliable estimates based on nationally representative data is an additional source of uncertainty in the DHS population estimate.

The most direct way to mitigate this uncertainty and improve this aspect of DHS's population estimate would be for the U.S. Census Bureau or a private-sector research body to produce nativity-specific mortality tables.

3) Collect additional information about foreign-born departures

DHS's estimate of the legally resident foreign-born population also must control for departures by those known to have legally entered the United States. Since 2015, DHS has collected reliable data on alien departures by air and sea. DHS does not have reliable departure data for aliens who departed prior to 2015, and DHS continues to lack comprehensive data on land departures. These gaps mean that DHS relies on statistical models of LPR, nonimmigrant, and other foreign-born departures to build its estimate of the legally resident population, adding to uncertainty in the population estimate.

The most direct way to strengthen the estimates of LPR emigration rates and nonimmigrant population size would be for the U.S. Census Bureau or a private-sector research body to collect survey data on emigration rates specific to LPRs and on the current residency status of nonimmigrants not known to have departed. Upon expansion of DHS's entry/exit tracking system to include exits via land, DHS will rely on these data to estimate nonimmigrant population size in the future.

4) Implement the DHS IDII

Currently, DHS maintains four dozen separate immigration data systems, most of which are organized by events, rather than by individuals. As a result, individuals applying for multiple benefits, admitted multiple times, or adjusting their immigration status touch several different data systems, and it is often difficult to match their records across systems. One result is that DHS's population estimates do not incorporate information comprehensively about when and how legally resident foreign-born individuals change their status over time, which may cause certain groups to be under- or over-estimated. For example, DHS's estimate of legally resident nonimmigrants does not account for nonimmigrants who adjust their status to become LPRs; and DHS's estimate of the length of residency for LPRs does not account for aliens who make multiple trips to the United States prior to adjusting status.

In general, the DHS Immigration Data Integration Initiative (IDII) would mitigate problems from siloed data systems by copying data from multiple systems into a single environment, linking records at the person level, and using linked records to build a single, comprehensive, and

---

[7] Singh, Gopal K., and Barry A. Miller, "Health, Life Expectancy, and Mortality Patterns among Immigrant Populations in the United States." *Canadian Journal of Public Health/Revue Canadienne De Sante'e Publique* 95, 3 (2004): I14-21.

integrated dataset. [8,9] As of December 2018, the IDII has linked records from 19 DHS and U.S. Department of Justice data systems related to the immigration enforcement lifecycle, but it does not include a comprehensive set of immigration benefits data yet.

Once the IDII dataset includes more robust person-level benefits data, DHS expects to use IDII data to match each admissions record to the alien's current status more accurately and to determine whether the alien has departed the country, strengthening and simplifying the creation of legally resident and illegal population estimates. Person-centric data would offer numerous additional benefits, including faster and more comprehensive reporting and the improved ability to conduct evidence-based planning and analysis on the end-to-end immigration benefits and enforcement systems. [10]

5) Assign unique identifiers to nonimmigrants

Currently, DHS does not assign unique (person-specific) identifiers to nonimmigrant visitors. The lack of such identifiers is an additional obstacle to including comprehensive data on adjustments of status in DHS's estimate of the nonimmigrant population because it means that any effort to link nonimmigrant and LPR records must employ less reliable matching methodologies rather than one-to-one unique identifier matches.

If DHS were to begin assigning unique identifiers such as Alien Numbers to nonimmigrants and ensuring that these data are carried forward to their I-485 applications for adjustment of status, it would simplify efforts to link nonimmigrant and LPR data systems going forward. Unique identifiers would allow DHS to account for adjustments of status in its nonimmigrant population estimate more easily, thereby strengthening its illegal alien estimate. Additionally, assigning unique identifiers would strengthen and simplify DHS's ability to track suspected nonimmigrants who overstay the time permitted for their visits.

6) Capture additional immigration benefits data in electronic systems

Currently, U.S. Citizenship and Immigration Services (USCIS) collects certain data elements on paper benefits application forms that are not captured in DHS's electronic record systems. Examples include certain information about aliens' occupations and education levels, among

---

[8] DHS operational components also link certain records to build integrated datasets for operational purposes. For example, the U.S. Customs and Border Protection (CBP) Arrival and Departure Information System links records from a number of ICE, CBP, USCIS, and U.S. Department of State systems to create a person-centric view of the immigration lifecycle for nonimmigrants admitted to the United States; and the USCIS Customer Profile Management Service serves as a person-centric repository for information about aliens requesting immigration benefits. The IDII is working with CBP and USCIS to leverage existing efforts to link immigration records and build an integrated dataset.

[9] In principle, the most direct way to mitigate siloed data would be to build a single person-centric immigration data system that would be shared by all of DHS's operational Components. A front-end system would offer certain additional operational benefits, but building such a system is believed to be far costlier and more disruptive to current operations than constructing a back-end analytic dataset, pending the results on an ongoing analysis of alternatives.

