



Published in final edited form as:

*Ann Am Acad Pol Soc Sci.* 2010 September ; 631(1): 124–140. doi:10.1177/00027162110373329.

## Immigration Statistics for the 21st Century

**Douglas S. Massey**

106 Wallace Hall Princeton University Princeton, NJ 08544 dmassey@princeton.edu(609)  
258-4949

### Abstract

Of the three main contributors to population growth—fertility, mortality, and net migration—the latter is by far the most difficult to capture statistically. This article discusses the main sources of federal statistical data on immigration, each with its own characteristic set of strengths, weaknesses, possibilities, and limitations in the context of the interested social scientist. Among the key limitations, the article argues, are the elimination of parental birthplace from the Census and the lack of complete data concerning the legal statuses of the U.S. population. This article will conclude with suggestions on remedying such deficiencies, at relatively low marginal cost, such as the inclusion of questions on parental birthplace, instituting a regular survey of randomly selected legal immigrants, and the use of the “two-card method” in statistical data.

### Keywords

immigration population growth; assimilation; Current Population Survey; Yearbook of Immigration Statistics; Public Use Microdata Sample; American Community Survey

---

Immigration has become a major force in American population dynamics. Throughout the 1990s and early 2000s, it accounted for about a third of annual population growth and almost half of labor force growth. The foreign-born population of the United States currently stands at a record 38 million people who together constitute around 13 percent of the nation. Adding in the children of immigrants increases the figure to a third of the population. One in every three Americans is now in the first or second generation of U.S. residence. This figure is just the average, of course, and as one moves down the age distribution, the percentage of foreign-born and their children rises, meaning that the future of the United States will be determined very much by what happens to these two generations of people. Having accurate statistics on first- and second-generation immigrants is essential to assessing the social, economic, and demographic future of the United States. Unfortunately, immigration statistics remain the weakest link in our demographic accounting system.

### THE PROBLEM OF IMMIGRATION MEASUREMENT

In part, this long-standing weakness stems from the nature of the underlying demographic phenomenon. Among the three contributors to population growth—fertility, mortality, and net migration—the latter is by far the most difficult to capture statistically. Vital events such as birth and death are simple to define and relatively easy to detect and record. Birth is an unambiguous occurrence that happens to all human beings once and only once. A successful birth initiates a human life, with all its manifold consequences for society, making it difficult to overlook. Although modern medicine has blurred the boundaries a bit, the start of life through birth is nonetheless clear and easy to detect and record with minimal state effort. After birth, of course, we age at a constant rate subject to a changing risk of mortality, but

---

death is another unambiguous event. As with birth, we all die once and only once, and the event yields a product that must be dealt with socially, culturally, and materially. In a well-functioning society, keeping a running count of the dead is not difficult because corpses are of inherent social concern, are difficult to obscure, and require some kind of disposal, easily bringing them to the attention of state officials charged with their registration.

Although birth and death are shaped by social forces, at base they are singularly biological events grounded in human physiology. In contrast, migration is not a biological so much as a social fact, and unlike a birth or death it may occur more than once or not at all. Human beings naturally move about in space, of course, and this movement is rooted in biology. Like all organisms, humans traverse the environment to access resources, survive, and reproduce. The transformation of this natural mobility into something labeled migration, however, is a social process that involves the consensual demarcation of lines in space by a human community that is ratified, supervised, and enforced by civil authority. At times, the socially defined line is marked in the real world and rendered explicit, as along the Mexico-U.S. border, but often the line exists only in people's minds or on a map, as with the border between Maryland and Washington, D.C.

Unlike a birth or death, therefore, a move is intrinsically a social event; whereas one cannot become un-born or un-dead, one can easily move back to where one comes from, not to mention proceeding to another destination and another one beyond that before then returning to the point of origin. As a result, moves are much harder for civil authorities to define and track, and they require greater effort and expense to record and tabulate. Moreover, the third component of population growth, net migration, is a product of not one but two complementary socially defined events—arrival and departure—and the nature and determinants of movement into and out of an area are often quite distinct, making modeling and forecasting inherently difficult.

These difficulties characterize mobility between as well as within countries, but crossing an international boundary complicates the inherent difficulties even further because the social construction of movement involves two different frames of social reference, which may be not only in disagreement but in competition. In addition, the measurement of international migration is inevitably bound up with issues of identity, citizenship, belonging, and entitlement, which create strong sentiments and interests attached to the definition of who is and who is not a migrant. An international move is much more than a demographic fact, therefore; it is a social, economic, and political event with strong and often competing interests in how it is defined and measured.

Because the definition of international migration is so political, the state and those who control it often seek to manipulate official categories and definitions to achieve political aims, make ideological points, or create the appearance of some socially favored version of reality. A simple change of location between Mexico and the United States means very different things, and carries very different societal implications, depending on whether the mover is classified as a legal U.S. resident, a temporary worker, a tourist, an exchange student, a business visitor, a refugee, or an unauthorized migrant, and many people on both sides of the border care very deeply about the category into which cross-border movers are placed.

## U.S. IMMIGRATION STATISTICS

Immigration statistics in the United States come from three principal sources: the Decennial Census (2010) administered by the U.S. Bureau of the Census, Current Population Survey (2010) jointly administered by the U.S. Department of Commerce and the Bureau of Labor Statistics, and the Office of Immigration Statistics (2010), administered by the U.S.

Department of Homeland Security. These federal entities yield somewhat overlapping but mostly complementary sources of information on immigration to the United States, each with its own characteristic set of strengths and weaknesses, possibilities and limitations.

