

Social Security

SOCIAL SECURITY BULLETIN Vol. 85, No. 2, 2025

IN THIS ISSUE:

- Measuring the Number of Unauthorized Immigrants in the United States: A Review of the Residual Estimation Method
- Measuring the Economic and Sociodemographic Characteristics of Unauthorized Immigrants in the United States with Survey Data
- A New Way to Estimate the Number of Unauthorized Immigrants in the United States

The *Social Security Bulletin* (ISSN 1937-4666) is published quarterly by the Social Security Administration, 250 E Street SW, 8th Floor, Washington, DC 20254.

Suggestions or comments concerning the *Bulletin* should be sent to ssb@ssa.gov.

Contents of this publication are not copyrighted; any items may be reprinted, but citation of the *Social Security Bulletin* as the source is requested. The *Bulletin* is available on the web at https://www.ssa .gov/policy/docs/ssb/.

Errata Policy: If errors that impair data interpretation are found after publication, corrections will be posted as errata on the web at https://www.ssa.gov/policy /docs/ssb/v85n2/index.html.

The findings and conclusions presented in the *Bulletin* are those of the authors and do not necessarily represent the views of the Social Security Administration.

SSA Publication No. 13-11700 Produced and published at U.S. taxpayer expense

Frank J. Bisignano Commissioner of Social Security

Douglas A. Turnbull Associate Commissioner for Analytical Services

Editorial and Production Staff Jessie Ann Dalrymple Margaret F. Jones Benjamin Pitkin Alison Sturgeon

Social Security Bulletin

Volume 85 • Number 2 • 2025

Articles

1 Measuring the Number of Unauthorized Immigrants in the United States: A Review of the Residual Estimation Method

by Harriet Duleep, Dave Shoffner, Robert V. Gesumaria, and Christopher R. Tamborini

In this first of three related articles on unauthorized immigration, the authors describe what has been, to date, the primary method of estimating the unauthorized immigrant population in the United States: the residual estimation method. This method entails subtracting an estimated population of authorized immigrants from an estimated population of all foreign-born U.S. residents based on national survey data; the residual result is presumed to constitute the unauthorized immigrant population. Federal agencies and independent research institutes have used the residual method to provide measures of the unauthorized immigrant population that are useful in modeling factors affecting Social Security, other government programs, and the overall economy. The authors describe the residual estimation process, present historical estimates it has produced, and review selected recent alternative estimation methods that either refine and extend the residual method or take a markedly different approach.

9 Measuring the Economic and Sociodemographic Characteristics of Unauthorized Immigrants in the United States with Survey Data by Christopher R. Tamborini, Harriet Duleep, Robert V. Gesumaria, and Dave Shoffner

In this second of three related articles on unauthorized immigration, the authors describe statistical techniques that researchers have used to estimate not only the size but also selected socioeconomic characteristics of the unauthorized immigrant population in the United States. Using data from national surveys, researchers impute the number of likely unauthorized immigrants as well as indicators of their earnings, education, occupation, and industry of employment. The authors describe the imputation methods those researchers have developed, review studies that have employed the methods and summarize selected findings, and discuss research on the correlation between U.S. residential permanence and human capital investment among likely unauthorized immigrants.

17 A New Way to Estimate the Number of Unauthorized Immigrants in the United States by Robert V. Gesumaria, Harriet Duleep, Christopher R. Tamborini, and Dave Shoffner

In this third of three related articles on unauthorized immigration, the authors introduce a new method of estimating the number of unauthorized immigrants in the United States by exploiting discrepancies between Current Population Survey (CPS) data and Social Security administrative data on Social Security numbers (SSNs). Potential unauthorized immigrant status is indicated when the SSNs reported by CPS respondents and the SSNs recorded in linked administrative data do not match. The authors use the nonmatching SSN data to identify likely unauthorized immigrants and apply a series of logical adjustments to refine the estimated population counts. Although the methodologies of the residual estimation method (described in the first of these three articles) and this new process are entirely different, the two approaches yield similar results. The consistency of estimates resulting from different methodologies suggests that both are likely accurate.

Measuring the Number of Unauthorized Immigrants in the United States: A Review of the Residual Estimation Method

by Harriet Duleep, Dave Shoffner, Robert V. Gesumaria, and Christopher R. Tamborini*

This first of three related articles on the U.S. unauthorized immigrant population discusses the predominant method of measuring that population: the residual estimation method, so named because an estimated population of authorized immigrants is subtracted from an estimated population of all foreign-born U.S. residents, with the residual result presumed to constitute the unauthorized immigrant population. We describe the method step by step and trace historical trends in the estimates it produces. We then differentiate between unauthorized immigrants who arrived via entry without inspection and those who overstayed a visa, noting that the shares who overstayed a visa have risen in recent years. We then discuss several recent studies: one that used a different methodology to estimate the unauthorized immigrant population, and others that proposed adjustments to the residual method.

Introduction

Hinting at the challenges of measuring their numbers, unauthorized immigrants are known by multiple labels-such as illegal immigrants, undocumented immigrants, extralegal immigrants, and unlawful permanent residents. Their elusive status makes estimating their current and historical numbers challenging. This article uses the terms "unauthorized immigration" and "unauthorized immigrants," which we define as foreign-born individuals who reside in the United States without a valid temporary visa, a permanent resident visa ("green card"), or U.S. citizenship. The importance of accurately measuring this population goes beyond issues directly affected by unauthorized immigration. For example, discussions about appropriate levels of legal immigration hinge on the assumed levels and trends in the unauthorized immigrant population.

This article attempts to summarize the complicated literature on measuring the U.S. unauthorized immigrant population. Following this introduction, it begins with a section that describes the residual estimation method—commonly known as simply the residual method—to date, the most widely used approach to estimating the U.S. unauthorized immigrant population. The second section presents estimates of the unauthorized immigrant population in the United States from various studies using the residual method. The third section highlights the increasing incidence of visa overstays among unauthorized

Selected Abbreviations

ACS	American Community Survey
CMS	Center for Migration Studies of New York
CPS	Current Population Survey
DHS	Department of Homeland Security
EWI	entry without inspection
MPI	Migration Policy Institute

^{*} Harriet Duleep is a researcher with the Social Security Administration (SSA); a research professor with the Public Policy Program, College of William and Mary; and a research fellow with the Institute for the Study of Labor (IZA) and with the Global Labor Organization. Dave Shoffner is an analyst, Christopher Tamborini is a researcher, and Robert Gesumaria is a researcher and IT specialist, all with SSA.

Note: Contents of this publication are not copyrighted; any items may be reprinted, but citation of the Social Security Bulletin as the source is requested. The Bulletin is available on the web at https://www.ssa.gov/policy/docs/ssb/. The findings and conclusions presented in the Bulletin are those of the authors and do not necessarily represent the views of the Social Security Administration.

immigrants and the significance of that trend for measuring the unauthorized immigrant population. The fourth section describes a widely critiqued study that estimated about twice the number of unauthorized immigrants as the residual-method studies. The fifth section reviews a revision to the residual method used in post-2010 research from the Center for Migration Studies of New York (CMS), a public policy educational institute. The article then closes with a concluding summary.

The Residual Method of Estimating the Unauthorized Immigrant Population

The residual method is used in computing various benchmark estimates, including those of the Department of Homeland Security (DHS) and independent think tanks such as the Pew Research Center (Pew), the Migration Policy Institute (MPI), and CMS.¹ Residualmethod estimates also underlie Social Security's actuarial forecasts and demographic assumptions.

As described in Warren and Passel (1987), the residual-method process begins with the use of national survey data, such as Census Bureau's annual American Community Survey (ACS) or the Annual Social and Economic Supplement to its Current Population Survey (CPS),² to estimate the total foreign-born population.³ The next step is to identify the foreign-born individuals who reside in the United States legally, using one or more of various possible data sources. For example, DHS uses its administrative records to identify all who are citizens or have legal permanent resident status. Administrative data from the Department of Health and Human Services' Office of Refugee Resettlement provide counts of refugees and asylees, which are added to the count of legal permanent residents. Researchers then adjust the sum of authorized immigrants by applying demographic statistical techniques to account for deaths, emigration, and new arrivals in the year. Alternatively, Pew, MPI, and CMS impute the likely authorized-resident population based on certain characteristics of foreign-born individuals available in the survey data. The Congressional Budget Office uses a similar imputation methodology to identify the legal status of foreign-born residents (Heinzel, Heller, and Tawil 2021). After subtracting the estimated total number of legal residents from the total foreign-born population, the residual population is identified as likely to be composed of unauthorized immigrants.

The results are adjusted by additional corrections that account for the likely undercount of the unauthorized immigrant population in the ACS and CPS data. Additional survey data provide indications of whether immigrant groups participated in the decennial census or a prior national survey, and how participation rates vary depending on selected respondent characteristics. Researchers may also consult survey and census results in Mexico, which provide additional information on how many Mexican immigrants might be missed in U.S. data sources. Such adjustments can affect the estimates by as little as 5 percent or as much as 15 percent and may vary by age, sex, years since arrival, and other factors.

The tabulation below shows three relatively recent estimates of the unauthorized immigrant population. Their similarity is not surprising: The DHS analysis of the unauthorized immigrant population in January 2018 describes the minor differences between its methodology and those used by Pew, MPI, and CMS, and acknowledges the similarity of the resulting estimates (Baker 2021).

Source	Year	Unauthorized immigrant population estimate (millions)
Pew	2017	10.5
CMS	2018	10.7
DHS	2018	11.4

SOURCES: Passel (2019); Warren (2019b); and Baker (2021).

Pew and CMS use federal surveys and the residual method to generate two major series of recurring estimates of the unauthorized immigrant population (Passel 2005, 2019; and Warren 1997, 2003, 2018a, 2018b, 2019a, 2019b). The CPS Annual Social and Economic Supplement and the ACS are the surveys most often used with the residual method for this purpose. Both surveys ask respondents where they were born and whether they are U.S. citizens. An advantage of using the ACS and the CPS over panel studies such as the Survey of Income and Program Participation (SIPP) is that their much larger sample sizes ostensibly provide more accurate data on immigrant populations, particularly at the state and local levels.⁴ Providing more recent measurements of the unauthorized immigrant population in the United States, DHS estimates 10.99 million as of January 2022 (Baker and Warren 2024), Pew estimates 11.0 million for 2022 (Passel and Krogstad 2024), and CMS estimates 10.94 million for 2022 (Warren 2024).

