United States General Accounting Office

GAO

Report to the Chairman, Subcommittee on Immigration, Committee on the Judiciary, U.S. Senate

September 1995

ILLEGAL IMMIGRATION

INS Overstay Estimation Methods Need Improvement

GAO/PEMD-95-20

GAO

United States
General Accounting Office
Washington, D.C. 20548

Program Evaluation and
Methodology Division

B-261245

September 26, 1995

The Honorable Alan K. Simpson
Chairman, Subcommittee on Immigration
Committee on the Judiciary
United States Senate

Dear Mr. Chairman:

This letter is in response to your April 18, 1995, request that we examine the basis for the Immigration and Naturalization Service (INS) estimates of overstays—that is, persons who entered the United States legally as visitors but did not leave under the terms of their admission. INS estimates that about half of the illegal resident aliens in the United States during recent years have been overstays. The other half are principally persons who crossed U.S. borders illegally.

Reliable and valid estimates of the number of overstays are important to public policy-making. Higher numbers of overstays might suggest, for example, the need for stricter policies or laws for issuing temporary U.S. visas to citizens of those countries whose travelers tend to overstay their visas in significant numbers. Higher numbers of illegal border crossers might suggest a greater need for actions such as further improving border security. Overstay data are also needed to monitor travel from countries whose citizens are not required to obtain a U.S. tourist visa.[1]

Background

The task of estimating overstays presents a difficult challenge. Although INS has a data system for tracking the dates when individual foreign visitors arrive in and depart from the United States, the agency cannot assume that all persons whom the system does not record as having left have, in fact, overstayed their lawful periods of entry.

[1]Twenty-three countries (mostly advanced industrial countries) are currently included in the Visa Waiver Pilot Program, extended through Sept. 30, 1996, by Public Law 103-416, sec. 210, Oct. 25, 1994. If the overstay rate INS estimates for a country in the pilot program is above a specified threshold, that country must be removed from the program. The full list of current visa-waiver pilot countries is given in appendix I, footnote 4, p. 17.

Most foreign visitors tracked by the INS data system travel to the United States by air.[2] The INS data system works as follows: Airline personnel distribute arrival/departure forms (I-94s) to foreign visitors.[3] An INS inspector collects the arrival portion of the form as each visitor enters the United States; at that time, the inspector staples the departure form in the visitor's passport. The airlines are responsible for collecting the departure forms when visitors leave and for sending those forms to INS. But for some visitors who may have actually left the United States, INS apparently has no record of the departures. It is thus not clear, in any particular instance, whether an "unrecorded departure form" represents a person who overstayed or whose actual departure went uncounted.

Objectives, Scope, and Methodology

Our objectives were to examine the basis of INS overstay estimates and, as appropriate, to suggest improved estimation procedures. In reviewing the agency's methodology, we examined published documents that describe INS procedures and interviewed staff about their methods of overstay estimation, the kinds of data that are available, and the potential for devising improved overstay estimates.[4] We focused primarily on tourist air arrivals since they represent the majority of the foreign visitors in the INS data system.[5] We developed new estimation procedures and applied them to tourist visitors arriving by air in the United States from nine countries from October 1990 to March 1991.[6] We also talked with officials from the Department of State about their use of INS estimates and asked a number of experts to review a draft of this report.[7]

We conducted our review in accordance with generally accepted government auditing standards.

[2]INS' Nonimmigrant Information System (NIIS), which tracks nonimmigrant visitors to the United States, does not include (1) certain categories of Canadian visitors or (2) Mexican visitors whose travel is limited to an area within 25 miles of the U.S.-Mexican border for periods of no more than 72 hours. The system also does not include foreign students (or their dependents) admitted to study at U.S. colleges, universities, or vocational schools.

[3]The information requested on the arrival form includes airline of entry, sex, date of birth, and certain other data. (The form is reproduced as figure I.1 in appendix I.)

[4]Our review did not include an assessment of the validity of the criteria INS used to select the "index" countries (which, as explained in a subsequent section, are used to estimate uncounted departures).

[5]Other than Mexican and Canadian visitors (who often enter this country by land), most foreign visitors arrive by air.

[6]INS provided us with these data from its Nonimmigrant Information System.

[7]The expert reviewers are listed in appendix III.

Results in Brief

Despite difficulties in obtaining accurate departure data, INS devised a creative approach for estimating overstays through estimating the number of uncounted departures (that is, "system error"). Specifically, INS determined that system error could be estimated by using data from countries for which it seems safe to assume there are few or no overstays (that is, "index countries"). INS applied its index countries' strategy using a "global" approach that requires the assumption that a single rate of system error applies to all countries worldwide.[8] We believe that this assumption is questionable and that improved approaches have the potential to reduce the uncertainty of current INS estimates.

We devised an alternative method for estimating overstays among foreign visitors who arrive by air. Our method is based on INS' index country strategy but uses more detailed INS data and avoids the global assumption. Our method also corrects an error in INS' computation formula[9] and uses appropriately weighted data. (INS did not weight the data.) When we applied our method to sample INS data for nine nonindex or sending countries, we found overstay estimates that are, on average, lower than corresponding INS estimates.[10] Specifically, our estimates are between 16 percent and 47 percent lower than INS'. (See table 1 on p. 8. While these are not worldwide estimates, the nine sending countries do account for over one-third of all overstays in the tourist air arrival class.)

INS' global approach provided a good starting point for estimating overstays. Our approach is different mainly in that it makes more limited assumptions. We believe this may reduce the uncertainty of the estimates. In response to early presentations of our work, INS and other experts suggested additional strategies for reducing the uncertainty of overstay estimates. (See appendix I.) We believe that INS should develop improved procedures for estimating overstays, using methods such as those discussed in this report.

[8]If this assumption does not hold true, the resulting estimates of overstays would be inaccurate—at least for certain countries and possibly worldwide.

[9]For tourists arriving by air, the resulting error in INS' estimate is on the order of 7 to 10 percent.

[10]These nine sending countries are the Bahamas, France, Germany, Haiti, India, Mexico, Philippines, Poland, and the former Soviet Union.

GAO's Analysis

INS considers the pool of unrecorded departure forms to represent two groups of persons—true overstays and uncounted departures.[11] For true overstays, no departure forms were ever turned in because these visitors never left the United States. For uncounted departures, there was a paperwork error of some sort; that is, these visitors did leave but their departures were not counted by INS (either because the airline did not collect their departure forms or for some other reason). INS has termed the unrecorded departure forms in this latter group "system error." Earlier we noted that one could not simply assume that all unrecorded departure forms represent overstays. Here we see that some unrecorded departure forms represent system error. In short, true overstays equal unrecorded departure forms minus the number attributable to system error.