### The Decennial Census

Since 1850 place of birth has been asked as a regular part of the U.S. census, in recent decades as part of the “long form” filled out by roughly a fifth of U.S. households. In the future, it will be included on the American Community Survey (2010), which is scheduled to replace the census long form beginning in 2010. A good estimate of net international migration between censuses may be derived simply by tabulating the foreign-born population by age and sex in census year  $t$  and then using life-table methods to project the population of foreigners ahead to census year  $t+10$ , and then subtracting this projected population from the enumerated census population of foreign birth at  $t+10$ .

Since one cannot be born into the population of immigrants (as a birth in the U.S. is by definition native), mortality and migration are the only forces affecting demographic change among immigrants. If mortality is stable, the life tables are accurate, and census enumeration is reliable or at least constant over time, this method yields an accurate estimate of net international migration over a decade, at least for persons aged 10 and over. Immigrants under the age of 10 enumerated at  $t+10$  can either be assumed to have entered since the last census. When the “rolling census” enabled by the American Community Survey (ACS) becomes operative after 2010, it will be possible, in theory at least, to use these methods to derive more frequent intercensal estimates of international migration.

The United States routinely publishes and disseminates detailed data on its foreign-born population. In contrast, governments in most other developed countries do *not* publicly release tabulations of the foreign-born, precisely because they do not wish to place accurate estimates of net immigration before the public for scrutiny and debate, underscoring the political nature of immigration data. Since 1970, those born outside the United States have also been asked to report their citizenship, year of entry, and English language ability, in addition to all the other variables asked of everyone in the census (age, gender, education, occupation, etc.). Through its website, the U.S. Census Bureau regularly releases reports and produces detailed tabulations of the foreign born population, and an interactive program allows users to construct their own cross-tabulations to describe specific national origin groups within particular geographic regions.<sup>1</sup> In addition, the Public Use Microdata Sample (PUMS) that is released after each decennial census allows anyone with programming skills to construct virtually any tabulation desired, subject only to the stricture that tabulations must be produced for geographic units with at least 100,000 inhabitants, in order to preserve confidentiality. These data are available for census years from 1850 to the present from the Integrated Public Use Microdata Series (2010) project at the University of Minnesota.

Using PUMS data analysts can also estimate sophisticated statistical models of immigrant adaptation and assimilation within various domains, such as language, education, employment, housing, marriage, family, and childbearing. While statistically sophisticated, however, census-based models of immigrant assimilation are flawed because the immigrants included in the census are a selective subset of all foreigners who ever migrated to the United States, representing those persons who survived the dual risks of mortality and return migration to be in the country when the census was taken. To the extent that either death or out-migration is correlated with the outcome being studied, cross-sectional models estimated from census data may produce biased and misleading results. Since immigrants are generally

---

<sup>1</sup>See [www.census.gov/population/www/socdemo/foreign/index.html](http://www.census.gov/population/www/socdemo/foreign/index.html).

young and unlikely to die, mortality is less serious as a source of bias than return migration, which may not only be sizeable but vary substantially from country to country.

A common approach to studying immigrant assimilation, for example, is to compare some outcome of interest—say, English language ability or wages—across year-of-entry cohorts. If, controlling for other factors, those who arrived ten to fifteen years ago speak better English and earn higher wages than those who arrived less than five or five to ten years ago, we infer a process of linguistic or economic assimilation. The problem with this inference, however, is that the cohorts are synthetic and not real. The apparent increase in wages and English ability with time spent in the United States may indeed mean that immigrants improve their earnings and English skills over time, or it could simply mean that immigrants who earned lower wages or had a difficult time learning English were more likely to return to their country of origin. The two scenarios—assimilation versus selection—are not mutually exclusive, of course, and from census data alone there is no way to distinguish between the two.

In addition to each person's birthplace, from 1870 to 1970 the U.S. decennial census also asked mother's and father's birthplace, thereby enabling demographers to identify second-generation immigrants separately from other native-born Americans. In any immigrant-receiving society, a critical juncture defining the long-term prospects for assimilation and integration is what happens to members of the second generation. However rapid and complete the assimilation of immigrants, they can never be natives. The U.S.-born children of immigrants, however, are natives by definition, and the extent to which they lag behind other natives in terms of socioeconomic achievement suggests the existence of barriers to assimilation, carrying the risk that a marginalized minority underclass might form in later generations.

For this reason, studies of immigration to the United States historically have paid considerable attention to the status and well-being of second-generation immigrants. Given questions on parental birthplace, members of the second generation are readily identifiable and can be compared directly with first-generation immigrants and other natives on any characteristic measured in the census. Moreover, biases stemming from selective out-migration, while not eliminated, are much reduced among those in the second generation, who generally display very low rates of out-migration to parents' country of origin. Detailed tabulations and statistical models estimated for second-generation immigrants thus provide a fairly accurate, real-time snapshot of how different national origin groups are doing in the United States, and what their long-term social and economic prospects might be.

Unfortunately, after 1970 the questions on parental birthplace were dropped from the decennial census, and despite vociferous protests from demographers they were not reinstated on the 1980, 1990, or 2000 censuses. As a result, during a period of mass immigration greater than any seen since the early twentieth century, in which record numbers of children were being born to immigrants in the United States, social scientists had no way to study the progress of the second generation using nationally representative data. Moreover, at this writing there is still no commitment to include parental birthplace questions on the American Community Survey, which will replace the census long form beginning in 2010.