Historical Residual-Method Estimates

Pew estimates that the number of unauthorized immigrants rose from about 3.5 million in 1990 to about 8.6 million in 2000 (Krogstad, Passel, and Cohn 2019). Average net growth in the unauthorized immigrant population was about 10 percent each year during that period. In the 2000s, the average net increase was steady at about 4 percent per year. The total unauthorized immigrant population peaked at about 12 million in 2007 then decreased to about 11 million in 2009. Since 2009, the number of unauthorized immigrants has stayed relatively flat at about 11 million (Warren 2018b).⁵ Different studies cover different time periods, but the trends they report follow a similar pattern.

Pew and CMS analyses show substantial declines in unauthorized immigration from Mexico, Eastern Europe, and South America since 2007. About 4.9 million Mexican unauthorized immigrants lived in the United States in 2017, compared with 6.9 million in 2007. Non-Mexican unauthorized immigration—mainly from Asia, Central America, and Africa—rose from about 5.3 million in 2007 to 5.5 million in 2017. Reflecting acute political and economic challenges, three "Northern Triangle" countries in Central America—Guatemala, Honduras, and El Salvador—have since 2015 been the origins of increasing unauthorized immigration (Krogstad, Passel, and Cohn 2019).

As net unauthorized immigration declined or slowed in recent years, the total estimated populations of unauthorized immigrants declined substantially in many states. From 2010 through 2017, their numbers dropped by about 500,000 in California, about 150,000 in New York, and about 75,000 in Illinois and Florida (Pew Research Center 2019). Consequently, the recentarrival share of the unauthorized immigrant population has declined: The fraction who have been in the United States for 10 or more years rose from about two in five persons in 2007 to about two in three in 2017 (Krogstad, Passel, and Cohn 2019).

Visa Overstays

Many persons enter the U.S. legally as tourists, students, temporary workers, or members of other permitted categories, but overstay the period designated in their visas. From 2009 to 2019, "the primary mode of entry for the unauthorized population [was] to overstay temporary visas" (Warren 2019b).

Despite their growing share of the unauthorized immigrant population, accurate estimates of their

numbers are elusive largely because there is no direct information on visa overstays. Warren and Warren (2013a, 2013b) and Warren (2019a, 2019b) propose an estimation method that uses indirect data sources. A key component of that method is to differentiate unauthorized immigrants who arrive via entry without inspection (EWI) from those who overstay their visas. This involves identifying the country of origin of immigrants estimated to be unauthorized under the residual method. Unauthorized immigrants from four countries—Mexico, Honduras, Guatemala, and El Salvador—are considered to be more likely to have entered without inspection, while those from other countries are considered more likely to have overstayed a visa.

Specifically, Warren and Warren (2013a) use data from the ACS on the resident population whose country of origin is Mexico, Honduras, Guatemala, or El Salvador, then subtract from that count the estimated number of legal residents (based on DHS data and demographic imputations). To account for slight variations in these assumptions, Warren and Warren multiply the residual estimate of the combined unauthorized immigrant population from Mexico, Honduras, Guatemala, and El Salvador by 90 percent; the result is the estimated EWI population. The rest of the estimated unauthorized immigrant population is assumed to be composed of visa overstays.

From 2000 to 2010, both the new EWI and the new visa overstay populations declined substantially (Warren and Warren 2013a, 2013b). By the 2010s, visa overstays constituted an expanding proportion of new unauthorized immigrants because the number of EWI arrivals was declining more rapidly than that of overstays. Nevertheless, among the total unauthorized population residing in the United States as of 2017, an estimated 4.9 million, or 46 percent, had overstayed a visa, and the other 54 percent had originally arrived via EWI (Warren 2019a).

Annual compilations of DHS data on lawful immigrant entries and deportations (DHS 2018) and on visa overstays (DHS 2020) help to inform estimates of the yearly arrivals of unauthorized immigrants. Yet converting these annual flow estimates into estimates of the total U.S. unauthorized immigrant population is hampered by the difficulty of estimating the numbers of unauthorized immigrants who emigrate, die, or attain legal residency.

DHS publishes annual visa-holder *Entry/Exit* Overstay Reports, of which the fiscal year 2019 edition (DHS 2020) covers the most recent prepandemic year and thus reflects the overstay flow of a typical year. It reports 574,740 suspected visa overstays in that year. Most of these individuals (92 percent) overstayed visas permitting temporary visits for business (B1) or tourism (B2), or came from countries that participate in the DHS Visa Waiver Program (VWP), which allows business or tourism visits of up to 90 days without a visa.⁶

The other 8 percent, or about 45,000 individuals, overstayed other types of visas in fiscal year 2019. Some of them likely held an Employee Authorization Document that allowed an extended period of employment and in some cases may have entitled them to obtain a Social Security number, unlike those who overstayed B1/B2 visas or visits from VWP participant countries.

An Alternative Methodology: Analyzing Border Crossings and Emigration

Because the unauthorized population is difficult to identify and measure, alternatives to the residual method should be encouraged, studied, and evaluated. Offering one such alternative, Fazel-Zarandi, Feinstein, and Kaplan (2018) estimate the size of the unauthorized immigrant population by analyzing cumulative inflows (EWI and visa overstays) and outflows (deportations, voluntary emigration, mortality, and legal status change). From statistics on border apprehensions and visa overstays, one can theoretically deduce the number of new unauthorized immigrant arrivals by imagining the flow in reverse: How many new arrivals are needed to offset the observed number of apprehensions and visa overstays? For outflows, the authors use data from DHS, the Centers for Disease Control and Prevention, Warren and Warren (2013b), and other sources.⁷ With this methodology, they estimate an unauthorized immigrant population of over 22 million, about twice the estimated counts based on the residual method.

The approach has been criticized by the Center for Immigration Studies (Camarota 2018), an independent research organization that advocates reduced immigration, and the CMS, which advocates increased immigration. Both groups take issue with using data on the apprehension of border crossers to measure inflows because the count of apprehension incidents exceeds the number of individuals apprehended.⁸ As such, they exaggerate the inflow of unauthorized immigrants from Mexico. Fazel-Zarandi, Feinstein, and Kaplan's model estimates that the U.S. unauthorized immigrant population from Mexico increased by 17.5 million in the 1990s. Warren (2018a) tests this estimate using Mexican census data⁹ for the period from 1990 to 2000. The population of Mexico was 86.1 million in 1990. From 1990 to 1999, 27.6 million births and 4.3 million deaths occurred there. Assuming zero net international migration during the decade, the population in 2000 would have been 109.4 million. The population in 2000 was 103.9 million, indicating a net migration of 5.5 million people from Mexico during the 1990s.

The 2000 U.S. census supports the 5.5 million emigrant statistic from Mexico in the 1990s: It counted 4.5 million immigrants from Mexico who entered from 1990 through 1999. Thus, official Mexican and U.S. statistics suggest that emigration from Mexico to the United States ranged from 4.5 million to 5.5 million people in the 1990s. This result implies that the estimate of 17.5 million in the Fazel-Zarandi, Feinstein, and Kaplan study overestimates the unauthorized population from Mexico by about 12 million.

That study may also overestimate the count of unauthorized immigrants in the United States by underestimating the extent to which these individuals subsequently emigrate. These discrepancies may help explain the disconnect between the unauthorized immigrant populations estimated by Fazel-Zarandi, Feinstein, and Kaplan and those derived from government statistics on births and the school-aged child population, which align more closely with the estimates produced by the residual method (Camarota 2018).

The CMS Methodology for Post-2010 Estimates

CMS used its own version of residual techniques to produce annual estimates of the unauthorized immigrant population from 2010 to 2019. The CMS estimates are based on ACS data for immigrants who arrived after 1981. CMS estimated that the total unauthorized immigrant population for 2010 was 11,725,000 (Warren and Warren 2013b), slightly more than the DHS estimate of 11.6 million (Baker 2021). However, for subsequent years, CMS altered its methodology for estimating components of immigrant population change (Warren 2021). Below, we summarize Warren's revised CMS methodology.

In the first step, CMS used 2010 ACS data on the foreign-born population from each of 145 countries

or areas of origin and used the conventional residual method to estimate the unauthorized immigrant population from each of those places of origin. In the second step, CMS applied "logical edits" to identify post-1981 immigrant arrivals from each place of origin who had likely attained legal status based on certain characteristics. For example, CMS considered individuals who worked in occupations that generally require legal status, had legal temporary migrant status, were immediate relatives of U.S. citizens, received public benefits that are restricted to legal residents, were aged 60 or older at entry, or were from countries from which most U.S. arrivals would be refugees to be legal residents. CMS then subtracted the numbers of these presumed legal residents from the conventional estimates of unauthorized immigrants to calculate an "edited population" of unauthorized immigrants from each country or area of origin. In the third step, CMS consulted independent databases to refine its figures, which resulted in estimates for 145 countries or areas of origin that were deemed plausible for each area and summed to 10,850,000.10

Next, a set of ratios was computed by dividing the residual-method estimate of the unauthorized immigrant population for 2010 by the edited population for each of the 145 countries or areas. These individual country ratios form the basis of the detailed CMS estimates for each year after 2010.

To illustrate the methodology step by step, Warren (2021) uses Mexico as the country-of-origin example. Using the residual method, CMS estimated that 6.138 million unauthorized residents were among the noncitizens from Mexico counted in the 2010 ACS. CMS found that, in all, the 2010 ACS counted 8.062 million noncitizen residents who arrived from Mexico after 1981. Of those, 1.645 million were determined to be likely legal residents, producing an edited population of 6.417 million. Dividing the number of unauthorized residents in 2010 estimated with the residual method (6.138 million) by the edited population (6.417 million) yields 0.956.11 CMS then multiplied the ratio 0.956 by the edited population calculations for each year 2011–2016 to estimate the unauthorized immigrant population from Mexico. This procedure was repeated for each country or area of origin.