In order to estimate true overstays, INS thus needs to estimate system error. INS does this using 11 "index countries" that it believes have so few true overstays that all unrecorded departure forms can be considered uncounted departures (system error).[12] If INS finds, for example, that 10 percent of airline travelers arriving from these index countries were not counted as departures, the agency then assumes that the percentage of uncounted airline departures is about 10 percent in all countries.[13] That is, INS uses one estimate of the percentage of uncounted departures for all countries of the world.

INS' specific method of estimating overstays from each country is to subtract an estimate of uncounted departures from the total number of unrecorded departure forms for each country, based on a global estimate of the proportion of departure forms that are missing because of system error. INS then totals the individual countries' overstay estimates to produce an estimate of the number of true overstays from all countries.[14]

[11]In this report, we use the terms "true overstay" and "overstay" interchangeably. We use the term "unrecorded departure forms" where INS has used the term "apparent overstays."

[12]The index countries are Australia, Belgium, Finland, Netherlands Antilles, New Zealand, Norway, Saudi Arabia, Sweden, Singapore, Suriname, and Switzerland. INS dropped Kuwait, a country it had originally included as an index country, because of increased overstays resulting from the Gulf war. With respect to the assumption that the index countries have virtually no true overstays, we note that all sizable countries, including index countries, probably have at least some "technical overstays." (See footnote 10 in appendix I.)

[13]As explained in appendix I, INS adds an adjustment to this figure when estimating uncounted departures for nonindex countries. This adjustment accounts for the fact that index countries may have slightly lower rates of uncounted departures than other countries. (An upward adjustment seems justified because one criterion INS used in selecting the index countries was low rates of unrecorded departure forms.)

[14]INS uses the same procedure to separately estimate overstays who arrived by air, land, and sea. (See appendix I, pp. 42-43, for a discussion of land overstays.)

While this method is creative, it makes a sweeping assumption. INS itself had characterized its procedures as more likely to overestimate than underestimate the number of overstays. Officials from the Department of State also questioned the accuracy of INS' estimates of overstays for certain countries.[15]

In a detailed review, we found that estimates based on INS' global approach are necessarily marked by considerable uncertainty because of the assumption that rates of system error are the same for index countries and all other countries. Although INS' procedures assume it is safe to ignore the potential for differences in system error between index countries and other countries, airline-specific data for the index countries showed evidence to the contrary: First, INS' airline-specific data confirmed that citizens of different countries tend to fly to the United States on different airlines. Second, INS' airline-specific data for index countries showed that there is substantial variation in system error across airlines.

We also found an algebraic error in INS' computation formula. INS' use of an incorrect formula resulted in an error on the order of 7 to 10 percent in INS' overstay estimates for air travelers.[16]

GAO's Method

We devised a method that uses INS' airline-specific data for index countries as the basis for estimating separate rates of system error for each airline. We explain this method in detail below, but the main point is that it allows us to estimate system error separately for each airline—rather than rely on the assumption that a single rate of system error applies to all flights from all countries of the world.

Specifically, we began by estimating system error separately for each airline that regularly flies between one or more index countries and the United States.[17] For example, we used the rate of unrecorded departure forms for index country passengers on United Airlines to estimate system

[15]That is, they told us they believed certain estimates did not seem plausible in view of the State Department's experience with or knowledge about those countries.

[16]The specific effect of the error in INS' computation formula—in and of itself—is to underestimate overstays. The specific size of the difference in estimates owing to this error varies according to the particular data in question. See appendix I, pp. 28-30, for a detailed explanation of the error.

[17]This approach was suggested by ongoing analyses at INS in which data on unrecorded departure forms were crosstabulated separately, airline by airline. It uses INS' index countries' strategy but carries it further.

error for all United Airlines flights.[18] We repeated this procedure, estimating system error separately for Delta Air Lines, Air Micronesia, and so forth, on the basis of index countries' data for each airline.

Then, we estimated overstays for nonindex countries—proceeding separately, country by country and airline by airline. For example, we estimated overstays from the Philippines who flew to this country on United Airlines using the specific rate of system error we had estimated for United Airlines; then we estimated overstays from the Philippines who flew in on Air Micronesia using the estimated rate of system error for Air Micronesia, and so forth. Essentially, we assumed that system error on United Airlines flights between the Philippines and the United States was about the same as for United Airlines flights between index countries and the United States. We made the same assumption for Air Micronesia flights, and so forth.

For some airlines, this approach was not possible because they do not regularly fly between any index country and the United States. For these "nonqualifying" airlines, additional assumptions were necessary. In making assumptions, we wished to be as conservative as possible. We therefore calculated two sets of estimates using alternative assumptions, as described below.

Assumption 1 for Nonqualifying Airlines: Same Overstay Rates

We first assumed that passengers on nonqualifying airlines have the same overstay rate as passengers from the same sending country who arrived in the United States on qualifying airlines (for which the separate calculations could be performed).

The example of Mexico can illustrate how our method works when this assumption is used. Separate estimates of system error were available for American Airlines, Delta Air Lines, United Airlines, Air Micronesia, and other "qualifying" airlines (which fly between index countries and the United States)—but not for airlines such as Mexicana de Aviacion and Aeromexico. We were thus able to calculate separate overstay estimates for Mexican visitors who arrived on each qualifying airline. For Mexican visitors who arrived on "nonqualifying" airlines that do not regularly fly

[18]That is, INS selected the index countries because visitors from these countries are apparently so unlikely to overstay that it may be reasonable to assume that virtually none are true overstays. Assuming that virtually none of the United Airlines passengers from the index countries overstayed, the rate of unrecorded departure forms calculated for index countries' visitors who flew in on United Airlines represents a rate of uncounted departures—or system error for that airline. Following the INS procedure described in footnote 13 (and in appendix I), we added one standard deviation to the rate of uncounted departures calculated for index countries' passengers on United Airlines. Unlike INS, we always used weighted data.

between index countries and the United States (such as Mexicana de Aviacion and Aeromexico), we assumed that the overstay rate was the same as for Mexican visitors arriving on qualifying airlines. Using these procedures, we obtained overstay estimates for nine nonindex, or sending, countries.

Assumption 2 for Nonqualifying Airlines: Same Rates of System Error

We also performed a second set of calculations for the nine countries, using the alternative assumption that, for a particular sending country, passengers on both types of airlines (qualifying and nonqualifying) had the same rates of system error.[19]

Again using the example of Mexico to illustrate, we calculated separate overstay estimates for passengers on each qualifying airline (American Airlines, United, Delta, Air Micronesia, and so forth)—exactly as before. We then assumed that for Mexicana de Aviacion, Aero Mexico, and other nonqualifying airlines, the rate of system error was the same as calculated for Mexicans who arrived on qualifying airlines.