### **The Current Population Survey**

The elimination of parental birthplace from the census created a huge gap in immigration statistics and brought about a crisis in immigration research that was only partially mitigated by the fielding of specialized surveys of second-generation immigrants in specific metropolitan areas (described below). Although statisticians and demographers were unable

to achieve the reinstatement of parental birthplace questions on the census, in 1996 they won a partial victory by getting them placed on the demographic supplement to the Current Population Survey (CPS). The CPS is administered by the U.S. Census Bureau on behalf of the U.S. Bureau of Labor Statistics and is the source of monthly employment statistics, the most important being the unemployment rate. Every March the CPS includes a detailed supplement of demographic questions that historically have provided the basis for intercensal population estimates.

The inclusion of questions on parental birthplace, along with items on place of birth, year of entry, and citizenship, gave researchers access to nationally representative data with which to study the second generations for the first time since 1970. Although asking about parental birthplace on the CPS represents a clear step forward for the nation's statistical system, it by no means makes up for the loss of parental birthplace questions on the census itself, or the American Community Survey going forward. The CPS is a very small sample compared with the census long form or ACS. As a result, CPS data yield stable estimates only for large immigrant groups and sizable geographic areas. It can be used effectively, for example, to study second-generation Mexican immigrants in California, but is of little use if one seeks information about second-generation Koreans in Oregon—the sample will just be too small. In addition, the CPS does not cover the entire U.S. population, only the civilian, noninstitutionalized population, and it does not include as many variables as the census.<sup>ii</sup>

The main contribution of the CPS to immigration statistics is to provide reliable annual estimates of the size and composition of first- and second-generation immigrants at the national level, and to allow investigators to estimate models of first- and second-generation assimilation at the national level for immigrants in general and for larger groups on a timely basis. In studying immigrant assimilation, however, the CPS suffers the same weakness as the census: its cross-sectional design requires the construction of synthetic cohorts subject to biases stemming from selective mortality and out-migration.

### The U.S. Department of Homeland Security

Prior to 2002, the principal bureaucratic agency responsible for collecting and disseminating data on immigrants was the Office of Immigration Statistics of the U.S. Immigration and Naturalization Service, which was located in the Department of Justice. In that year, however, the office was transferred to the Bureau of Citizenship and Immigration Services in the newly created Department of Homeland Security. The core missions of the Departments of Justice and Homeland Security are quite different, of course, and the shift between agencies generally reduced the quantity and accessibility of immigration data released by the Office of Immigration Statistics, a fact that can be appreciated simply by comparing the thickness of the 2001 and 2008 Yearbook of Immigration Statistics, the agency's key product. Whereas the former contained 300 pages of information, the latter included just 110 pages.

What the Office of Immigration Statistics does is extract information from records of specific administrative encounters between foreigners in the United States and the immigration bureaucracy. Wherever somebody enters the United States as a permanent resident alien, for example, he or she turns in a form at a U.S. port of entry and the information from this form is transferred to the Office of Immigration Statistics, which uses it to keep a running tabulation of new arrivals and their characteristics. Unlike the census, however, the information available on the permanent residence form is quite limited.

---

<sup>ii</sup>Extracts derived from the March supplement of the CPS may be downloaded directly from the Census Bureau website at [www.census.gov/cps](http://www.census.gov/cps).

Although it yields detailed information about visa category and sponsorship, it offers only basic data on demographic background (age, gender, marital status, place of birth, and place of intended residence), very limited information on socioeconomic status (there is no information on education, for example), and virtually nothing about an immigrant's circumstances before or after entry.

The arrival of nonpermanent migrants is also tracked by the Office of Immigration Statistics, which publishes detailed counts of people in various visa categories—temporary workers, exchange visitors, students, diplomats, intra-company transferees, tourists, investors, and so on—but very little about the characteristics of the people in these classifications and nothing about what happens to them once they are in the United States. The office also keeps a running count of apprehensions along the Mexico-U.S. border as well as arrests and deportations within the United States. In the case of both temporary and apprehended migrants, the office publishes next to nothing about the characteristics of the people involved.

Each year the Office of Immigration Statistics publishes its Yearbook of Immigration Statistics,<sup>iii</sup> though typically at a lag of a year or more. Specific tables from the Yearbook are available for download in Excel format, giving analysts direct access to detailed tabulations of legal permanent residents, nonimmigrants, refugees and asylees, naturalizing immigrants, and enforcement actions. In all cases, the data correspond only to the gross inflow of people, however. The United States has no system in place to monitor the number and characteristics of people who leave the country, whether native or foreign. Moreover, in tabulating nonimmigrants and enforcement actions, the numbers refer to events and not people. Those who enter on temporary visas and those who are apprehended without documents can and do encounter the immigration bureaucracy multiple times in the course of a year, making counts in these categories poor indicators of either temporary or undocumented migration. Arrests of undocumented migrants along the Mexico-U.S. border, especially, are of limited use in trying to assess the volume of undocumented migration because apprehensions fluctuate not only with the number of people attempting to cross the border, but also with the number of officers and quantity of resources devoted to patrolling it.

## COVERAGE AND GAPS IN THE SYSTEM

Despite the shortcomings described below, the U.S. system of immigration statistics is much better than those prevailing in most other developed countries in terms of quantity, quality, and accessibility. Tabulations of the foreign-born enumerated by the census yield detailed information about the origins and characteristics of immigrants to the United States, and the routine release of census microdata samples enables investigators to estimate sophisticated models of immigrant adaptation and assimilation. Moreover, whereas the decennial nature of the census in the past made timeliness an issue, the “rolling census” enabled after 2010 by the American Community Survey will, in theory, allow more frequent updating of net immigration based on a large, census-like sample.