[10] See DHS Office of Immigration Statistics, "2014 Southwest Border Encounters: Three-Year Cohort Outcomes Analysis," August 2018, https://www.dhs.gov/publication/2014-southwest-border-encounters-three-year-cohort-outcomes-analysis.

other data elements. As USCIS moves increasingly to electronic filing systems, DHS could capture most or all data elements on immigration benefits forms in its electronic systems. Such data could be used to add descriptive data to DHS's lawfully present population estimates, and would allow DHS to further inform Congress and the public about the effects of U.S. immigration policy and migration flows.

7) Launch a new version of NIS

The New Immigrant Survey (NIS) is a multi-cohort panel study of new LPRs. A survey pilot study was implemented by the Immigration and Naturalization Service in 1996, and cohorts of LPRs were surveyed in 2003 and from 2007 to 2009. The study analyzed a nationally representative sample of LPRs and included questions about return migration by LPRs to their countries of origin along with extensive additional questions about LPRs' labor market experiences and contributions to the U.S. economy, the relationship between lawful and unlawful immigration, family reunification, assimilation of first- and second-generation immigrants, health and well-being, and other factors.[11]

DHS could work with other Federal agencies and private-sector partners to launch a new version of the NIS. Such a survey could be used to strengthen DHS's estimates of foreign-born departures and mortality, and would have the further benefit of providing extensive additional policy-relevant data on immigrant integration.

8) Include questions about adjustments to LPR status in the Census Survey of Income and Program Participation

The U.S. Census Survey of Income and Program Participation (SIPP) is a national panel survey based on a stratified sample of the U.S. noninstitutionalized population. The survey is designed to measure income, labor force participation, social program participation and eligibility, and general demographic characteristics of people participating in existing federal, state, and local programs.

The 2008 version of the SIPP included questions about whether foreign-born residents had arrived with an immigrant visa and whether they had adjusted to LPR status within the United States; but the 2014 version of the SIPP only asked about status at arrival. Private-sector researchers have used the 2008 SIPP data along with a series of "logical edits" to impute information about legal and illegal immigration status in the ACS, but they are unable to do so with the more limited data in the 2014 SIPP.[12] If the U.S. Census Bureau were to include questions about adjustments of status in the next version of the SIPP, private-sector and potential DHS researchers could use this data source once again in order to produce detailed demographic estimates of the illegal alien population.

---

[11] See Princeton University, "The New Immigrant Survey," https://nis.princeton.edu/index.html.
[12] See Migration Policy Institute, "Unauthorized Immigrant Population Profiles," https://www.migrationpolicy.org/programs/us-immigration-policy-program-data-hub/unauthorized-immigrant-population-profiles.

IV. Conclusion/DHS Action Plan

Some of the potential improvements identified in this report would be costly or difficult to implement; however, this report does not address these challenges or weigh these costs against the potential improvements to DHS population estimates.

OIS will work with the U.S. Census Bureau to explore the possibility—subject to available resources and legal requirements—of additional research to assess the size of the undercount of foreign-born populations in the ACS, to develop nativity-specific mortality estimates, and to collect additional information about foreign-born departures.

OIS will work to include other potential improvements identified in this report as DHS implements changes to its current data systems:
- Once DHS completes a reliable land-based entry/exit system to complement the existing air- and sea-based system, OIS will incorporate departure information derived from entry/exit data in its nonimmigrant population estimate.
- OIS will support implementation of the IDII as part of a broader effort to strengthen DHS's reporting and analysis of immigration data. In particular, OIS will work to incorporate core immigration benefits datasets into the IDII dataset during FY 2019, including comprehensive data on LPR admissions, naturalizations, and nonimmigrant arrivals. Once the IDII establishes a comprehensive person-centric dataset for reporting and analysis, OIS will use person-centric data to include information about adjustments of status in its population estimates systematically.

Should DHS and external partners take additional actions identified in this report, OIS will include the additional resulting data in its population estimates:
- If DHS begins assigning unique identifiers to nonimmigrants, OIS will use such identifiers to link nonimmigrant and LPR records more reliably, both for the purposes of strengthening its population estimates and to strengthen the IDII comprehensive integrated dataset.
- If DHS begins capturing additional benefits data in its electronic systems, OIS will work to incorporate such data in its research products.
- If DHS works with other partners to launch a new version of NIS, OIS will work to incorporate NIS data in its research products.
- If the U.S. Census Bureau resumes asking about adjustments to LPR status in a future version of SIPP, OIS will work to incorporate SIPP imputations in its research products.

Appendix: List of Abbreviations and Definitions

Abbreviation	Definition
ACS	American Community Survey
CBP	U.S. Customs and Border Protection
CMS	Center for Migration Studies
DHS	Department of Homeland Security
FY	Fiscal Year
IDII	Immigration Data Integration Initiative
LPR	Lawful Permanent Resident
NIS	New Immigrant Survey
OIS	Office of Immigration Statistics
SIPP	Survey of Income and Program Participation
USCIS	U.S. Citizenship and Immigration Services