GAO's Estimates and INS'

We believe our method represents an improvement on two counts: First, the assumptions required are more limited than those required by INS' global method because, to the extent possible, our method relies on airline-by-airline estimates of system error. Second, we used a correct, mathematically derived formula, as well as appropriately weighted data.[20]

Table 1 presents our two sets of estimates, based on the alternative assumptions for nonqualifying airlines, for nine sending countries.[21]

[19]Our second set of calculations is like INS' method in that it makes assumptions about rates of uncounted departures (system error) but not about overstay rates. But it differs from INS' in that our assumptions about rates of system error are more limited; that is, unlike INS', our assumptions about system error do not involve airlines that do not fly between the country in question and the United States.

[20]In INS' calculation of system error, a large index country and a small index country are given equal weight. All our calculations are weighted to represent the relevant number of visitors from each country. For example, in estimating system error for a particular airline, data for index countries' passengers on that airline are appropriately weighted to reflect the number of passengers from each index country who flew here on that airline.

[21]Of course, comparisons of our estimates to INS' for the nine sending countries are not necessarily indicative of results that would be obtained if estimates were calculated for all countries worldwide.

Table 1: GAO and INS Estimates of Overstays Among Tourist Air Arrivals From Nine Countries[a]

		Overstay estimates			
		GAO airline-by-airline method		INS global method	
Continent and sending country	Expected departures[b]	Using assumption 1[c]	Using assumption 2[d]	Incorrect formula[e]	Corrected formula[f]
North America					
Bahamas	112,717	928 (0.8%)	3,866 (3.4%)	6,889 (6.1%)	7,457 (6.6%)
Haiti	20,382	1,904 (9.3%)	3,238 (15.9%)	3,208 (15.7%)	3,461 (17.0%)
Mexico	399,355	7,098 (1.8%)	15,646 (3.9%)	18,697 (4.7%)	20,270 (5.1%)
Europe					
France	220,815	2,152 (1.0%)	1,786 (0.8%)	3,921 (1.8%)	4,300 (1.9%)
Germany	462,508	1,542 (0.3%)	1,257 (0.3%)	0 (0)	0 (0)
Poland	17,042	3,790 (22.2%)	6,038 (35.4%)	5,919 (34.7%)	6,378 (37.4%)
Soviet Union	22,220	2,308 (10.4%)	1,926 (8.7%)	2,181 (9.8%)	2,356 (10.6%)
Asia					
India	21,878	1,683 (7.7%)	1,753 (8.0%)	2,103 (9.8%)	2,272 (10.4%)
Philippines	29,905	3,337 (11.2%)	3,755 (12.6%)	3,946 (13.2%)	4,258 (14.2%)
Total nine countries	**1,306,822**	**24,742 (1.9%)**	**39,265 (3.0%)**	**46,864 (3.6%)**	**50,752 (3.9%)**
Total worldwide	**5,654,346**	[g]	[g]	**126,167 (2.2%)**	[g]

(Table notes on next page)

[a]These nine countries represent 37 percent of the overstays that INS estimated for tourist air arrivals. Tourist air arrivals account for 76 percent of all the overstays INS estimates. The sample data cover the period October 1990 to March 1991 and are from NIIS, which does not include (1) certain categories of Canadian visitors or (2) Mexicans with border-crossing cards that allow them to travel within 25 miles of the U.S.-Mexican border for periods of 72 hours or less. This table also excludes travelers who are not tourists—for example, visitors for business purposes and temporary workers—and land and sea arrivals. However, the majority of foreign visitors captured by the NIIS data are tourists who arrive by air. All estimates in this table were computed using estimates of system error based on the following index countries: Australia, Belgium, Finland, Netherlands Antilles, New Zealand, Norway, Singapore, Suriname, Sweden, and Switzerland.

[b]The percentage of expected departures on qualifying airlines were: Bahamas, 25.8; Haiti, 66; Mexico, 50.9; France, 83; Germany, 81.5; Poland, 23.1; Soviet Union, 15.8; India, 68.7; and the Philippines, 53.9.

[c]Assumption 1 is that visitors flying to the United States on nonqualifying airlines are characterized by the same overstay rate as visitors from the same country flying here on qualifying airlines.

[d]Assumption 2 is that visitors flying to the United States on nonqualifying airlines are characterized by the same rate of system error as visitors from the same country flying here on qualifying airlines.

[e]INS' calculations. In preparing these estimates, INS used its usual procedures and did not correct the computational error in its formula.

[f]Our calculations, based on INS' global estimate of system error and the correct formula, which is presented in appendix I. Results should be considered approximate because the system error (as calculated by INS) was rounded to the nearest tenth of one percent.

[g]GAO did not calculate worldwide overstay estimates.

Table 1 also presents estimates of overstays based on INS' global method; one column presents estimates calculated by INS using the incorrect formula, and the other column presents estimates that we calculated using INS' general global method but correcting the error in INS' formula.[22]

While results vary for individual countries, overstay estimates are, on average, lower when using the more detailed airline-by-airline method than when using the global method.[23] The contrast between results using the global method and corresponding results using the airline-by-airline method is greatest when the global method is applied using the correct computation formula.

[22]Neither of the two sets of estimates based on INS' global method was calculated using weighted data.

[23]A key reason is that the airline-by-airline approach recognizes that visitors from Mexico, the Philippines, and many other nations are much more likely to fly to the United States on American carriers than to use index countries' airlines such as Air New Zealand. The American carriers have relatively high rates of uncounted departures—that is, relatively high rates of "system error." (See table I.2 in appendix I.) The relatively high rates of system error translate to lower estimates of overstays, since the number of unrecorded departure forms is divided into system error (uncounted departures) and overstays.

GAO's Alternative Assumptions

The airline-by-airline overstay estimates shown in table 1 are, on average, lowest when passengers on nonqualifying airlines are assumed to have the same overstay rates as passengers from the same country who arrived on qualifying airlines.[24] Although the nine countries' overstay estimates based on assumption 1 (totaling about 25,000) are dramatically lower than INS', these might be the better of the two sets of airline-by-airline estimates.[25] However, taken together, our two sets of estimates for overstays from the nine sending countries (25,000 and 39,000) define a range that depicts the likely uncertainty of the estimates that arises from the lack of detailed data for visitors who fly here on nonqualifying airlines.

There is a specific logic for using the two alternative assumptions. That is, the known quantity for each sending country and airline is the number—or percentage—of unrecorded departure forms. If the percentage of unrecorded departure forms is higher for passengers on "nonqualifying" airlines than for passengers on "qualifying" airlines (as is the case for Mexico, for example), this difference is likely attributable either (1) to the fact that system error is higher for nonqualifying airlines (whereas equal overstay rates apply for passengers on both groups of airlines) or (2) to the fact that the overstay rate is higher for the nonqualifying airlines (whereas the same rate of system error applies) or (3) to some combination of higher system error and higher overstay rates for the nonqualifying airlines.[26] A range of estimates based on our two alternative assumptions covers these three possibilities.[27] For each country, the size of the range is determined by the empirically observed difference between the rates of unrecorded departure forms for qualifying and nonqualifying airlines.