As we have seen, however, the census and ACS suffer from two serious weaknesses as data sources for the study of immigration. First, they yield cross-sectional data files with variables defined at a single point in time, forcing investigators to study the inherently longitudinal process of immigrant assimilation using synthetic cohorts constructed from data on year of entry, which yields results that are potentially biased by selective mortality and emigration. Second, as currently constituted, neither the census nor the ACS enables the

---

<sup>iii</sup>Available from [www.dhs.gov/files/statistics/publications/yearbook.shtm](http://www.dhs.gov/files/statistics/publications/yearbook.shtm).

identification of second-generation immigrants, a critical defect that greatly undermines our ability to evaluate how well different national-origin groups are integrating within the United States.

The latter deficit in the U.S. statistical system has been only partially mitigated by the inclusion of questions on parental birthplace in the March demographic supplement to the Current Population Survey. The addition of immigration items to the CPS—including questions on place of birth, year of arrival, and citizenship in addition to parental birthplace—enables analysts to construct a yearly statistical profile of first- and second-generation immigrants from variables available in the CPS and to estimate multivariate models of achievement for both groups compared with native persons of native parentage. However, as already noted, the CPS covers only the noninstitutionalized civilian population, and the background variables it contains are not as exhaustive as those in the census. Moreover, the relatively small size of the CPS sample restricts profiling and modeling to only the largest national-origin groups and more populous geographic areas.

As described above, the U.S. Department of Homeland Security's Office of Immigration Statistics publishes annual counts of persons entering the United States as legal immigrants, temporary migrants, and refugees and asylees, as well as counts of apprehensions, deportations, and naturalizations. This information comes from records of bureaucratic events that occur within the system of immigrant management, which yield data about only the most basic characteristics of legal immigrants and little information about those naturalizing, arrested, and deported beyond national origin. Moreover, frequencies in the latter three categories represent events and not people, and in the case of legal and temporary immigrants the numbers and characteristics refer only to arrivals, with no accounting or adjustment made for departures. At best, therefore, they yield annual information on gross inflows to the United States, even though outflows of immigrants are substantial and vary markedly from group to group.

None of the former sources is able to measure adequately the current or former legal status of immigrants. Although the Office of Immigration Statistics explicitly presents data by legal category, its coverage of the undocumented population is limited to the subset of those who are arrested or apprehended, and even the most basic personal characteristics are not gathered or reported. Likewise, the Department of Homeland Security reports in some detail the current status of arriving resident aliens and nonimmigrants, but offers limited information about prior experience in the United States.

The lack of a measure of prior U.S. experience is a serious problem, as the vast majority of “new” legal immigrants have already spent time in the country in some other status. Moreover, the amount of this prior experience and the status in which it was accumulated vary greatly from country to country. For example, whereas two-thirds of new permanent resident aliens from Mexico have prior experience as undocumented migrants (with another 15 percent as legal visitors, 1 percent as temporary workers, and 2 percent as students), among new permanent residents from Korea, only 19 percent have prior undocumented experience (compared with 41 percent as legal visitors, 3 percent as temporary workers, and 24 percent as students).

Moreover, the census, ACS, and CPS do not gather any information on the legal status of the foreign-born except for U.S. citizenship. Noncitizens come in a variety of legal categories, however, each of which carries different rights and privileges, with very different implications for status and well-being in the United States. Legal resident aliens have most, but not all of the rights of citizens, whereas asylees and refugees have fewer rights, and temporary workers have fewer still. The privileges of people entering on student, exchange,

and temporary work visas are highly constrained, and those in the country without authorization have no rights at all. All these people are included among the foreign-born enumerated in the census, ACS, and CPS, and once again the distribution varies widely from group to group.

The inability to measure and control for legal status creates a major source of unmeasured heterogeneity leading to significant bias in models of immigrant assimilation and incorporation. Among major immigrant groups, the share of undocumented migrants ranges from a low 12 percent among Indians to a high of 67 percent among Hondurans. In general, Latin American immigrants are much more likely to be undocumented, with figures of 66 percent for those from Guatemala, 60 percent from Mexico, 54 percent from Brazil, 50 percent from El Salvador, and 41 percent from Ecuador, compared with just 13 percent for China, 23 percent for Korea, and 12 percent across all other countries. If undocumented migrants suffer greater exploitation and discrimination than other immigrants in American society, then estimates of parameters in assimilation models will be severely biased for Latin Americans relative to other immigrant groups.

## FILLING THE GAPS

The foregoing review has uncovered a variety of salient gaps in the U.S. system of immigration statistics:

1. The ACS, which replaces the census long form in 2010 and constitutes the most important source of information on the characteristics of immigrants, does not include questions on parental birthplace, making it impossible to identify and study members of the second generation using the nation's premier data source.
2. Data currently available from the Census Bureau and Department of Homeland Security are exclusively cross-sectional, whereas processes of immigration adaptation, adjustment, and assimilation are inherently longitudinal.
3. Data collected on legal immigrants by the Department of Homeland Security are limited to information included on the form turned in by people applying for permanent resident status and omit critical characteristics such as education, prior migratory experience, and home country circumstances.
4. There is currently no nationwide source of information on the number, characteristics, and behavior of undocumented migrants, who constitute nearly a third of all immigrants present in the United States and up to two-thirds of some groups.
5. Although undocumented and temporary legal migrants are included in the census, CPS, and ACS, they cannot be individually identified, making it impossible to profile the undocumented population and introducing bias into models of immigrant assimilation and attainment.