CMS used the same country-specific ratios to estimate the unauthorized immigrant population each year from 2011 through 2016 because the 2010 population numbers and ratios were "the cumulative result of legal and undocumented entries throughout the entire 28-year period from 1982 to 2010. If the proportion of unauthorized to legal entries in 2011 deviated from that long-term trend, the effect on the ratio, and thus the estimate [for 2011], would likely be small. Examination of the annual ACS data for noncitizens, and of annual DHS data for legal permanent residents... admitted, shows that arrivals and population trends for nearly all countries [of origin] tend to be fairly stable over time" (Warren 2021). In any event, annual variations in the ratios for each of the 145 countries or areas would likely tend to offset each other over time, such that the total unauthorized immigrant population estimates would be stable in the period 2011–2016. However, for its annual estimates for 2017–2019, CMS devised a set of procedures to revise the ratios as needed. CMS based those revisions on administrative data and estimates of noncitizen deaths and emigration from DHS (Warren 2021).

Regardless of the rigor and sophistication of the methodologies they use, how do we know whether the estimates from CMS—and other organizations accurately reflect the size of the unauthorized immigrant population? The validity of residual-method estimates of the unauthorized immigrant population rests in part on the adequacy of the adjustments for undercount in the ACS. CMS evaluated the accuracy of those adjustments in part by comparing its unauthorized immigrant population estimates with those of reliable administrative sources such as DHS' data sets on Deferred Action for Childhood Arrivals (DACA) applicants and Temporary Protected Status (TPS) beneficiaries, the latter group being principally from El Salvador, Honduras, and Haiti.

If the CMS estimates of the total unauthorized immigrant population are too low, then the number of DACA applicants reported by DHS would be much higher than the CMS projections. In fact, DHS reported about 800,000 DACA applicants for 2010, considerably *lower* than the approximately 1.2 million applicants projected by CMS and other organizations. The number of DACA applicants likely was lower than the actual population of individuals eligible to apply for DACA because, as in other legalization programs, some who are eligible do not apply, for various reasons (Warren 2021).

It is difficult to estimate the number of TPS beneficiaries from specific countries such as El Salvador, Honduras, and Haiti and for specific periods of entry. The DHS data, based on administrative records, provide an approximate number that should have been counted in the ACS. Warren (2021) observes that "for each country, the CMS estimates are higher than the DHS data. The CMS and DHS numbers differ because of sampling variability in the ACS, [the] timing of the estimates (2015 vs. 2017), and other differences in the underlying data. The similarity of these figures, however, provides additional strong support for the overall [unauthorized immigrant] population size and the adequacy of [each organization's] adjustments for undercount in the ACS."

Conclusion

This first of three articles on unauthorized immigration focuses on the predominant method of measuring the unauthorized immigrant population in the United States: the residual estimation method. It also highlights the ever-changing nature of unauthorized immigration: Past trends do not always predict future patterns. For instance, Warren (2024) finds that after a decade of decline, the U.S. unauthorized immigrant population increased by 650,000 in 2022.

Refining the residual method and developing potential new methods—and analyzing their strengths and weaknesses—illuminates the various complexities of measuring unauthorized immigration. For instance, Fazel-Zarandi, Feinstein, and Kaplan (2018), and its critique (Warren 2018a), highlight that different immigrant groups have different patterns of subsequent emigration from the United States. Accounting for such differences is important in measuring their populations, earnings, family structures, and other characteristics, as our second article (Tamborini and others 2025) shows.

This article also describes a methodological extension of the residual method used in estimating post-2010 unauthorized immigrant populations. The methodology has evolved to account for immigrants' countries of origin and mode of arrival (EWI or visa overstay). Our second article also explores additional statistical techniques that may be better able to identify unauthorized immigrant populations and provide a deeper understanding of their characteristics. Our third and concluding article (Gesumaria and others 2025) introduces a new method of estimating the unauthorized immigrant population.

Notes

Acknowledgments: We conducted this study under the direction and with the support of Mark J. Warshawsky. We thank Stephanie Myers, Anya Olsen, Mark Sarney, Steve Goss, Tokunbo Oluwole, Ben Danforth, Kent O. Morgan, Steve Robinson, Michael Morris, and Gayle Reznik.

We would especially like to thank Mark Regets (Senior Fellow, National Foundation for American Policy, former immigration expert with the National Science Foundation), Robert Warren (Senior Visiting Fellow at the Center for Migration Studies of New York, who served as a demographer for 34 years with the Census Bureau and the former Immigration and Naturalization Service), and Ben Pitkin (editor, *Social Security Bulletin*).

¹ The benchmark estimates from DHS and these independent think tanks in turn provide the basis for Census Bureau and Congressional Budget Office estimates. The Congressional Research Service and the University of Pennsylvania's U.S. budget model also rely on estimates generated by this method. Note that CMS used the residual method as part of its estimation procedure for 2010 but has revised it for its post-2010 estimates, as we will describe.

² The Census Bureau conducts the CPS for the Bureau of Labor Statistics.

³ Pew, MPI, and CMS have all used CPS data. However, their more recent estimates (from 2005 forward) have been based on the ACS.

⁴ The ACS, CPS, and SIPP are nationally representative household surveys administered by the Census Bureau. For more information on ACS, see https://www.census .gov/programs-surveys/acs. For more information on the CPS, see https://www.census.gov/programs-surveys/cps /technical-documentation.html. For detailed descriptions of the SIPP data, see https://www.census.gov/programs -surveys/sipp.html.

⁵ There is little evidence, if any, that the recession stopped the upward trend in unauthorized immigrant population after 2008. More likely, a long-term increase in departures coincided with declining arrivals, mostly from Mexico. If the recession had been the cause of a decrease after 2008, the unauthorized immigrant population would have increased thereafter, but it did not.

⁶ Because DHS cannot acquire complete departure records for many VWP participants, it likely overestimates the number of overstays for visitors from those countries.

⁷ The authors note that "voluntary emigration rates are the largest source of outflow and the most uncertain based on limited data availability."

⁸ Some individuals are apprehended multiple times, on repeated instances of returning to a U.S. residence from visits to Mexico.

⁹ Specifically, statistics from the National Institute of Statistics and Geography of Mexico.

¹⁰ The 10,850,000 figure reflects the estimated undocumented immigrant population based on ACS data prior to adjusting for undercount. That adjustment produces the 11,725,000 population estimate (Warren 2014).

¹¹ The ratio for Mexico is the highest for any country. Most countries of origin with relatively large unauthorized immigrant populations have ratios in the range of about 0.80 to 0.92.

References

- Baker, Bryan. 2021. Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2015–January 2018. Washington, DC: Department of Homeland Security, Office of Immigration Statistics. https://ohss.dhs.gov/topics/immigration/unauthorized /population-estimates.
- Baker, Bryan, and Robert Warren. 2024. *Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2018–January 2022.* Washington, DC: Department of Homeland Security, Office of Homeland Security Statistics. https://ohss.dhs.gov/topics /immigration/unauthorized/population-estimates.
- Camarota, Steven A. 2018. "New Estimate of 22 Million Illegal Immigrants Is Not Plausible." https://cis.org /Camarota/New-Estimate-22-Million-Illegal-Immigrants -Not-Plausible.
- [DHS] Department of Homeland Security. 2018. 2018 Yearbook of Immigration Statistics. Washington, DC: DHS, Office of Immigration Statistics. https://ohss.dhs.gov /topics/immigration/yearbook/2018.

Fazel-Zarandi, Mohammad M., Jonathan S. Feinstein, and Edward H. Kaplan. 2018. "The Number of Undocumented Immigrants in the United States: Estimates Based on Demographic Modeling with Data from 1990 to 2016." *PLoS ONE* 13(9): e0201193. https://doi .org/10.1371/journal.pone.0201193.

Gesumaria, Robert V., Harriet Duleep, Christopher R. Tamborini, and Dave Shoffner. 2025. "A New Way to Estimate the Number of Unauthorized Immigrants in the United States." *Social Security Bulletin* 85(2): 17–23.

Heinzel, Julia, Rebecca Heller, and Natalie Tawil. 2021. "Estimating the Legal Status of Foreign-Born People." CBO Working Paper No. 2021-02. Washington, DC: Congressional Budget Office. https://www.cbo.gov /publication/57022.

Krogstad, Jens Manuel, Jeffrey S. Passel, and D'Vera Cohn. 2019. "5 Facts About Illegal Immigration in the U.S." Pew Research Center Short Read. https://www .pewresearch.org/short-reads/2019/06/12/5-facts-about -illegal-immigration-in-the-u-s/.

Passel, Jeffrey S. 2005. Unauthorized Migrants: Numbers and Characteristics. Washington, DC: Pew Hispanic Center. https://www.pewresearch.org/race-and-ethnicity /2005/06/14/unauthorized-migrants/.

- ------. 2019. "Measuring Illegal Immigration: How Pew Research Center Counts Unauthorized Immigrants in the U.S." Pew Research Center Short Read. https:// www.pewresearch.org/short-reads/2019/07/12/how-pew -research-center-counts-unauthorized-immigrants-in-us/.
- Passel, Jeffrey S., and Jens Manuel Krogstad. 2024. "What We Know About Unauthorized Immigrants Living in the U.S." Pew Research Center Short Read. https://www .pewresearch.org/short-reads/2024/07/22/what-we-know -about-unauthorized-immigrants-living-in-the-us/.
- Pew Research Center. 2019. "Unauthorized Immigrant Population Trends for States, Birth Countries and Regions." https://www.pewresearch.org/race-and-ethnicity/feature /unauthorized-trends/.
- Tamborini, Christopher R., Harriet Duleep, Robert V. Gesumaria, and Dave Shoffner. 2025. "Measuring the Economic and Sociodemographic Characteristics of Unauthorized Immigrants in the United States with Survey Data." *Social Security Bulletin* 85(2): 9–15.
- Warren, Robert. 1997. Estimates of the Undocumented Immigrant Population Residing in the United States: 1996.Washington, DC: Immigration and Naturalization Service.
 - ———. 2003. Estimates of the Unauthorized Immigrant Population Residing in the United States: 1990 to 2000. Washington, DC: DHS.
 - ------. 2014. "Democratizing Data About Unauthorized Residents in the United States: Estimates and Public-Use Data, 2010 to 2013." *Journal on Migration and Human Security* 2(4): 305–328. https://journals.sagepub.com /doi/10.1177/233150241400200403.

------. 2018a. "The Estimate of 22 Million Undocumented Residents Is Vastly Overstated, For Reasons Described Here." *PLoS ONE* 13(9): e0201193. Reader comment (December 12). https://journals.plos.org/plosone/article /comment?id=10.1371/annotation/b53601d7-9b42-450a -b23c-39f4d889fcbd.