[24]Specifically, our overstay estimate of 24,742 for the nine sending countries, based on assumption 1, is 47-percent lower than INS' estimate of 46,864. By contrast, our estimate for the same nine countries based on assumption 2 (39,265) is 16-percent lower than INS' estimate of 46,864.

[25]It may be more plausible to assume equal overstay rates for passengers from the same country than to assume equal rates of uncounted departures (system error) for passengers on the two types of airlines ("qualifying" airlines for which separate estimates are possible and "nonqualifying" airlines for which separate estimates cannot be made). Assuming equal rates of uncounted departures (system error) seems less plausible because (1) the airlines are responsible for collecting the forms and sending them to INS, and (2) as shown in table I.2 in appendix I, rates of system error vary across airlines—even for passengers from the same index country.

[26]For cases where the percentage of unrecorded departure forms is lower for nonqualifying airlines (for example, Germany), analogous logic applies. That is, the lower rates are likely to be explained by lower system error, lower overstay rates, or some combination of these two.

[27]Notably, however, the range does not include the possibility that where the percentage of unrecorded forms is higher, either system error or overstay rates could be lower.

Limitations and Potential Improvements

All overstay estimates require some assumptions—and are therefore subject to some degree of uncertainty. At the time we drafted this report, the only estimates other than those that we made were INS' estimates based on its global approach. The method we developed reduces the uncertainty relative to INS' global approach because the assumptions are more limited. However, it is important to realize that even the range for airline-by-airline estimates described above does not take account of all sources of uncertainty. There is variation in system error across index countries—even for passengers on a particular airline. And it is certainly possible that, even on a particular qualifying airline, system error might be higher (or lower) for passengers from sending countries (for example, the Bahamas or Mexico) than it is for passengers from index countries.[28]

Although we believe the procedures we devised are a step in the right direction, we also believe further improvements are possible. In responding to an oral briefing on our work, INS officials agreed that our method constitutes a step forward, and they suggested modifications that build upon—and that they believe might improve—our approach. Specifically, INS suggested increasing the number of index countries and using more detailed data on variation in system error by port of entry. Other expert reviewers provided a variety of suggestions, including using more complex models and conducting special studies to obtain empirical estimates of system error in nonindex countries.[29]

Recommendation to the Commissioner of INS

We recommend that the Commissioner of INS have new overstay estimates prepared for air arrivals from all countries, using improved estimation procedures such as those discussed in this report, including, as appropriate, the potential improvements suggested by INS or by reviewers of this report.

Agency Comments

INS agreed that data on the individual airlines could improve its estimates of overstays. INS developed a new version of the airline-by-airline approach and revised its criteria for selecting index countries—replacing its "original" set of index countries with a new set of 17. In its written comments, INS compared estimates obtained with its old and new

[28]Since differences in system error do exist for passengers from different index countries even when traveling on the same airline, it is likely that system error for individual sending countries differs from the index countries' estimate to some degree. And there is no way of knowing, at present, whether the differences occur about equally in each direction (and so would tend to balance each other, on average, worldwide) or whether they are predominantly in one direction.

[29]See pp. 38-40 for examples of the kinds of empirical studies that might be appropriate.

estimation procedures. [30] INS concluded that (1) the two methods produced differing estimates for certain countries, and (2) the worldwide estimate of overstays obtained with its new procedures (104,000 for the visitor category and time period in question) essentially matched the number estimated with its old procedures (107,000). More recently, INS told us that it will no longer use its previous approach and has adopted the airline-by-airline method (including the corrected estimation formula).

By adopting an airline-by-airline approach and correcting its formula, we believe INS has taken an important step forward. This is partly because achieving more accurate estimates of overstays from individual countries is essential for certain policy-related decisions—particularly those involved in administering the Visa Waiver Pilot Program. In addition, contrary to INS' conclusion, we believe that the agency's change to an airline-by-airline method may make a substantial difference in the worldwide estimate. As we explain in appendix II, depending on which sets of index countries INS uses, the worldwide overstay estimate achieved with airline-by-airline procedures may be 17- to 25-percent lower than the global estimate. (See pp. 56-58.)

Although INS' adoption of an airline-by-airline method is an important step in the right direction, we believe the improvement of overstay estimation methods is a "work in progress." As discussed in appendix II, we believe INS' selection of index countries deserves further study. Also, certain aspects of INS' airline-by-airline procedures differ from those that we used, and we have some potentially important concerns about those differences.[31] Finally, the experts who reviewed a draft of this report suggested possible strategies for reducing the uncertainty that characterizes both our estimates and INS', but INS has not had an opportunity to consider the experts' suggestions.[32] For these reasons, we believe our recommendation to the Commissioner of INS is still appropriate.

As agreed with your office, unless you publicly announce its contents earlier, we plan no further distribution of this report until 30 days from its date of issue. At that time, we will be sending copies of this report to the Immigration and Naturalization Service, the Department of State, and

[30]The full text of INS' comments and our detailed response are in appendix II.

[31]See pp. 59-60.

[32]See pp. 38-42 for the experts' suggestions.

other interested parties. We will also make copies available to others upon request.

If you or your staff would like to discuss any of the issues we present here, please call me at (202) 512-2900 or Judith A. Droitcour, who served as project director on this study, at (202) 512-5885. Other major contributors to this report are listed in appendix IV.

Sincerely yours,

Joseph F. Delfico
Acting Assistant Comptroller General

Contents

Figures

Abbreviations

CD	Counted departures
EXD	Expected departures
INS	Immigration and Naturalization Service
NIIS	Nonimmigrant Information System
OV	Overstays
R(OV)	Rate of overstay
TD	True departures
UCD	Uncounted departures
URDF	Unrecorded departure forms

Technical Appendix

This appendix presents detailed information on the following technical topics: (1) the estimation of "system error" (percentage of uncounted departures) and INS' "global" assumption that the level of system error is the same worldwide (across all airlines and all countries of citizenship); (2) the formula we developed for estimating overstays (given an estimate of system error) and an explanation of the error we identified in INS' computation formula; and (3) a new set of procedures for estimating overstays that, by incorporating data on specific airlines, avoids the need for INS' global assumption and, at the same time, uses the correct computation formula and appropriately weighted data. Additional sections of this appendix discuss (4) possible further improvements in estimation procedures and (5) problems with overstay estimation procedures for visitors arriving by land.

Background

Each year, millions of foreigners legally enter the United States as "nonimmigrant" visitors (that is, on a temporary basis and for a specific purpose, such as tourism). The INS Nonimmigrant Information System tracks records of legal nonimmigrant visitors by country of citizenship; visa category (tourist visit, business, or other purpose such as temporary work); mode of travel (air, land, or sea); and for each air passenger, the airline on which he or she entered the United States.[1] As indicated by the NIIS data, the large majority of foreign visitors are tourists who enter by air. (It is important to note that many Canadian visitors are not counted by this system. The same is true for many Mexican visitors who use border-crossing cards.[2] Data on foreign students entering the United States to attend school are maintained in a separate system and are not part of the estimates discussed in this report.)