In response to these gaps, academic investigators have launched several data collection efforts on their own, often with federal support. To provide better data on undocumented migrants, researchers at the University of Guadalajara and Princeton University obtained support from the National Institute of Child Health and Human Development to launch the Mexican Migration Project (2010) and the Latin American Migration Project (2010). Since 1987, Jorge Durand and I have annually supervised teams of interviewers to conduct representative surveys in Mexican sending communities and their branch communities in the United States, and since 1998 they have applied the same methods in other nations throughout Latin America.<sup>iv</sup> While not necessarily representative of all undocumented migrants from the region, the two projects nevertheless provide detailed and reliable

information on the behavior and characteristics of a large sample of undocumented migrants from the most important source countries.

To overcome the reliance of investigators on cross-sectional datasets plagued by selection bias, other researchers launched the New Immigrant Survey (2010) with funding from a consortium of federal agencies, including the National Institute of Child Health and Human Development, the National Institute on Aging, the National Science Foundation, the Department of Health and Human Services, and the Department of Homeland Security. In 2003, James Smith at RAND, Mark Rosenzweig at Yale, Guillermina Jasso at New York University, and I, working at Princeton, fielded a survey of around 8,600 new legal immigrants who had just received their residence documents, and in 2008 and 2009 they reinterviewed the same people to create a longitudinal panel study of immigration. The baseline survey is publicly available on the Internet,<sup>v</sup> and while the second-wave survey is still being processed, it will ultimately be linked to the first wave and be made publicly available to users via the project website.

Several other researchers have sought to overcome the lack of information on the second generation by undertaking specialized surveys of immigrants and their children in particular metropolitan areas. With funding from the National Institutes of Health, the National Science Foundation, the Sloan Foundation, and the Russell Sage Foundation, Alejandro Portes at Princeton University and Rubén Rumbaut at the University of California at Irvine launched the Children of Immigrants Longitudinal Study (2010), which surveyed immigrant children who arrived in the country at a young age or who were born in the United States of immigrant parents. They focused on the Miami and San Diego metropolitan areas with a baseline interview fielded in 1992 and 1993 and then reinterviewed the same respondents in 1995 and 1996 and 2001 and 2003, thereby creating an unusually rich and detailed longitudinal file that is disseminated from the project website.<sup>vi</sup>

With funding from the Russell Sage Foundation, two additional projects were launched to study second-generation immigrants in New York and Los Angeles. The Immigrant Second Generation in Metropolitan New York Study (2010), directed by John Mollenkopf and Philip Kasinitz of the City University of New York and Mary Waters of Harvard University, surveyed second-generation immigrants from Colombia, Ecuador, Peru, the Dominican Republic, the West Indies, China, and Russia, along with Puerto Ricans, native whites, and native blacks, who were 18 to 32 years old and who had grown up in and around New York. The Immigration and Intergenerational Mobility in Metropolitan Los Angeles Study (2010), directed by Rubén G. Rumbaut and Frank Bean at the University of California at Irvine, along with colleagues Leo Chavez, Jennifer Lee, Susan Brown, Louis DiSipio, and Min Zhou, in 2004 surveyed young adults aged 20 to 39 from six foreign-born and foreign-parentage groups, including persons from Mexico, Vietnam, the Philippines, Korea, China, Guatemala, and El Salvador, as well as native-born and native-parentage Mexican-Americans, non-Hispanic whites, and non-Hispanic blacks in the Los Angeles metropolitan area.<sup>vii</sup>

These privately administered but often publicly funded databases fill, at least partially, some of the gaps in the federal system of immigration statistics. Nonetheless, several salient gaps remain and require attention. The most pressing need is for the inclusion of questions on parental birthplace in the ACS. This action would make data on second-generation

<sup>iv</sup>This data can be viewed via the project websites at [mmp.opr.princeton.edu](http://mmp.opr.princeton.edu) and <http://lamp.opr.princeton.edu>.

<sup>v</sup>Available at [nis.princeton.edu](http://nis.princeton.edu).

<sup>vi</sup>Available at [cmd.princeton.edu/data%20CILS.shtml](http://cmd.princeton.edu/data%20CILS.shtml).

<sup>vii</sup>The data are available from the Inter-University Consortium for Political and Social Research at [www.icpsr.umich.edu/cocoon/ICPSR/STUDY/22627.xml](http://www.icpsr.umich.edu/cocoon/ICPSR/STUDY/22627.xml).

immigrants routinely available as part of the “rolling census” that will commence after 2010. At a time when the population of second-generation immigrants is larger than ever, and when so many public debates revolve around the status and well-being of the children of immigrants, it is unconscionable to lack national statistics on the second generation. Although parental birthplace questions may not be mandated by specific legislation, the welfare of the newest generation of Americans is central to the mission of virtually every branch of the federal government, from the Department of Health and Human Services to the Department of Education to the Department of Justice.