------. 2018b. "The US Undocumented Population Fell Sharply During the Obama Era: Estimates for 2016." New York, NY: Center for Migration Studies of New York. https://cmsny.org/publications /warren-undocumented-2016/.

------. 2019a. "Detailed Estimates of the Overstay Population Residing in the United States in 2017." New York, NY: Center for Migration Studies of New York. https:// cmsny.org/visa-overstay-population-warren-120219/.

------. 2019b. "US Undocumented Population Continued to Fall from 2016 to 2017 and Visa Overstays Significantly Exceeded Illegal Crossings for the Seventh Consecutive Year." *Journal on Migration and Human Security* 7(1): 19–22. https://journals.sagepub.com /doi/10.1177/2331502419830339. ——. 2021. "In 2019, the US Undocumented Population Continued a Decade-Long Decline and the Foreign-Born Population Neared Zero Growth." *Journal on Migration and Human Security* 9(1): 31–43. https://journals .sagepub.com/doi/10.1177/2331502421993746.

------. 2024. "After a Decade of Decline, the US Undocumented Population Increased by 650,000 in 2022." *Journal on Migration and Human Security* 12(2): 85–95. https://journals.sagepub.com /doi/10.1177/23315024241226624.

Warren, Robert, and Jeffrey Passel. 1987. "A Count of the Uncountable: Estimates of Undocumented Aliens Counted in the 1980 United States Census." *Demography* 24(3): 375–393. Warren, Robert, and John Robert Warren. 2013a. "A Review of the Declining Number of Visa Overstays in the U.S. from 2000 to 2009." New York, NY: Center for Migration Studies of New York. https://cmsny.org/wp -content/uploads/Re-formatted-short-paper-5.17.13.pdf.

------. 2013b. "Unauthorized Immigration to the United States: Annual Estimates and Components of Change, by State, 1990 to 2010," *International Migration Review* 47(2): 296–329. https://onlinelibrary.wiley.com/doi /abs/10.1111/imre.12022.

Measuring the Economic and Sociodemographic Characteristics of Unauthorized Immigrants in the United States with Survey Data

by Christopher R. Tamborini, Harriet Duleep, Robert V. Gesumaria, and Dave Shoffner*

In addition to estimating the number of unauthorized immigrants residing in the United States, scholars and policymakers are interested in immigrants' economic and sociodemographic characteristics. In this article, the second of three related articles, we describe methodological techniques used in estimating the population of unauthorized immigrants and their characteristics based on data from national surveys. Because large surveys typically do not ask about legal status, researchers use imputation techniques to identify potentially unauthorized immigrant populations. When applied to data from the Current Population Survey, the American Community Survey, and the Survey of Income and Program Participation, these techniques illuminate differences between likely authorized and unauthorized immigrants in such characteristics as job stability, occupation, and industry of employment. We review studies that have used such techniques. We also discuss research on the correlation between residential permanence and human capital investment among unauthorized immigrants.

Introduction

Accurate measurement of the various trends, aspects, and outcomes of unauthorized immigration is challenging, as sources of information tend to be limited or indirect. Nevertheless, survey results have been used to estimate the size of the unauthorized immigrant population, as discussed in our preceding article, and as a source of information on the characteristics of that population (Capps, Bachmeier, and Van Hook 2018). In this article, we describe some of the common methodological techniques that have been applied to survey data to estimate population characteristics-such as earnings, employment, and household composition-of unauthorized immigrants. We first discuss how imputation methods are used to identify potential unauthorized immigrants among the foreign-born residents counted in national surveys. We then summarize results from a selection of the existing literature on this topic. We follow that with a brief

discussion of the relationship between the duration of U.S. residence and human capital investment, and a concluding summary.

Imputing Unauthorized Status

Most surveys, including large national federal surveys, provide no direct measure of noncitizens' legal status. Therefore, researchers interested in examining the characteristics of the unauthorized immigrant population—and how they may differ from

Selected AbbreviationsACSAmerican Community SurveyCPSCurrent Population SurveyLPRlawful permanent residentMMPMexican Migration ProjectSIPPSurvey of Income and Program Participation

^{*} Christopher Tamborini is a researcher and Dave Shoffner is an analyst with the Social Security Administration (SSA). Harriet Duleep is a researcher with SSA; a research professor with the Public Policy Program, College of William and Mary; and a research fellow with the Institute for the Study of Labor (IZA) and with the Global Labor Organization. Robert Gesumaria is a researcher and IT specialist with SSA.

Note: Contents of this publication are not copyrighted; any items may be reprinted, but citation of the Social Security Bulletin as the source is requested. The Bulletin is available on the web at https://www.ssa.gov/policy/docs/ssb/. The findings and conclusions presented in the Bulletin are those of the authors and do not necessarily represent the views of the Social Security Administration.

those of documented noncitizen immigrants—must use indirect measures. To this end, two techniques logical imputation and statistical imputation—are often used to identify potentially unauthorized populations and to estimate the likely distribution of those populations across economic and demographic characteristics.

Logical imputation starts with identifying survey variables that are associated with authorized U.S. residence among noncitizens: for example, being a veteran or a government employee, having a certain occupational specialty, or receiving public benefits such as Medicaid coverage or Supplemental Security Income. Respondents with such characteristics are removed from the study population, and the remaining pool of potentially unauthorized immigrants is further reduced through multiple adjustments that vary depending on the research methodology. For example, some strategies randomly distribute immigrants into authorized and unauthorized pools to reflect a target benchmark based on estimates from the Department of Homeland Security or independent organizations such as the Pew Research Center. Others use information about immigrants' characteristics from the Census Bureau's Survey of Income and Program Participation (SIPP), a national longitudinal survey that follows panels of respondents over 1- to 5-year spans, with follow-up surveys administered to panel members in multiple waves. For its 1996 through 2008 panels, SIPP supplemented its core survey with a separate migration module, which asked immigrant respondents about their legal status upon arrival in the United States and whether that status had changed up to the time of the survey (Tamborini and Villarreal 2021).

A good example of logical imputation is found in Bachmeier, Van Hook, and Bean (2014). The authors use SIPP migration-module data to sort respondents into likely authorized and likely unauthorized groups. In the first step, the authors assign foreign-born individuals who report U.S. citizenship to the likely authorized group. Noncitizen immigrants who report entering the United States as lawful permanent residents (LPRs) are likewise sorted into the likely authorized group.¹ Respondents who are not U.S. citizens and did not enter as LPRs but who report changing to LPR status after arriving are also considered likely authorized (although this question last appeared in questionnaires for the 2008 panel).²

Other works that impute legal status employ "data fusion," or analyzing the characteristics of SIPP migration-module respondents and applying those distributions—as predictors of likely authorized or unauthorized status—to data for foreign-born respondents from another survey that features a much larger respondent sample, such as the Census Bureau's Current Population Survey (CPS) or American Community Survey (ACS) (Van Hook and others 2015; Capps, Bachmeier, and Van Hook 2018).^{3,4} In other words, educational, income, and other characteristics of SIPP respondents that can be gleaned from the migration module are applied to the larger CPS or ACS samples. For this technique to be valid, the variables of interest must be observed in both the SIPP's and the larger survey's samples (Van Hook and others 2015).

One potential drawback of statistical imputation from one survey to another is that it is quite complex and requires a number of additional assumptions to be made beyond logical allocation. Further, the publicuse data for the SIPP migration module reports only LPR status and excludes information on students, workers, and other noncitizens with legal temporary resident status. The 2008 SIPP panel also lacks individual variables for country of birth, which biases the country-of-origin information drawn from other SIPP panels. Moreover, although the ACS and the CPS are conducted annually, the SIPP is not;⁵ this makes the fusion of cross-survey data on the characteristics of unauthorized immigrants difficult for short or medium time periods.

Despite the depth of information on immigrant characteristics available from the SIPP migration module, its usefulness for estimating the size of the unauthorized immigrant population is limited. As noted earlier, the SIPP is not administered to a new (and expansive) sample every year like the ACS or CPS. Rather, as a medium-term longitudinal survey, SIPP follows a panel over a span of 1–5 years, with follow-up survey waves administered during that period.⁶ Theoretically, one could use the logical allocation method with the migration-module results to estimate the likely unauthorized immigrant populations for 2004 and 2008 using only the SIPP panels for those years. However, trends and changes for 2005-2007 and 2009-2011 cannot be tracked using SIPP because it is not an annual survey, and panel attrition-and differential selection out of the panel by documentation status-might introduce biases in the survey waves for the later years of the panel.

Furthermore, the SIPP migration module included a key question that was used to proxy for documentation status (specifically, whether an immigrant had changed from non-LPR to LPR status since arrival in the United States); but the migration module was eliminated from SIPP panels after 2008. Without this question, an estimated 5–10 percent of LPRs could be unidentifiable in the more recent SIPP data, resulting in an over-estimated count of unauthorized immigrants. The ACS and CPS do not contain similar questions.⁷

A potential drawback of all the survey-based methods is that they likely undercount all immigrants, particularly those who are unauthorized (Baker 2021; Passel and Cohn 2016; Passel and Krogstad 2023; Van Hook and others 2014). This might appear to be of greater concern for research that aims to estimate the size of the unauthorized immigrant population than for efforts to compare the characteristics of potentially unauthorized and likely authorized immigrants. However, unauthorized immigrants who respond to surveys are not representative of the entire immigrant population and therefore may bias the populationcharacteristics estimates of immigrants overall and by legal status (Capps, Bachmeier, and Van Hook 2018). The accuracy of measures of legal status based on logical imputation or statistical imputation relies on the accuracy of the survey-based variables used.

Selected Findings from the Literature

Using the methods described above and national survey data, a small but growing body of literature has attempted to identify the characteristics of the U.S. immigrant population by legal status. We highlight selected findings from a sample of the existing studies here; this is not an exhaustive literature review.

Hall, Greenman, and Farkas (2010) use logical imputation and SIPP migration-module data from the 1996 and 2001 panels to examine differences in working conditions across four groups: likely authorized Mexican immigrants, likely unauthorized Mexican immigrants, U.S.-born Mexican-Americans, and U.S.-born non-Hispanic White people. Their analysis suggests that among male Mexican immigrants, those who are likely unauthorized are concentrated in lowerskilled service jobs and earn 17 percent less than their likely authorized counterparts. The corresponding wage advantage for likely authorized female Mexican immigrants is 9 percent. The authors also find lower returns on human capital and slower wage growth for likely unauthorized male Mexican immigrants than for their likely authorized counterparts: the return to education for the former is half the return for the latter.