INS also attempts to determine whether or not each foreign visitor (identified by name, date of birth, country of citizenship, and passport number) exits the United States by 9 months after his or her expected date of departure. INS' use of a minimum 9-month overstay period for purposes of calculating the overstay estimates discussed in this report means that transients who overstayed for only a few weeks or months are not included in these figures. Rather, INS defines estimated overstays as

[1]The NIIS data system includes a record of information that visitors provide on arrival/departure forms, including sex, age, and so forth; port of entry is also recorded in the data system. We note that an unknown number of persons fraudulently enter this country as tourists. If the fraud is successful, the record of such a person's seemingly legitimate entry would be tracked together with legal entries in the NIIS system.

[2]For Mexicans with border-crossing cards, travel is limited to an area within 25 miles of the U.S.-Mexican border for periods no longer than 72 hours.

foreign visitors who—having been here for a year or more—appear to have settled in the United States.[3] However, as explained below, there is (1) considerable error in counting visitor departures and thus (2) uncertainty in the estimation of overstays.

Briefly, when a foreign visitor legally enters the United States, he or she fills out an arrival/departure form (I-94 form; see figure I.1).[4] The arrival portion of the I-94 form is detached and retained by INS. The departure portion of the form is stapled to the visitor's passport. Each visitor is supposed to turn in the departure portion of the form when exiting the United States. The INS data system attempts to match the arrival with the corresponding departure portion of the form. Thus, it is possible to determine the number of foreign visitors for whom departure forms have not been recorded ("unrecorded departure forms").

In some instances, unrecorded departure forms correspond to visitors who overstayed. Often, however, departure forms have gone unrecorded because either (1) the forms were not turned in, not collected, or if collected by an airline, not returned to INS; or conceivably, (2) the forms were returned to INS but not correctly recorded as of a 9-month period following the expected departure date.

[3]For example, a tourist with a 3-month visa who overstays for 9 months has been in the United States for a year. Note that the procedures described here estimate the number of new illegal immigrants who are overstays. Certain analyses transform these estimates to estimates of the number of illegal immigrants—living here as of any one date—who initially came in legally and overstayed. To perform this transformation, it is necessary to estimate the number of overstays who returned to the home country before the date in question, the number who adjust to a legal status before that time, and so forth.

[4]Visitors from countries in the Visa Waiver Pilot Program use form I-94W. The current Visa Waiver Pilot Program countries are Andorra, Austria, Belgium, Brunei, Denmark, Finland, France, Germany, Iceland, Italy, Japan, Liechtenstein, Luxembourg, Monaco, Netherlands, New Zealand, Norway, San Marino, Spain, Switzerland, Sweden, and United Kingdom. Effective April 1, 1995, Ireland was also designated as a pilot program country, but with probationary status.

Figure I.1: Arrival/Departure Form I-94 for Foreign Visitors to the United States

Front

U.S. Department of Justice
Immigration and Naturalization Service

OMB 1115-0077

Admission Number

Welcome to the United States

601408904 00

I-94 Arrival/Departure Record - Instructions

This form must be completed by all persons except U.S. Citizens, returning resident aliens, aliens with immigrant visas, and Canadian Citizens visiting or in transit.

Type of print legibly with pen in ALL CAPITAL LETTERS. Use English. Do not write on the back of this form.

This form is in two parts. Please complete both the Arrival Record (Items 1 through 13) and the Departure Record (Items 14 through 17).

When all items are completed, present this form to the U.S. Immigration and Naturalization Service Inspector.

Item 7 -If you are entering the United States by land, enter **LAND** in this space. If you are entering the United States by ship, enter **SEA** in this space.

Form I-94 (04-15-86)Y

Admission Number

601408904 00

Immigration and Naturalization Service

I-94
Arrival Record

1.Family Name
2.First (Given) Name
3.Birth Date (Day/Mo/Yr)
4.Country of Citizenship
5.Sex (Male or Female)
6.Passport Number
7.Airline and Flight Number
8.Country Where You Live
9.City Where You Boarded
10.City Where Visa Was Issued
11.Date Issued (Day/Mo/Yr)
12.Address While in the United States (Number and Street)
13.City and State

Departure Number

601408904 00

Immigration and Naturalization Service

I-94
Departure Record

14.Family Name
15.First (Given) Name
16.Birth Date (Day/Mo/Yr)
17.Country of Citizenship

See Other Side **STAPLE HERE**

Back

This Side For Government Use Only

Primary Inspection

Applicant's Name
Date Referred ____ Time ____ Insp. # ____

Reason Referred

☐ 212A ☐ ☐ ☐ PP ☐ Visa ☐ Parole ☐ SLB ☐ TWOV

☐ Other

Secondary Inspection

End Secondary Time ____ Insp. # ____

Disposition

18. Occupation	19. Waivers
20. INS File **A -**	21. INS FCO
22. Petition Number	23. Program Number
24. ☐ Bond	25. ☐ Prospective Student

26. Itinerary/Comments

27. TWOV Ticket Number

Warning - A nonimmigrant who accepts unauthorized employment is subject to deportation.
Important - Retain this permit in your possession; ***you must surrender it when you leave the U.S.*** Failure to do so may delay your entry into the U.S. in the future.
You are authorized to stay in the U.S. only until the date written on this form. To remain past this date, without permmission from immigration authorities, is a violation of the law.
Surrender this permit when you leave the U.S.:
- By sea or air, to the transportation line;
- Across the Canadian border, to a Canadian Official;
- Across the Mexican border, to a U.S. Official.

Students planning to reenter the U.S. within 30 days to return to the same school, see "Arrival-Departure" on page 2 of Form I-20 **prior to surrendering this permit.**

Record of Changes

Port: **Departure Record**

Date:

Carrier:

Flight #/Ship Name:

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

Source: Immigration and Naturalization Service.

System Error

The fact that many visitors who actually leave the United States are not on record as having departed constitutes a major problem for overstay estimation. For air passengers, the airlines have been charged with collecting departure forms and transferring these to INS. The agency plans to begin a new airline data collection effort soon; while INS characterizes this effort as still "in the design phase," it may have some potential for improving INS air departure data. One other possible—but likely expensive—solution would be for INS officials to collect departure forms at airports.[5] For land travelers, the key reason why departure records are missing is that roadways exiting the United States lack facilities for collecting departure forms.[6]

Although uncollected I-94 departure forms seem to be the chief problem, some portion of uncounted departures may possibly result from other factors, such as the loss of forms after they have been collected, incorrect keypunching (which can prevent a match between entry and exit forms), or inadvertent deletion of information from a computer file.[7]

INS uses the term "system error" to refer to uncounted departures. Estimates of system error are not negligible and seem quite large relative to the number of overstays. Yet no one knows, with any precision, how large the system error component is. Consequently, any estimate of overstays is necessarily somewhat uncertain.