A second recommendation is to institute an annual survey of new legal immigrants to the United States. There is no reason why data on this population need be limited to information collected on government forms. At the point when new legal immigrants turn in the forms to claim permanent resident status, a subset could easily and quite inexpensively be randomly surveyed to collect detailed information on their social and economic characteristics, including information about social connections to other immigrants in the United States and to friends and relatives at home. Recent work has clearly established the importance of migrant networks and social capital in the process of international migration, yet virtually no information is available on the subject from federal sources. In addition, the survey could ask about circumstances in the home country prior to leaving for the United States in order to quantify and analyze the potential costs and benefits of international migration.

Finally, one of the most vexing problems facing immigration statisticians is the measurement of undocumented migration. Current population estimates are generated using an indirect demographic approach that accounts for changes in the population of foreign birth stemming from permanent legal immigration, temporary legal migration, mortality, and out-migration and then derives the population of undocumented migrants as a residual. The latest estimates prepared by the Office of Immigration Statistics put the figure at around 11 million persons, roughly a third of the total immigrant population of the United States (Hoefler, Rytina, and Baker 2010).

Residual estimates, however, are notoriously sensitive to errors in the components, and while mortality and in-migration by temporary and permanent legal migrants are known with some precision, the measurement of out-migration itself rests on assumptions and estimates. It would be useful, therefore, to have an independent means of estimating the size of the undocumented population. Such an estimate could be achieved by periodically administering a “two card” method developed by the U.S. Government Accountability Office (2009) as part of the Current Population Survey.

Under this method, separate randomly chosen subsamples of respondents are asked about their current immigration status and asked to classify themselves into categories listed on a card, but for one set of respondents the card includes “currently undocumented” grouped along with “legal permanent resident” and “refugee/asylee” (Card 1); on the card presented to the other set of respondents, “currently undocumented” is grouped with “citizen” and “temporary visitor,” leaving “legal permanent resident” and “refugee/asylee” in a separate category (Card 2). The number of undocumented migrants is then estimated by subtracting the number who answered yes to the category containing just legal immigrants and refugees/asylees on Card 2 from the number who answered yes to the category containing these groups plus the undocumented on Card 1.

This method overcomes the implicit threat posed by asking whether someone is undocumented because those without documents only ever admit to being in a category that contains undocumented migrants among other kinds of migrants. That way it is impossible to identify any *individual* as being undocumented, and respondents are unafraid to respond

truthfully. Unfortunately, the same feature that guarantees confidentiality also makes it impossible to link individual characteristics to legal status, so that it is not possible to define the characteristics of the undocumented population. Nonetheless, regular implementation of the two-card method on the CPS would provide an independent estimate of the size of the undocumented population to serve as a check on validity of the indirect demographic methods currently in use.

## CONCLUSION

Compared with the statistical systems of other nations, the United States does a fairly good job in developing and publishing statistics on immigration. The regular release of data on persons of foreign birth enumerated on the Decennial Census, the ACS, and the CPS provides detailed information on changes in the number of immigrants and their characteristics, whereas data tabulated by the Office of Immigration Statistics at the Department of Homeland Security provide basic data on new legal permanent and nonpermanent immigrants. Although the resulting data are cross-sectional, a new federally funded survey, the New Immigrant Survey, promises to provide researchers with longitudinal data on immigrant assimilation. Moreover, using indirect methods federal statisticians can estimate the size of the undocumented population, and though the characteristics and behavior of undocumented migrants cannot be studied using federal sources, data from the Mexican Migration Project and the Latin American Migration Project offer access to this information.

Figure 1 illustrates how different data sources can be combined to shed light on the controversial subject of Mexico-U.S. migration. Our southern neighbor is the largest single source of immigrants to the United States, accounting for around a third of the foreign-born population, and it is the overwhelming source for most undocumented migration, accounting for around 60 percent of the population. The figure summarizes the flow of Mexican immigrants into the United States over the past century in three categories: legal immigrants, temporary legal workers, and undocumented migrants. The figures on legal immigration and temporary worker migration come from the 2008 Yearbook of Immigration Statistics (Office of Immigration Statistics 2009), whereas the net inflow of undocumented migrants was estimated by applying probabilities of undocumented migration estimated using data from the Mexican Migration Project (see Durand and Massey 2004) to census counts of Mexicans on both sides of the border. In the data series on legal immigration, former undocumented migrants who legalized under the 1986 Immigration Reform and Control Act are excluded from the counts, as they are already included in the estimated flow of undocumented migrants.

The figure shows that large-scale Mexico-U.S. migration is hardly new; it has developed and evolved over more than a century. The origins go back to 1907, when the United States negotiated a cutoff of labor migration from Japan and U.S. employers immediately began recruiting in Mexico. Once the U.S. entered the First World War, the government set up its own temporary worker program; during the 1920s Mexican migration surged as quotas were imposed to curtail immigration from Europe. With the onset of the Great Depression in late 1929, however, Mexicans were no longer welcome, and U.S. officials launched a mass deportation campaign that from 1929 to 1939 forcibly expelled 469,000 Mexican citizens from the United States, bringing about a definitive end to Mexico-U.S. migration for a decade.

As quickly as the welcome mat for Mexicans was withdrawn with the onset of the Great Depression, however, it was put out again once the United States entered the Second World War. A few months after Pearl Harbor, the U.S. government approached Mexico to

negotiate a bilateral temporary worker agreement known as the Bracero Accord. As shown in the figure, the period from 1950 to 1965 was dominated by the rise and fall of this temporary worker program. The era started small with just 4,200 migrant workers in 1942, but by 1945 the program had expanded to nearly 50,000 per year. After a brief lull in 1946, the program resumed steady growth and peaked in 1956 at 445,000 workers.