Using similar methods and more recent (2004 and 2008) SIPP panels, Greenman and Hall (2013) address variation in educational attainment among Mexican

and Central American immigrants by legal status. They find lower high school graduation and college enrollment rates among the likely unauthorized, a differential not explained by family background. Hall, Greenman, and Yi (2018) use data from the 1996, 2001, and 2004 SIPP panels to examine job mobility among likely unauthorized immigrants and find that those from Mexico and Central America have lower job mobility than likely authorized immigrants from the same areas. Moreover, when unauthorized immigrants changed jobs (either within or across firms), their rates of switching to similar jobs (rather than upward transitions) were higher than those of U.S.born workers and likely authorized immigrants.

Using logical imputation methods based on data from the SIPP core and migration modules, Tamborini and Villarreal (2021) explore differences in job stability among immigrants during the Great Recession by likely legal status and Hall, Musick, and Yi (2019) study family composition among Hispanic immigrant households. Tamborini and Villarreal find that likely unauthorized immigrants faced greater job instability, particularly underemployment, during and after the Great Recession than did legal resident immigrants. Hall, Musick, and Yi find that unauthorized Hispanic immigrants exhibited more complex living arrangements than other groups did, such as being more likely to reside with extended family and nonfamily members. Over the observation period, likely unauthorized Hispanic immigrants also experienced greater family instability (in terms of changing family size and structure) than other groups did.

Other research uses ACS data to examine differences in immigrant characteristics by legal status. Passel and Cohn (2016) examine occupation and industry of employment differences between likely unauthorized and legal immigrants using a probabilistic process to impute legal resident status for survey respondents based on age, region of birth, family relationships, and other demographic characteristics. This method extends the residual method of estimating the size of the unauthorized immigrant population, which we described in the first of these three related articles (Duleep and others 2025). Passel and Cohn find substantial within-group variation in the occupation and industry of employment among immigrants by legal status. Likely unauthorized immigrants tend to be employed in low-skilled occupations characterized by informal and nonstandard work arrangements, including landscaping, foodservice, and hospitality. They are also concentrated in construction and farmwork:

Passel and Cohn estimate that in 2014, unauthorized immigrants constituted 15 percent of workers in construction and 26 percent of those in farming.

Borjas (2017a and 2017b) uses data files constructed by Pew Research Center analysts to examine the labor market characteristics of immigrants by legal status based on ACS and CPS data. Using a variant of a probabilistic logical imputation method described in Passel and Cohn (2014), Borjas also finds labor market differences between unauthorized immigrants, authorized immigrants, and the U.S.-born population. Consistent with Hall, Greenman, and Farkas (2010) findings using the SIPP, Borjas (2017a) observes that wages and returns on education are lower for likely unauthorized immigrants than for authorized immigrants and U.S.-born workers, with legal status associated with wages that are between 6 percent and 14 percent higher. Borjas (2017b) also finds substantially higher labor force participation and employment rates among likely unauthorized immigrant men than the likely authorized group (as does Albert 2021). By contrast, among women, unauthorized immigrants experienced substantially lower employment rates than their likely authorized counterparts. Bean, Brown, and Bachmeier (2015) impute the legal status of Mexican immigrants using the 2012 ACS and find lower earnings among likely unauthorized men.

Permanence of U.S. Residence and Investment in Human Capital

In estimating the number of unauthorized immigrants and in understanding their characteristics, it is important to differentiate between those who stay in the United States permanently and those with temporary U.S. residency.⁸ Unauthorized immigrants are more likely to return to their countries of origin than authorized immigrants are (Sohn and others 2023).

The Mexican Migration Project (MMP), a joint Mexican and American interdisciplinary research effort established in 1982, gathers information on migrants who are—at least initially—relatively transient. The MMP conducts interviews in the winter months, when many migrants return to their home country to join their families. Out-migrant samples are also taken, matching those communities with migrants residing in the United States.⁹ The MMP data reveal a population that mostly lacks U.S. legal status, whose members transit back and forth between the United States and Mexico, and who generally experience low U.S. earnings growth. In another MMP study, Massey (1986) probes the role of permanence and finds that migrants form social and economic ties as they accrue time in the United States, which increases the chances that they will attempt to settle permanently. With time, migrants bring their family members and, with greater permanence, they secure more stable, better paying jobs. These data have also been used to analyze the role of economic and social factors in the decision to attempt unauthorized migration (Ryo 2013).

Using data collected by the China International Migration Project, Chunyu (2011) traces the work trajectories of immigrants from China's Fujian province, the source of the largest wave of Chinese emigration in the 1990s. Like their Mexican counterparts, these immigrants are mostly unauthorized, with low levels of education: 41 percent possess no more than an elementary-school education. Yet, in contrast with the Mexican unauthorized immigrants, few return to China.

A window on the effect of permanence within the more generally transient Mexican unauthorized population is opened by examining individuals who applied for legal status under the 1986 Immigration Reform and Control Act (IRCA). Under IRCA, 1.7 million persons were legalized by 1990, 1.3 million of whom were Mexican. Individuals could attain legal status if they could show "long-term" U.S. residence.¹⁰ Thus, those who applied for legal status are a relatively permanent subset of the unauthorized population.

From IRCA's processing system, the Legalized Population Survey (LPS) data file was created, with information on the jobs and earnings of these individuals at three points in time-when they first entered the United States, when they sought legal permanent residence, and several years thereafter. Using the 1989 LPS, Powers and Seltzer (1998) find that real median earnings rose 21 percent for unauthorized immigrant men between their initial U.S. job and the time they applied for legal status. Using a scale that reflects the relative economic status indicated by detailed occupations, Powers and Seltzer also find meaningful earnings gains for the study population in the period before they attained legal status.¹¹ The study results suggest that within a population generally characterized by impermanence and low earnings mobility, earnings growth exists for those who reside for longer periods in the United States.

Using data from various university and local government surveys conducted in southern California,

Cornelius and Marcelli (2000) find that permanent settlement of Mexican migrants in the United States began to increase in the 1970s and accelerated during the 1980s. Permanence thus varies across groups of unauthorized immigrants as well as over time for the same group: A historically transient group may begin to trend toward more permanence, and permanence affects certain characteristics—such as earnings—of unauthorized immigrant groups in ways that affect the accuracy of unauthorized immigrant population estimates.

Conclusion

Scholars and policymakers explore the extent to which the economic and social characteristics of immigrants vary by legal status. This article describes some methodological strategies that have been employed to identify the characteristics of potential unauthorized immigrants using data from national surveys. We have focused on methods of imputing legal status and then have summarized results from selected studies that follow such strategies. Future work would benefit from developing more precise measures, if possible, of immigrants' legal status. Perhaps current methods that essentially impute legal status could be combined with information indicating whether the survey data can be matched to administrative data records. The third of our three related articles (Gesumaria and others 2025) examines such potential survey-administrative data linkages.

Notes

Acknowledgments: We conducted this study under the direction and with the support of Mark J. Warshawsky. We thank Stephanie Myers, Anya Olsen, Mark Sarney, Steve Goss, Tokunbo Oluwole, Ben Danforth, Kent O. Morgan, Steve Robinson, Michael Morris, and Gayle Reznik. We would especially like to thank Mark Regets (Senior Fellow, National Foundation for American Policy, former immigration expert with the National Science Foundation), Robert Warren (Senior Visiting Fellow at the Center for Migration Studies of New York, who served as a demographer for 34 years with the Census Bureau and the former Immigration and Naturalization Service), and Ben Pitkin (editor, *Social Security Bulletin*).

¹ To protect respondent confidentiality, the SIPP groups all immigrants who were non-LPRs at U.S. entry under a single "other" category. In addition to unauthorized immigrants, that category includes workers, refugees, asylees, tourists, business travelers, and diplomats and other political representatives with legal temporary resident status. ² Imputed values for all questions used to infer documentation status are not considered in the assignment of likely status.

³ The CPS is conducted by the Census Bureau for the Bureau of Labor Statistics.

⁴ Keister and Aronson (2017) also use data fusion, applying characteristics found in SIPP data to results from the Federal Reserve Board's Survey of Consumer Finances.

⁵ For some SIPP panels, particularly newer panels, follow-up waves are conducted annually; but these yearly reiterations are not equivalent to fielding the survey to new participants each year, as the ACS and CPS do.

⁶ SIPP waves are fielded at regular intervals that vary from panel to panel.

⁷ The questions about migration that remain in the more recent SIPP panels are less detailed than the previous versions were. For example, the 2014 and later SIPP panels do not ask whether non-LPR arrivals had subsequently attained LPR status.

⁸ Duleep and Regets (1994, 1999, and 2002) and Duleep and others (2020) explore the role of permanence in human capital investment among immigrants and model its effects.

⁹ See Massey and Zenteno (1999) for further information. The collected data, compiled in a comprehensive database, has formed the foundation for Orrenius and Zavodny (2003), Donato, Durand, and Massey (1992), and numerous other studies.

¹⁰ For applicants granted specialized agricultural worker status, the requirements for legalization were much more lenient (only 90 days of continuous agricultural employment in the past year). Comparing their experiences with those of long-term U.S. resident immigrants thereby provided a potential natural experiment on the effect of permanence.

¹¹ The initial occupational status scores of this population placed the unauthorized immigrants in the lowest onefifth of all U.S. occupations. By the time they applied for legal status, these immigrants were no longer in the lowest occupational-status quintile.

References

- Albert, Cristoph. 2021. "The Labor Market Impact of Immigration: Job Creation Versus Job Competition." *American Economic Journal: Macroeconomics* 13(1): 35–78.
- Bachmeier, James D., Jennifer Van Hook, and Frank D. Bean. 2014. "Can We Measure Immigrants' Legal Status? Lessons from Two U.S. Surveys." *International Migration Review* 48(2): 538–566. https://doi.org/10.1111 /imre.12059.

Baker, Bryan. 2021. Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2015–January 2018. Washington, DC: Department of Homeland Security, Office of Immigration Statistics. https://ohss.dhs.gov/topics/immigration/unauthorized /population-estimates.