[5]We have not reviewed INS data collection plans in detail and therefore cannot comment on the likelihood that they will actually result in improved overstay information in the coming years. A bill sponsored by the Chairman, House Subcommittee on Immigration and Claims, would "establish a pilot program in which officers of the [INS] collect a record of departure for every alien departing the United States and match the records of departure with the record of the alien's arrival in the United States." (Title I, sec. 113, of H.R. 2202, 104th Cong., 1st sess.) The pilot program would be operated "at no less than 3 of the 5 air ports of entry with the heaviest volume of incoming traffic from foreign territories" and would provide indications of the cost of such a program nationwide.

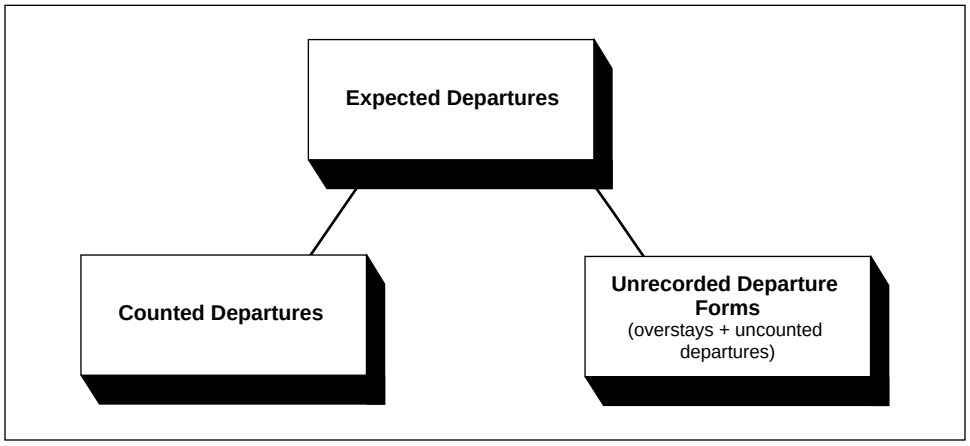
[6]See the section on land entries on pp. 42-43.

[7]INS recognizes that system error includes lost departure forms, inadvertent keying errors, and errors resulting from the fact that some extensions of stay are not entered into the NIIS data system. According to INS, a contractor currently keypunches the number on the arrival form, although departure forms are scanned. Some NIIS data were lost for fiscal year 1990 because of a computer error; only indirect estimates of overstays are therefore possible for fiscal year 1990 (although these limitations were not noted by INS in publications of overstay estimates). Other errors have been detected in fiscal year 1993 and 1994 data, and they are currently not usable for making overstay estimates. Errors that preclude the use of all data for certain fiscal years are conceptually distinct from the "system error" that pervades the data that must be used to obtain overstay estimates. Nevertheless, the fact that the larger errors occur alerts one to the possibility of substantial numbers of smaller computer errors in the data set. Although we did not collect information on "nonmatches" (departure form records that could not be matched to arrivals), we note that such information might inform the estimation procedure.

The estimation problem can be briefly described as follows. As shown in figure I.2, the three major known quantities in the INS Nonimmigrant Information System are

- the total number of expected departures, that is, the number of departures that should occur by the time the visitors' legal periods of stay (plus the subsequent 9-month period) have expired;
- the number of counted departures, that is, the number of foreign visitors who turned in departure forms that were received and correctly recorded and stored by INS;[8] and
- the difference between expected departures and counted departures, which we term "unrecorded departure forms."

Figure I.2: Known Quantities in INS Data on Foreign Visitors



In sum, the estimation challenge is that the "unrecorded departure forms" category includes both those foreign visitors who failed to depart (overstays) and those who did depart but were not counted as having departed (uncounted departures or system error). The data do not indicate how many foreign visitors with unrecorded departure forms fall into each of these two groups. It is clear, however, that being able to estimate the number of uncounted departures (system error) would allow the number of overstays to be estimated.

[8]Although the number of counted departures is taken as known, some uncertainty may pertain to this category. That is, some overstays (such as those who had entered with fraudulent documents or who had fraudulently obtained valid documents) might find a way to turn in their departure forms in order to create the impression that they had left—when in fact they had not.

INS Devised a Strategy to Estimate System Error

INS devised a creative strategy to estimate system error. The key is to examine the rates of unrecorded departure forms in "index" countries—such as Australia, Sweden, Finland, and Switzerland—for which INS believes it is safe to assume that there are only negligible numbers of overstays.[9] If this is so, then for visitors who are citizens of these (index) countries, the rates of unrecorded departure forms and the rates of uncounted departures (or system error) would, for all intents and purposes, be equal.[10]

Initially, INS defined 12 index countries: Australia, Belgium, Finland, Kuwait, Netherlands Antilles, New Zealand, Norway, Saudi Arabia, Singapore, Suriname, Sweden, and Switzerland. Later, Kuwait was dropped from the list (because of overstays resulting from the Gulf war).[11] INS reported that the 12 countries were selected according to five criteria, one of which is low rates of unrecorded departure forms.[12]

INS realized that the data for foreign visitors from the index countries provide a potential window onto the level of system error that occurs in other countries. INS applied this strategy separately for visitors who arrived by air, land, and sea. This approach involves an assumption that travelers who arrive by one mode of transportation generally depart by the same mode—for example, that visitors who arrive by air generally depart by air. This may be less likely for citizens of certain countries. Notably, Mexicans and Canadians who arrive by air may be more likely to depart by land than visitors from countries that do not border the United States.

[9]See Robert Warren, "Annual Estimates of Nonimmigrant Overstays in the United States: 1985 to 1988," in Frank D. Bean, Barry Edmonston, and Jeffrey S. Passel, Undocumented Migration to the United States: IRCA and the Experience of the 1980s (Washington, D.C.: Urban Institute Press, 1990), pp. 77-110.

[10]To the extent that visitors from index countries do overstay, system error would be overestimated. As one reviewer of this report pointed out, some visitors from index countries may be "technical overstays" in that if they had applied for visa renewal or other legal status, it probably would have been granted. This means that estimates of system error may include some "technical overstays." If rates of "technical overstays" are the same for index and nonindex countries, then rates of system error estimated for nonindex countries encompass the technical overstays. Such persons would be included in rates of system error rather than being counted as overstays (in the estimates presented in this report). However, because rates of technical overstays may differ across countries, this represents a source of uncertainty in all current estimates.