During the latter half of the 1950s the number of bracero migrants averaged 433,000 per year and then declined to a yearly average of 233,000 during the early 1960s. Because Mexico was not under quota limitations, legal Mexican immigration also grew during the bracero era, going from around 2,200 per year in 1942 to more than 55,000 in 1963. With ample opportunities for legal entry through the early 1960s, undocumented migration was virtually nonexistent, and border apprehensions averaged just 30,000 per year.

In 1965, however, the United States unilaterally canceled the Bracero Program and passed amendments to the Immigration and Nationality Act that imposed the first numerical limits on legal immigration from Mexico, restrictions that grew increasingly stringent with amendments passed in 1976 and 1978. Whereas in 1959 Mexicans had access to 438,000 temporary work visas and an unlimited number of residence visas, by 1979 just 1,725 work visas were issued, and the annual quota for Mexican immigrants who were not immediate relatives of U.S. citizens stood at just 20,000.

As the demand for Mexican workers continued to grow and the supply of legal workers dried up, the gap was filled by unauthorized workers. Net undocumented migration rose from zero in the early 1960s to reach 150,000 in the late 1970s and then climbed to around 300,000 in the late 1980s. During this time legal immigration fluctuated between 50,000 and 100,000 persons per year. After 1990 rising border enforcement and the implementation of ever harsher anti-immigrant policies steadily cut the net undocumented inflow, which fell to around 200,000 per year in 2000, and then plummeted to reach zero by 2008.

It was not only enforcement actions that caused the decline in undocumented migration, however, for just as U.S. policies were closing one door to the country, they quietly and without fanfare opened another. Temporary worker migration began to increase in response to farm labor shortages in the late 1990s, but after 2001 as undocumented migrants found it increasingly difficult to get to the fields, the number of temporary work visas issued to Mexican farm workers surged. In 1990 only 17,000 temporary workers entered from Mexico, but by 2008 the number had reached 360,000, a figure not seen since the heyday of the Bracero Program. Workers who would have entered the U.S. as undocumented migrants in 1988 were coming in as legal temporary workers in 2008.

After 1990 the number of legal immigrants entering the United States also continued to rise, despite numerical restrictions. Ironically, the rising tide of legal immigration from Mexico was itself an unintended consequence of the increasing pressure put on permanent residents by U.S. legislation, which barred noncitizens from key entitlement programs, put them at risk of expulsion for earlier immigration violations, and denied access to due process in immigration proceedings. In response, Mexican resident aliens increasingly sought protection by acquiring U.S. citizenship. A rising rate of naturalization carries important implications for immigration because each new citizen acquires the right to bring in parents, spouses, and minor children without numerical limitation while also gaining the right to petition for the entry of brothers and sisters under quotas.

The data thus suggest that we have entered a new period in the history of Mexico-U.S. migration. After years of growth, undocumented migration has come to a halt and the undocumented population has stabilized at around 12 million. Those already in the country show no inclination to return to Mexico, and, in fact, rates of undocumented out-migration

are themselves near zero. At the same time, the entry of Mexicans through legal channels—as either permanent residents or temporary workers—is approaching record levels.

It remains to be seen what the future will bring, but understanding the history of Mexico-U.S. migration in this way would have been impossible without immigration statistics that were put together from publicly available sources supported by the federal government. The principal weakness in the current system of immigration statistics at this point is the lack of data for a large sample of second-generation immigrants. Other problems are the limited amount of information available for new permanent residents and the reliance on indirect methods to estimate the size of the undocumented population. These deficiencies can be remedied, however, at relatively low marginal cost by including questions on parental birthplace on the ACS, by instituting a regular survey of randomly selected legal immigrants as they file the paperwork to receive their documents, and by including questions for the two-card method periodically on the CPS.