Bean, Frank D., Susan K. Brown, and James D. Bachmeier. 2015. Parents Without Papers: The Progress and Pitfalls of Mexican American Integration. New York, NY: Russell Sage Foundation.

Borjas, George J. 2017a. "The Earnings of Undocumented Immigrants." NBER Working Paper No. 23236. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w23236.

------. 2017b. "The Labor Supply of Undocumented Immigrants." *Labour Economics* 46: 1–13. https://doi .org/10.1016/j.labeco.2017.02.004.

Capps, Randy, James D. Bachmeier, and Jennifer Van Hook. 2018. "Estimating the Characteristics of Unauthorized Immigrants Using U.S. Census Data: Combined Sample Multiple Imputation." *The ANNALS of the American Academy of Political and Social Science* 677(1): 165–179. https://doi.org/10.1177/0002716218767383.

Chunyu, Miao David. 2011. "Earnings Growth Patterns of Chinese Labor Immigrants in the United States." Paper prepared for the 2011 Annual Meeting of the Population Association of America, Washington, DC.

Cornelius, Wayne A., and Enrico A. Marcelli. 2000. "The Changing Profile of Mexican Migrants to the United States: New Evidence from California and Mexico." IZA Discussion Paper No. 220. Bonn: Institute for the Study of Labor (IZA).

Donato, Katharine M., Jorge Durand, and Douglas S. Massey. 1992. "Changing Conditions in the US Labor Market: Effects of the Immigration Reform and Control Act of 1986." *Population Research and Policy Review* 11(2): 93–115.

Duleep, Harriet Orcutt, and Mark C. Regets. 1994. "The Elusive Concept of Immigrant Quality." Program for Research on Immigration Policy Discussion Paper No. PRIP-UI-28. Washington, DC: The Urban Institute.

------. 1999. "Immigrants and Human-Capital Investment." *American Economic Review* 89(2): 186–191. https://doi .org/10.1257/aer.89.2.186.

------. 2002. "The Elusive Concept of Immigrant Quality: Evidence from 1970–1990." IZA Discussion Paper No. 631. Bonn: Institute for the Study of Labor (IZA). https://www.iza.org/publications/dp/631/the-elusive -concept-of-immigrant-quality-evidence-from-1970-1990.

Duleep, Harriet, Mark C. Regets, Seth Sanders, and Phanindra V. Wunnava. 2020. *Human Capital Investment: A History of Asian Immigrants and Their Family* *Ties*. Cham, Switzerland: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-47083-8.

Duleep, Harriet, Dave Shoffner, Robert V. Gesumaria, and Christopher R. Tamborini. 2025. "Measuring the Number of Unauthorized Immigrants in the United States: A Review of the Residual Estimation Method." *Social Security Bulletin* 85(2): 1–8.

Gesumaria, Robert V., Harriet Duleep, Christopher R. Tamborini, and Dave Shoffner. 2025. "A New Way to Estimate the Number of Unauthorized Immigrants in the United States." *Social Security Bulletin* 85(2): 17–23.

Greenman, Emily, and Matthew Hall. 2013. "Legal Status and Educational Transitions for Mexican and Central American Immigrant Youth." *Social Forces* 91(4): 1475–1498. https://doi.org/10.1093/sf/sot040.

Hall, Matthew, Emily Greenman, and George Farkas. 2010. "Legal Status and Wage Disparities for Mexican Immigrants." *Social Forces* 89(2): 491–513. https://doi .org/10.1353/sof.2010.0082.

Hall, Matthew, Emily Greenman, and Youngmin Yi. 2018. "Job Mobility Among Unauthorized Immigrant Workers." *Social Forces* 97(3): 999–1028. https://doi .org/10.1093/sf/soy086.

Hall, Matthew, Kelly Musick, and Youngmin Yi. 2019."Living Arrangements and Household Complexity Among Undocumented Immigrants." *Population and Development Review* 45(1): 81–101.

Keister, Lisa A., and Brian Aronson. 2017. "Immigrants in the One Percent: The National Origin of Top Wealth Owners." *PLoS One*. 12(2). https://doi.org/10.1371 /journal.pone.0172876.

Massey, Douglas S. 1986. "The Settlement Process Among Mexican Migrants to the United States." *American Sociological Review* 51(5): 670–684.

Massey, Douglas S., and Rene M. Zenteno. 1999. "The Dynamics of Mass Migration." *Proceedings of the National Academy of Sciences of the United States of America* 96(9): 5328–5335.

Orrenius, Pia M., and Madeline Zavodny, 2003. "Do Amnesty Programs Reduce Undocumented Immigration? Evidence from IRCA." *Demography* 40(3): 437–450. https://doi.org/10.2307/1515154.

Passel, Jeffrey S., and D'Vera Cohn. 2014. Unauthorized Immigrant Totals Rise in 7 States, Fall in 14: Decline in Those from Mexico Fuels Most State Decreases. Washington, DC: Pew Research Center. https://www .pewresearch.org/race-and-ethnicity/2014/11/18 /unauthorized-immigrant-totals-rise-in-7-states-fall-in-14/.

------. 2016. Size of U.S. Unauthorized Immigrant Workforce Stable After the Great Recession. Washington, DC: Pew Research Center. https://www.pewresearch.org /race-and-ethnicity/2016/11/03/size-of-u-s-unauthorized -immigrant-workforce-stable-after-the-great-recession/.

- Passel, Jeffrey S., and Jens Manuel Krogstad. 2023. "What We Know About Unauthorized Immigrants Living in the U.S." Pew Research Center Short Read. https://www .pewresearch.org/short-reads/2024/07/22/what-we-know -about-unauthorized-immigrants-living-in-the-us/.
- Powers, Mary G., and William Seltzer. 1998. "Occupational Status and Mobility Among Undocumented Immigrants by Gender." *International Migration Review* 32(1): 21–56.
- Ryo, Emily. 2013. "Deciding to Cross: Norms and Economics of Unauthorized Migration." *American Sociological Review* 78(4): 574–603.
- Sohn, Heeju, Anne R. Pebley, Amanda Landrian Gonzalez, and Noreen Goldman. 2023. "Deportations and Departures: Undocumented Mexican Immigrants' Return Migration During Three Presidential Administrations." *Proceedings of the National Academy of Sciences* 120(9): e2212184120. https://doi.org/10.1073/pnas.2212184120.

- Tamborini, Christopher R., and Andrés Villarreal. 2021. "Immigrants' Employment Stability Over the Great Recession and Its Aftermath." *Demography* 58(5): 1867–1895. https://doi.org/10.1215/00703370-9429459.
- Van Hook, Jennifer, James D. Bachmeier, Donna L. Coffman, and Ofer Harel. 2015. "Can We Spin Straw Into Gold? An Evaluation of Immigrant Legal Status Imputation Approaches." *Demography* 52(1): 329–354. https:// doi.org/10.1007/s13524-014-0358-x.
- Van Hook, Jennifer, Frank D. Bean, James D. Bachmeier, and Catherine Tucker. 2014. "Recent Trends in Coverage of the Mexican-Born Population of the United States: Results from Applying Multiple Methods Across Time." *Demography* 51(2): 699–726. https://doi.org/10.1007 /s13524-014-0280-2.

A New Way to Estimate the Number of Unauthorized Immigrants in the United States

by Robert V. Gesumaria, Harriet Duleep, Christopher R. Tamborini, and Dave Shoffner*

This article introduces a new method of estimating the number of unauthorized immigrants in the United States by exploiting discrepancies between Current Population Survey (CPS) data and Social Security administrative data on Social Security numbers (SSNs). Potential unauthorized immigrant status is indicated when the SSNs reported by CPS respondents and the SSNs recorded in linked administrative data do not match. We use nonmatching SSN data to identify likely unauthorized immigrants and apply a series of logical adjustments to refine the estimated population counts. The resulting estimates are consistent with those calculated using the residual estimation method, which we described in the first of this group of three related articles. Because the residual method and this new method take entirely different approaches, the similarity of their results is mutually reinforcing.

Introduction

This article introduces a new method of estimating the unauthorized immigrant population in the United States by exploiting discrepancies between data from the Census Bureau's Current Population Survey (CPS) and administrative records maintained by the Social Security Administration (SSA). In some cases, the Social Security numbers (SSNs) reported by survey respondents do not match the SSNs in the administrative data. Among foreign-born individuals, these SSN nonmatches often indicate unauthorized status. Our new *CPS-SSA nonmatch* estimation method separately addresses the two causes of nonmatching SSNs: individuals without a valid SSN (Type 1 nonmatches) and individuals who use a valid SSN that belongs to another person (Type 2 nonmatches).

The article comprises six sections, beginning with this introduction. The second section presents background information on the CPS-SSA data linkage. The third and fourth sections respectively describe the circumstances that lead to Type 1 and Type 2 SSN nonmatches, and our method of using each type of mismatched data as a step in the process of estimating the unauthorized immigrant population. The fifth section offers information for analysts planning to use or improve the CPS-SSA nonmatch method. The sixth section concludes by comparing the numbers of unauthorized immigrants estimated using our CPS-SSA nonmatch method with those estimated using the residual method, which the first of these three related articles describes (Duleep and others 2025).

Background

Matching federal survey data with administrative data records is a critical research tool, and the SSN plays a crucial role in matching CPS data with SSA data

Selected Abbreviations			
CPS	Current Population Survey		
DHS	Department of Homeland Security		
SSA	Social Security Administration		
SSN	Social Security number		

^{*} Robert Gesumaria is a researcher and IT specialist with the Social Security Administration (SSA). Harriet Duleep is a researcher with SSA; a research professor with the Public Policy Program, College of William and Mary; and a research fellow with the Institute for the Study of Labor (IZA) and with the Global Labor Organization. Christopher Tamborini is a researcher and Dave Shoffner is an analyst, also with SSA.

Note: Contents of this publication are not copyrighted; any items may be reprinted, but citation of the Social Security Bulletin as the source is requested. The Bulletin is available on the web at https://www.ssa.gov/policy/docs/ssb/. The findings and conclusions presented in the Bulletin are those of the authors and do not necessarily represent the views of the Social Security Administration.

files (Aziz, Kilss, and Scheuren 1978; Delbene 1979; Duleep 1986). Nevertheless, over time, CPS respondents have become increasingly reluctant to provide their SSN and thereby enable the administrative-data linkages. To overcome this obstacle, in the early 2000s,

the Census Bureau stopped directly requesting an SSN. Instead, under a new methodology, a respondent is informed that the survey data will be matched with other federal data for research purposes. Unless the respondent opts out, the Census Bureau then combines SSN application information from SSA's [Numerical Identification System data] file with address records from the [Internal Revenue Service], SSA, and other sources to determine the respondent's correct SSN. Once a match is found, survey and administrative data for the respondent are linked (McNabb and others 2009).