[11]In the sample data that INS provided to us for calculating airline-by-airline estimates, Saudi Arabia was not included. Therefore, in our trial estimates of the new procedures—and in the comparison estimates that we subsequently asked INS to provide to us based on the agency's usual procedures—both Kuwait and Saudi Arabia were omitted from the set of index countries.

[12]The other criteria are "(1) low backlogs for immigrant visas; (2) low numbers of applicants for legalization; (3) low numbers of alien apprehensions by the INS; (4) low estimates of undocumented aliens counted in the 1980 Census" (Warren, 1990).

INS Calculated System Error for All Airlines Grouped Together

Using the index country data for visitors who arrived by air, INS calculated a single estimate of system error that combined data for all index countries' air passengers and the airlines they used to fly into the United States.[13] By calculating a single rate of system error—and applying it globally—INS procedures assumed that this "single point estimate of system error applies to every country of origin within a particular category of admission." (Warren, 1990, p. 81.) An INS official noted that "it is possible that some areas of the world have different rates of system error. For example, some airlines might do a better job of emphasizing the collection of I-94 departure forms." (Warren, 1990, p. 81.)

In fact, certain index countries and airlines appear to be considerably more assiduous in collecting departure forms than others. Given this fact, INS' approach rests on assuming that the rate of system error estimated on the basis of index countries' air passengers (and the airlines that they use to fly into the United States) is, on average, at least reasonably representative of system error for air passengers from other countries—and the different airlines that they may use to fly into the United States. Of course, no data exist to show whether system error is the same or different in index countries and in other (nonindex) countries; this lack of data represents the major source of uncertainty in current overstay estimates.

The implications of using a global estimate of system error—as opposed to an airline-by-airline approach—are as follows. If the "typical airline" serving the index countries (that is, serving Australia, Switzerland, and so forth) is more assiduous in collecting departure forms than the typical airline serving a given nonindex country (for example, India), then system error is actually lower for the airlines serving the index countries than for the airlines serving India.[14] But the INS procedures would not account for this, so system error would be underestimated for India (or an analogous sending country). Consequently, INS' estimated number of overstays for India (or an analogous country) would be too high. The logical converse could also be true in some instances.

Because the airlines collect the forms, we believe they are a key factor in estimating system error. But various other factors may also play a role in

[13]Separate calculations were made for tourists and business travelers and for air, land, and sea arrivals. In these calculations, INS weighted data for each index country equally (rather than according to the numbers of visitors).

[14]We refer here to the potential difference between the rate of uncounted departures for visitors from index countries (that is, the average taken across those countries) and the rate of uncounted departures for visitors from the sending country.

determining the rate of system error for visitors from a particular country.[15]

INS Used the Average Plus One Standard Deviation

INS calculates its "global" estimate of system error by taking the average percent of unrecorded departures across the 11 index countries plus the standard deviation of the 11 numbers.[16] (See table I.1.) INS adds one standard deviation to the average in order to avoid underestimating system error.[17] Adding an adjustment of this sort seems to be justified because one criterion INS used to select the index countries was low rates of unrecorded departure forms. The fact that INS used this criterion means that the index countries may not only have zero (or near-zero) rates of overstays, but they may also have lower-than-average rates of system error. We believe, as INS does, that adding an adjustment may help to correct this possible bias. However, we have not assessed the adequacy of this corrective factor.

[15]For some countries, system error is likely to be different than for most other countries. For at least two reasons, system error may be higher for Mexico and Canada than for other countries. First, as previously noted, Mexicans and Canadians who arrive by air may be more likely to depart by land than air arrivals from index countries (which do not border the United States). Because system error is much higher for land departures, INS' estimate of system error for air arrivals from Mexico may underestimate of the true level of system error characterizing Mexican air arrivals. Second, the airlines' collection of departure forms may be more lax on flights to Mexico and Canada because regulations do not require collection of departure forms for foreign visitors making short "side trips" to Mexico or Canada during a longer visit in the United States. These considerations suggest that overstays might be overestimated for Mexico and Canada. The reason is that underestimates of system error translate directly to overestimates of overstays. (A somewhat similar situation might apply for visitors who arrive by air from one or more Caribbean nations because they may be more likely to depart by sea than visitors from index countries.)

[16]In these calculations, INS takes a simple, rather than a weighted, average. That is, INS gives large and small index countries equal weight.

[17]For the example shown in table I.1, this procedure produced an estimate of system error of 9.52 percent—that is, 8.19 percent rate for unrecorded departure forms in the index countries plus 1.33 percent representing the standard deviation.

Table I.1: INS' Computation of System Error[a]

Country of citizenship	Percentage apparent overstay
Australia	8.09
Belgium	10.21
Finland	5.71
Kuwait[b]	8.65
Netherlands Antilles	9.60
New Zealand	7.92
Norway	9.06
Saudi Arabia	8.36
Singapore	5.79
Suriname	7.91
Sweden	8.81
Switzerland	8.14

Note: Mean apparent overstay (average unrecorded departure forms): 8.19%; standard deviation: 1.33%; and estimated system error: 9.52%.

[a]Tourist air arrivals, October 1986-March 1987. INS calculated the mean with each country (large or small) given an equal weight.

[b]Kuwait was later dropped from the list of index countries (because the Gulf war increased the number of overstays), yielding 11 index countries.

System Error Varies

We believe a main limitation of the INS estimator is its global quality. That is, INS calculated one estimate of system error based on the airlines that index country passengers flew to enter the United States—and used this estimate for all sending countries and all airlines that citizens of the sending countries flew. But system error is not constant.

Examining index country data for one 6-month period, we found that the rate of unrecorded departure forms, which INS takes as the rate of uncounted departures in index countries, varies across airlines as well as across countries. (See table I.2.) Dividing the airlines into two groups—those with generally lower levels of system error and those with higher levels—it is apparent that the airlines with lower levels of system error tend to be based in index countries, mainly Australia and New Zealand or northern European countries. By contrast, those airlines with higher levels of system error are—in the majority of cases—American carriers based in the United States.

Table I.2: Variation in Percentage of Uncounted Departures (System Error) for Selected Index Countries[a]

	Index country							
Qualifying airline	Australia	New Zealand	Belgium	Finland	Singapore	Norway	Sweden	Switzerland
Lower levels of system error								
Air New Zealand-Int.	2.4%	3.2%						
Air New Zealand	2.9	3.5						
KLM Royal Dutch			3.5%				4.6%	
Qantas	4.3	4.9					4.0	
Icelandair						2.3%	5.7	
Singapore Airlines	5.4				3.1%			
SAS Scandinavian				3.3%		4.4	5.2	
Finnair				2.6			8.4	
Higher levels of system error								
Air Link Corp.				4.6		4.9	7.3	
American Airlines	4.4	4.8	5.8				6.8	7.3%
United Airlines	4.5	5.1			8.0			
Delta	4.9							7.0
British Airways	4.1	3.3		7.4		7.5	9.1	8.1
Air Micronesia	6.0	6.1					6.7	8.3
Pan American	5.8	6.1	7.2	6.0		9.6	10.6	13.1
Virgin Atlantic	8.7	9.1						
Northwest Orient					4.9		13.5	
TWA	6.8		14.1			12.9	19.1	18.5

[a]Based on tourist air arrivals from October 1990 to March 1991. Data are provided for cells representing 1,000 or more passengers. The table includes only those index countries with 1,000 or more passengers on each of two or more airlines and only those qualifying airlines with 1,000 or more passengers from each of two or more countries. It shows variation across countries for a given airline—and variation across airlines for a given country (for cells in which reasonably reliable estimates of system error can be made).