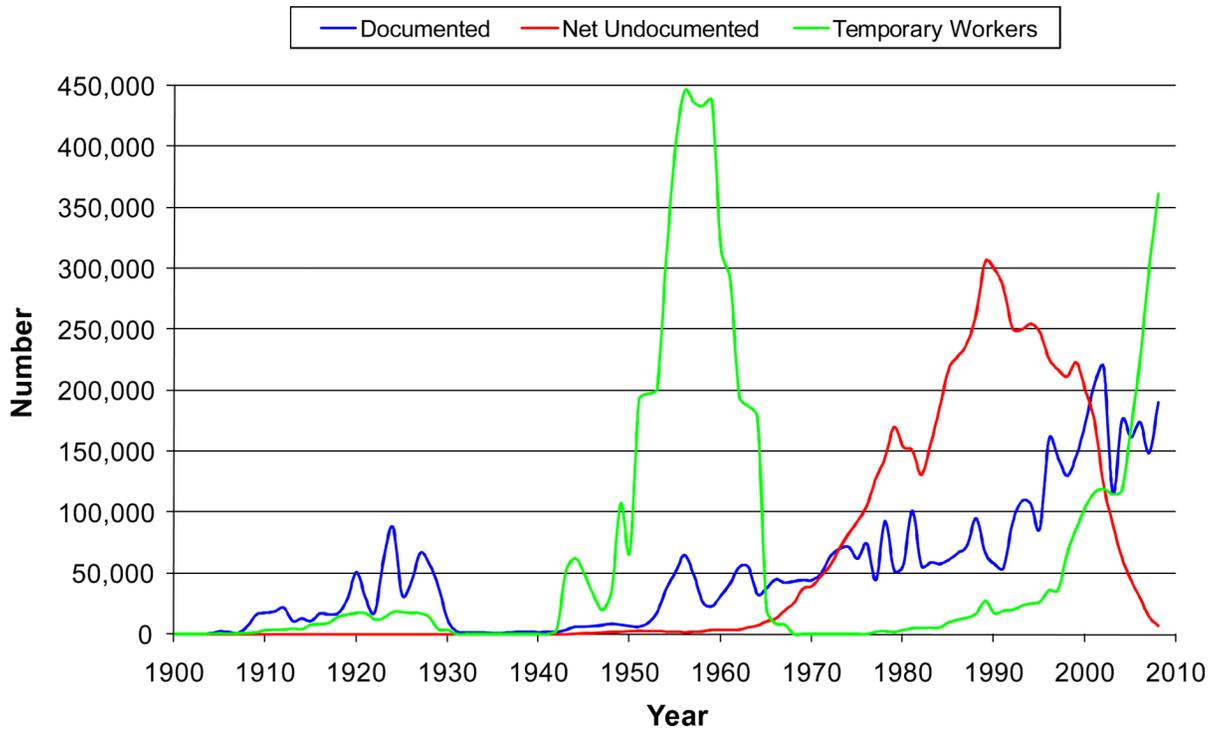
## Biography

Douglas S. Massey is the Henry G. Bryant Professor of Sociology and Public Affairs at Princeton University and the president of the American Academy of Political and Social Science. His research focuses on international migration, race and housing, discrimination, education, urban poverty, and Latin America, especially Mexico. His most recent books are *New Faces in New Places: The Changing Geography of American Immigration* (Russell Sage Foundation 2008) and *Categorically Unequal: The American Stratification System* (Russell Sage Foundation 2007).

## References

- American Community Survey. Homepage of the American Community Survey. U.S. Bureau of the Census; Washington, DC: 2010. <http://www.census.gov/acs/www/> [April 21, 2010]
- Children of Immigrants Longitudinal Study. Homepage of the Children of Immigrants Longitudinal Study. Center for Migration and Development, Princeton University; Princeton, NJ: <http://cmd.princeton.edu/data%20CILS.shtml> [April 21, 2010]
- Current Population Survey. Homepage of the Current Population Survey. U.S. Bureau of the Census and Bureau of Labor Statistics; Washington, DC: 2010. <http://www.census.gov/cps/> [April 21, 2010]
- Decennial Census. Homepage of the U.S. Census Bureau. U.S. Bureau of the Census; Washington, DC: 2010. <http://www.census.gov/> [April 21, 2010]
- Durand, Jorge; Massey, Douglas S. *Crossing the Border: Research from the Mexican Migration Project*. Russell Sage Foundation; New York: 2004.
- Hoefer, Michael; Rytina, Nancy; Baker, Brian C. *Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2009*. Office of Immigration Statistics, U.S. Office of Homeland Security; Washington, DC: 2010. <http://www.dhs.gov/files/statistics/publications/#2>
- Immigrant Second Generation in Metropolitan New York Study. Website of Immigrant Second Generation in Metropolitan New York Study. Russell Sage Foundation; New York: 2010. [http://www.russellsage.org/programs/main/immigration/the\\_immigrant\\_second\\_generation\\_in\\_metropolitan\\_new\\_york](http://www.russellsage.org/programs/main/immigration/the_immigrant_second_generation_in_metropolitan_new_york) [April 21, 2010]
- Immigration and Intergenerational Mobility in Metropolitan Los Angeles Study. Homepage of the Immigration and Intergenerational Mobility in Metropolitan Los Angeles Study. Russell Sage Foundation; New York: 2010. [http://www.russellsage.org/programs/main/immigration/immigration\\_and\\_intergenerational\\_mobility\\_in\\_metropolitan\\_los\\_angeles](http://www.russellsage.org/programs/main/immigration/immigration_and_intergenerational_mobility_in_metropolitan_los_angeles) [April 21, 2010]
- Integrated Public Use Microdata Studies. Homepage of IPUMS-USA. Minnesota Population Center, University of Minnesota; Minneapolis, MN: 2010. <http://usa.ipums.org/usa/> [April 21, 2010]

- Latin American Migration Project. Homepage of the Latin American Migration Project. Office of Population Research, Princeton University; Princeton, NJ: 2010. <http://lamp.opr.princeton.edu/> [April 21, 2010]
- Mexican Migration Project. Homepage of the Latin American Migration Project. Office of Population Research, Princeton University; Princeton, NJ: 2010. <http://mmp.opr.princeton.edu/> [April 21, 2010]
- New Immigrant Survey. Homepage of the New Immigrant Survey. Office of Population Research, Princeton University; Princeton, NJ: 2010. <http://nis.opr.princeton.edu/> [April 21, 2010]
- Office of Immigration Statistics. 2008 Yearbook of Immigration Statistics. U.S. Department of Homeland Security; Washington, DC: 2009.
- Office of Immigration Statistics. Immigration Statistics Homepage. U.S. Department of Homeland Security; Washington, DC: 2010. <http://www.dhs.gov/files/statistics/immigration.shtm> [April 21, 2010]
- U.S. Government Accountability Office. Estimating Irregular Migration in a Survey: The 'Two-Card Follow-Up' Method."United Nations, Department of Economic and Social Affairs, Population Division (2009).. Seventh Coordination Meeting on International Migration; 2009. Working Paper No. ESA/P/WP.208



**Figure 1.**  
Mexican migration to the United States in three categories.