Importantly, SSNs are not disclosed in any data set used for research. Instead, to protect the individuals' identifies, they are replaced with coded proxy identifiers.¹

Methods

To estimate the size of the unauthorized immigrant population, we examine linked CPS and SSA data for foreign-born individuals and quantify the instances in which the reported SSNs do not match. A Type 1 nonmatch occurs if there is no valid SSN: The respondent either has no SSN at all or has a fraudulent SSN, meaning that it was fabricated by the respondent or by his or her employer. A Type 2 mismatch occurs if the respondent has a valid SSN, but that SSN legitimately belongs to another person. For this analysis, we specifically use SSA's Numerical Identification System (Numident) file and CPS Annual Social and Economic Supplement (ASEC) data for 2006 through 2016. We then compare the prevalence of CPS-SSA data nonmatches for immigrants with the nonmatch rate of U.S.-born respondents, as detailed below.

Type 1 Nonmatches

The CPS results contain data for individuals with Type 1 nonmatches (no SSN or a nonvalid SSN), but the Social Security administrative record system does not. The survey and administrative data for these individuals therefore cannot match. Table 1 shows the number and prevalence of these nonmatches among persons aged 15 or older. The number of foreign-born nonmatches in Table 1 might appear to provide logical estimates of the size of the unauthorized immigrant population. Yet SSN reporting errors can cause mismatches between the CPS data and SSA records that have nothing to do with unauthorized immigration. Because there are no unauthorized immigrants among the native-born population, the nonmatch rate for that group provides a control for estimating the shares of nonmatches that occur for reasons other than unauthorized immigration.

To estimate the percentages of foreign-born nonmatches that are due to unauthorized immigration, we subtract the nonmatch rate among the native-born respondents from the nonmatch rate for all immigrants. We then compute the unauthorized immigrant population by multiplying the total foreign-born population by the resulting percentage, as shown in Table 1.

One concern arises: The match probabilities may correlate with demographic and socioeconomic characteristics, and the distribution of these variables may differ between the foreign-born and U.S.-born populations. To address this concern, we reweighted the native-born sample to align with the foreign-born sample by age, sex, and education. We omit these results, however, because reweighting the sample on these characteristics barely changed the estimated unauthorized immigrant population.

Type 2 Nonmatches

When a CPS respondent is an unauthorized immigrant who uses or has used someone else's valid SSN, the SSN will appear in both the CPS and the SSA data. To estimate the number of unauthorized immigrants who use, or have used, another person's valid SSN, we distinguish between two types of SSN matches in the CPS and SSA data. In an affirmative match, the individual's CPS data and the SSA data match on key variables, such as sex and birth year. A dubious match occurs when the individual's CPS and SSA data do not match on key variables. Thus, we define a match as dubious if the individual's CPS data and SSA records differ either in sex or in age (if by more than 5 years). The count of dubious matches among the foreign-born may provide a good estimate of the number who are using someone else's valid SSN. We refer to the number of dubious matches divided by the total number of all matches (affirmative plus dubious) in the CPS and SSA data as the discrepancy rate. As we did with Type 1 nonmatches, we use the U.S.-born population as a control to account for Type 2 discrepancies that are caused by reasons

Table 1.

Estimating the unauthorized immigrant population aged 15 or older: Foreign-born U.S. population, and adjustments accounting for Type 1 SSN nonmatches, 2006–2016

	Foreign-borr	population	Type 1 SSN nonmatches		Estimated percentage of	Estimated unauthorized	
			Foreign-born		U.Sborn (as a	the foreign-born	immigrant
				Share of	percentage	population	population
		Share of U.S.		foreign-born	of total	who are	based
Survey		population		population	U.Sborn	unauthorized	on Type 1
year	Number	(%)	Number	(%)	population)	immigrants ^a	nonmatches ^b
2006	33,571,249	14.40	9,924,156	29.56	8.54	21.02	7,055,651
2007	35,063,411	14.86	10,242,443	29.21	8.36	20.85	7,311,041
2008	35,180,322	14.77	10,547,283	29.98	9.30	20.68	7,276,442
2009	34,884,933	14.53	10,280,222	29.47	9.36	20.11	7,016,271
2010	35,682,735	14.73	8,117,290	22.75	9.53	13.22	4,716,841
2011	36,479,785	14.95	6,812,580	18.67	8.59	10.08	3,679,485
2012	38,195,263	15.42	6,979,920	18.27	9.35	8.92	3,407,508
2013	38,517,423	15.41	8,872,147	23.03	9.98	13.05	5,027,376
2014	39,212,327	15.53	8,969,626	22.87	10.20	12.67	4,971,488
2015	40,556,084	15.89	9,559,729	23.57	10.48	13.09	5,311,430
2016	41,346,254	16.03	9,879,161	23.89	10.78	13.11	5,420,932

SOURCE: Authors' calculations based on CPS-ASEC and Social Security administrative data.

NOTE: A CPS respondent with no SSN or a fabricated SSN is a Type 1 SSN nonmatch.

a. Equals the nonmatch share of foreign-born population minus the nonmatch share of U.S.-born population.

b. Equals the foreign-born population times the estimated percentage who are unauthorized.

other than unauthorized immigration. We subtract the discrepancy rate for the U.S.-born population from the discrepancy rate for the foreign-born population to determine the *net* discrepancy rate.

We apply the net discrepancy rate to the number of all immigrants who have CPS-SSA data matches (whether affirmative or dubious). The resulting number is added to the unauthorized immigrant population that was estimated based on Type 1 nonmatches (Table 2).

Comparing CPS-SSA Nonmatch Method and Residual Method Estimates

As the first of our three articles discusses, the residual method of estimating the unauthorized immigrant population includes steps that account for survey undercounts and for immigrants who entered the United States with legal temporary visas but then overstayed them.

Adjusting for Undercounts

Researchers account for American Community Survey and CPS undercounts by adjusting their estimated counts of unauthorized immigrants upward by

Social Security Bulletin, Vol. 85, No. 2, 2025

5 percent to 15 percent. We use 10 percent, the midpoint of those adjustments, to offset undercounting in the CPS. Specifically, we assume that the unauthorized immigrant population figures based on nonmatching SSNs in Table 2 represent 90.9 percent of the true population (that is, the population accounting for CPS undercounts). Table 3 shows the figures adjusted to equal 100 percent of those counts.

Accounting for Visa Overstays

B1 tourist visas and B2 business trip visas account for about 92 percent of visa overstays (Department of Homeland Security [DHS] 2021). Most people who overstay tourist or business visas are unlikely to have an SSN history. Yet, many visitors holding other types of visas are eligible for temporary employment authorization. This suggests that we could subtract the number of individuals who overstayed a B1 or B2 visa from the total number of visa overstays to estimate the number of visa overstays with an SSN history. Unfortunately, estimates of the population of visa overstays are not available; however, estimated net annual flows in visa overstays are. Warren (2019) estimates that 46 percent of the unauthorized immigrant population in 2017 overstayed a visa and DHS (2021) reports that

Table 2.

Estimating the unauthorized immigrant population aged 15 or older: Adding Type 2 SSN nonmatches to the Type 1 nonmatch analysis results, 2006–2016

Survey year	Estimated unauthorized immigrant population based on Type 1 SSN nonmatches only (from Table 1)	Net discrepancy rate ^a (%)	Additional unauthorized immigrants based on Type 2 SSN nonmatch analysis	Total estimated unauthorized immigrant population based on nonmatching SSNs
2006	7,055,651	3.85	271,942	7,327,593
2007	7,311,041	4.72	345,011	7,656,052
2008	7,276,442	4.23	307,913	7,584,355
2009	7,016,271	3.61	253,429	7,269,700
2010	4,716,841	63.64	3,001,877	7,718,718
2011	3,679,485	90.79	3,340,527	7,020,012
2012	3,407,508	104.34	3,555,428	6,962,936
2013	5,027,376	35.14	1,766,858	6,794,234
2014	4,971,488	31.51	1,566,572	6,538,060
2015	5,311,430	29.24	1,552,917	6,864,347
2016	5,420,932	31.29	1,696,076	7,117,008

SOURCE: Authors' calculations based on CPS-ASEC and Social Security administrative data.

NOTES: A CPS respondent with no SSN or a fabricated SSN is a Type 1 SSN nonmatch.

A CPS respondent with a valid SSN that legitimately belongs to another person is a Type 2 SSN nonmatch.

a. The difference in SSN-match discrepancy rates between U.S.-born and foreign-born CPS respondents.

Table 3.

Adjusting the estimated unauthorized immigrant population: Survey undercount, visa overstays, and expanding the analyzed population from ages 15 or older to all ages, 2006–2016

	Total actimated	Adjusting to account for—				
	unauthorized immigrant population based on		Visa overstays not captured in nonmatching	The all-ages unauthorized immigrant population if the share of the population aged 0–14 equals—		
Survey year	nonmatching SSNs (from Table 2)	CPS undercount: Add 10%	SSN analysis: Add another 5%	6.9% ^a	20% ^b	
2006	7,327,593	8,060,352	8,463,370	9.047.343	10,156,044	
2007	7.656.052	8.421.657	8.842.740	9.452.889	10.611.288	
2008	7,584,355	8,342,791	8,759,931	9,364,366	10,511,917	
2009	7,269,700	7,996,670	8,396,504	8,975,863	10,075,805	
2010	7,718,718	8,490,590	8,915,120	9,530,263	10,698,144	
2011	7,020,012	7,722,013	8,108,114	8,667,574	9,729,737	
2012	6,962,936	7,659,230	8,042,192	8,597,103	9,650,630	
2013	6,794,234	7,473,657	7,847,340	8,388,806	9,416,808	
2014	6,538,060	7,191,866	7,551,459	8,072,510	9,061,751	
2015	6,864,347	7,550,782	7,928,321	8,475,375	9,513,985	
2016	7,117,008	7,828,709	8,220,144	8,787,334	9,864,173	

SOURCE: Authors' calculations based on CPS-ASEC and Social Security administrative data.

a. Assumes the share of the population aged 0–14 is lower among unauthorized immigrants than in the overall U.S. population.

b. Assumes the share of the population aged 0–14 is similar between unauthorized immigrants and the overall U.S. population.

in 2019, about 8 percent of overstays held nonbusiness or nontourist visas. Eight percent of 46 percent is about 3.7 percent. To reduce the risk of underestimating overstay incidence, we adjust 3.7 percent up to 5 percent, then add that 5 percent to the estimated population of unauthorized immigrants in Table 3.

Estimates for All Ages

Our estimates are calculated for the unauthorized immigrant population aged 15 or older but the residual method estimates, discussed in the first of these three articles, are calculated for the all-ages population. The final step of our estimation method is to reconcile that difference. For the period 2006-2016, about 20 percent of the U.S. population was aged 14 or younger (Census Bureau 2023). Yet the motivations and the logistics of undocumented immigration are likely to result in a disproportionally low presence of children younger than 15 in the unauthorized immigrant population. Among foreign-born U.S. residents who arrived in the period 1982-2019, the 2019 American Community Survey found that 6.9 percent were younger than 15. Table 3 therefore shows our computations with both 6.9 percent and 20 percent adjustments to provide alternative estimates of the all-ages unauthorized immigrant population.

Data Limitations and Notes for Future Research

We likely underestimate the percentage of unauthorized immigrants who overstay their visas and have a valid SSN because the estimates are based on flow data rather than on "snapshot" data for entire populations at particular points in time. If available, snapshot data should be used to inform these estimates.

Our estimates ignore individuals who overstayed a visa but now reside outside the United States and assume that individuals overstaying a B1 or B2 business or tourist visa do not have any administrative records at SSA. We are not certain whether our algorithm counts holders of F1 visas, who are eligible for Optional Practical Training (which can last from 6 months to 27 months), as authorized or unauthorized immigrants.

The Census Bureau's Person Identification Validation System (PVS) matches survey responses with SSA data without disclosing SSNs. The PVS uses probabilistic matching to assign a unique Census Bureau identifier for each person (Wagner and Layne 2014). Analogous to data fingerprints, the unique non-SSN identifying information that the PVS uses will not find matches in SSA data for persons who have never applied for and received SSNs. Because these persons have never given their identifying information to SSA or the Internal Revenue Service, they have no data in the administrative records. Thus, the PVS allows us to infer that immigrant survey respondents who have no matching SSA data do not have a valid SSN, suggesting that they may be unauthorized immigrants.

Our methodology focuses on the number of unauthorized immigrants and not their characteristics, which we explore in the second of our three articles (Tamborini and others 2025). Subject to further investigation, the CPS-SSA nonmatch method may provide a convenient way to continuously measure both the size and characteristics of the U.S. unauthorized immigrant population.

Summary and Conclusion

Each year, the SSA actuaries forecast the financial status of the Old-Age, Survivors, and Disability Insurance programs by projecting U.S. labor force participation, earnings, and other variables. These long-term projections incorporate assumptions about the relationship between immigration and Social Security. In describing the unauthorized immigrant population and presenting methods for estimating its size, our three articles may provide insights to inform those assumptions.

To date, two estimation methodologies have dominated efforts to measure the number of unauthorized immigrants in the United States. The first, the residual method, is described in detail in the first of these three articles (Duleep and others 2025). It involves subtracting from the count of all foreign-born individuals residing in the United States the numbers with legalresident status. The results represent an estimate of the unauthorized immigrant population.

The second approach uses enforcement statistics such as border apprehensions. An attractive feature of this approach is that it starts with known information about who we are trying to measure—unauthorized immigrants. Yet a single person may cross the border and return multiple times. If each apprehension is counted as a new entrant, then this method overestimates the number of unauthorized immigrants. The number of border-crossing agents will also affect how many unauthorized immigrants are counted: with more agents, more apprehensions occur and are counted. Given these shortcomings, the enforcementstatistics estimation method is not used as often as the residual method, which is preferred by DHS and various research institutes.

Consistency of results implies accuracy, and studies that use the residual method find similar results. Skeptics note, however, that the accuracy of the residual method estimates are difficult to verify, given that they share a similar methodology. Perhaps all are consistently wrong? Would a valid but different methodology find similar results?

Motivating our study was a concern that the residual method may dramatically understate unauthorized immigration. To explore this concern, we developed an alternative estimation method. We employ a unique, restricted-use dataset linking data for respondents from multiple years of the CPS to their administrative records compiled at SSA. The CPS-SSA nonmatch method counts two types of unauthorized immigrants: those who do not have a valid SSN and those who use the valid SSN of another person. The CPS-SSA nonmatch method differs completely from the residual method. If our estimates of the unauthorized immigrant population are similar to residual method estimates, it cannot be due to methodologic similarities.

As discussed in the first of our three articles, the Center for Migration Studies of New York (CMS) has used its own version of residual techniques to produce annual estimates of the unauthorized immigrant population from 2010 to 2019, providing greater detail than the DHS estimates. Nevertheless, CMS estimated the total unauthorized immigrant population for 2010 at 11.7 million (Warren and Warren 2013), only slightly more than DHS' estimate of 11.6 million (Baker 2021). The results for our CPS-SSA nonmatch method and from CMS and DHS using the residual method are similar: We estimate an unauthorized immigrant population of 10.7 million in 2010 (Table 3, using the 20 percent adjustment to expand the counted population from those aged 15 or older to those of all ages).

Estimates of the unauthorized immigrant population over time using the residual and CPS-SSA nonmatch methods are also broadly similar. Table 3 shows that the CPS-SSA nonmatch estimates of the number of unauthorized immigrants peaked in 2010, followed by 4 consecutive years of small decreases. The numbers then increased in 2015 and 2016. Similarly, following 2010, CMS estimates using a modified residual method show several years of declining unauthorized immigration, until 2022 when the estimated number of unauthorized immigrants increased 6 percent (Warren 2024). In conclusion, our different methodology produces estimates of the size of the unauthorized immigrant population in the United States—and of unauthorized immigration trends—that are broadly similar to those produced using the residual method. We find no evidence that the residual method underestimates unauthorized immigration. The similarity in results is important both for national policy discussions about unauthorized immigrants in the United States and for the specific policy needs of Social Security.

Notes

Acknowledgments: We conducted this study under the direction and with the support of Mark J. Warshawsky. We thank Stephanie Myers, Anya Olsen, Mark Sarney, Steve Goss, Tokunbo Oluwole, Ben Danforth, Kent O. Morgan, Steve Robinson, Michael Morris, and Gayle Reznik. We would especially like to thank Mark Regets (Senior Fellow, National Foundation for American Policy, former immigration expert with the National Science Foundation), Robert Warren (Senior Visiting Fellow at the Center for Migration Studies of New York, who served as a demographer for 34 years with the Census Bureau and the former Immigration and Naturalization Service), and Ben Pitkin (editor, *Social Security Bulletin*).

¹ For more information on Census Bureau matching procedures, see Wagner and Layne (2014).

References

- Aziz, Faye, Beth Kilss, and Fritz Scheuren. 1978. 1973 Current Population Survey: Administrative Record Exact Match File Codebook, Part I—Code Counts and Item Definitions. Studies from Interagency Data Linkages, Report No. 8. DHEW Publication No. (SSA) 79-11750. Washington, DC: Department of Health, Education, and Welfare, SSA, Office of Research and Statistics.
- Baker, Bryan, 2021. Estimates of the Unauthorized Immigration Population Residing in the United States: January 2015–January 2018. Washington, DC: Department of Homeland Security, Office of Immigration Statistics. https://ohss.dhs.gov/topics/immigration/unauthorized /population-estimates.
- Census Bureau. 2023. "National Population by Characteristics: 2010–2019. Median Age and Age by Sex, Race, and Hispanic Origin, Annual Estimates of the Resident Population by Sex, Age, Race, and Hispanic Origin for the United States: April 1, 2010 to July 1, 2019." https://www.census.gov/data/tables/time-series/demo /popest/2010s-national-detail.html.
- Delbene, Linda. 1979. 1937–1976 Social Security Longitudinal Earnings Exact Match File. Studies from Interagency Data Linkages, Report No. 9. Washington, DC:

Department of Health, Education, and Welfare, SSA, Office of Research and Statistics.

- [DHS] Department of Homeland Security. 2021. *Fiscal Year 2020 Entry/Exit Overstay Report*. Washington, DC: DHS, U.S. Customs and Border Protection. https://www .dhs.gov/publication/entryexit-overstay-report.
- Duleep, Harriet Orcutt. 1986. "Measuring the Effect of Income on Adult Mortality Using Longitudinal Administrative Record Data." *The Journal of Human Resources* 21(2): 238–251.
- Duleep, Harriet, Dave Shoffner, Robert V. Gesumaria, and Christopher R. Tamborini. 2025. "Measuring the Number of Unauthorized Immigrants in the United States: A Review of the Residual Estimation Method." *Social Security Bulletin* 85(2): 1–8.
- McNabb, Jennifer, David Timmons, Jae Song, and Carolyn Puckett. 2009. "Uses of Administrative Data at the Social Security Administration." *Social Security Bulletin* 69(1): 75–84.
- Tamborini, Christopher R., Harriet Duleep, Robert V. Gesumaria, and Dave Shoffner. 2025. "Measuring the Economic and Sociodemographic Characteristics of Unauthorized Immigrants in the United States with Survey Data." *Social Security Bulletin* 85(2): 9–15.

- Wagner, Deborah, and Mary Layne. 2014. "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software." CARRA Working Paper No. 2014-01. Washington, DC: Census Bureau.
- Warren, Robert. 2019. "Detailed Estimates of the Overstay Population Residing in the United States in 2017." New York, NY: Center for Migration Studies of New York. https://cmsny.org/visa-overstay-population-warren -120219/.
- 2024. "After a Decade of Decline, the US Undocumented Population Increased by 650,000 in 2022."
 Journal on Migration and Human Security 12(2): 85–95. https://doi.org/10.1177/23315024241226624.
- Warren, Robert, and John Robert Warren. 2013. "Unauthorized Immigration to the United States: Annual Estimates and Components of Change, by State, 1990 to 2010," *International Migration Review* 47(2): 296–329. https:// onlinelibrary.wiley.com/doi/abs/10.1111/imre.12022.