Further consideration indicates that, when flying between the United States and other countries, airlines in the group with lower levels of system error generally tend to fly mainly between the United States and Australia, New Zealand, or northern European countries. By contrast, the airlines with higher levels of system error not only fly from index countries to the United States, but also serve many other countries around the world (for example, countries in Latin America or Asia), bringing travelers from these diverse regions to the United States. On a worldwide basis, these airlines with higher system error would logically be more frequently used by visitors to the United States than airlines in the other group.

Potentially significant patterns like these suggest that, if an estimation procedure could be based on the specific airlines that are actually used by visitors from each foreign country, that procedure might improve estimates of system error. And improved estimates of system error would reduce the uncertainty associated with overstay estimates based on INS' global approach.

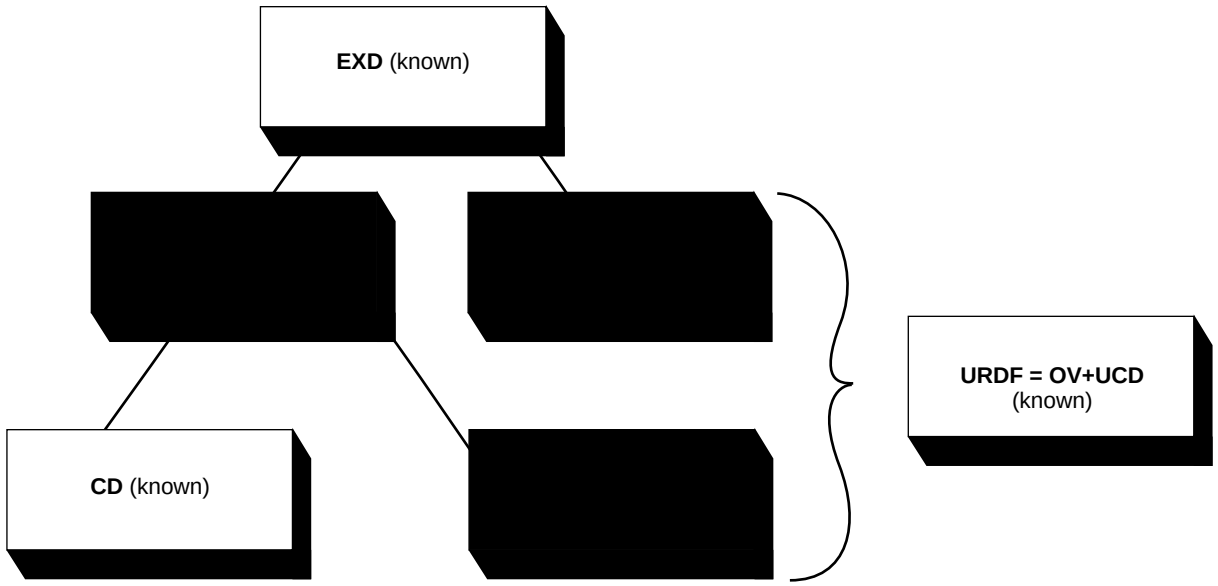
Using System Error to Estimate Overstays

INS uses its global system error figure to estimate the number of overstays from each sending country. In this section, we (1) present a formula for estimating the number of overstays, given an estimate of system error, and (2) contrast this formula with the INS computation formula (in which we found an error).

Figure I.3 provides the conceptual map needed to follow the ensuing discussion. Conceptually, expected departures may be divided into true departures and overstays. True departures can, in turn, be divided into counted departures (that is, visitors whose returned and recorded departure forms indicate that they have left the United States) and uncounted departures. The category of unrecorded departure forms includes both overstays and uncounted departures.[18]

[18]The categories shown in figure I.3 may be applied to foreign visitors in a specific visa category, arriving via a specific mode of travel (air, land, or sea) from a particular sending country. They could also be applied to visitors arriving on a particular airline, if an airline-by-airline approach were used.

Figure I.3: Conceptual Map for Estimating System Error



EXD = number of expected departures
TD = number of true departures
CD = number of counted departures
UCD = number of uncounted departures
OV = number of overstays
URDF = OV + UCD = number of unrecorded departure forms

Note: URDF is known because URDF = EXD – CD (from figure I.2).

Terminology for Computations

As previously explained, the known quantities are:

- the expected departures (EXD),
- the counted departures (CD), and
- the unrecorded departure forms (URDF) category, which is obtained by subtracting counted departures from expected departures.

True departures (TD), overstays (OV), and uncounted departures (UCD or system error) must be estimated.

As discussed in the previous section, the purpose of estimating system error using index countries' data is to estimate the percentage and number of uncounted departures for every sending country—and thereby to

estimate the number of overstays. But an appropriate formula must be used.

Two additional terms are needed for this discussion.

- First, as mentioned above, true departures can be divided into counted departures and uncounted departures. The symbol P refers to the true proportion of departures that are uncounted (that is, the true rate of system error). Specifically,

$$P = UCD/TD.$$

- Second, the symbol P′ refers to an estimate of system error (an estimate of P), defined as the average proportion of uncounted departures in the index countries plus one standard deviation. Specifically,

$$P' = \text{Average } (UCD/TD)_i + SD,$$

where the subscript i denotes index countries and SD refers to the standard deviation across the index countries.

That is, the rate of system error (P) is estimated using index countries' data. P′ is an estimate of the proportion of true departures that went uncounted. As appropriate, P can be estimated for all visitors arriving by air from all countries taken together (using all data on passengers from index countries) or for some subset of passengers such as those arriving on a specific airline (using data on index countries' passengers who arrived on that airline).

Appropriate Formula and INS Error

We derived the following formula to estimate overstays:

$$OV = EXD - (CD/(1 - P'))$$

Figure I.4 shows the derivation of this formula.

Figure I.4: Estimating Overstays for Any One Country

A. Overstays as a function of expected departures and true departures:

EXD = TD + OV	(from figure I.3)
OV = EXD – TD	(algebra)

B. To estimate true departures:

TD = CD + UCD	(from figure I.3)
TD – UCD = CD	(algebra: subtracting UCD from both sides of equation)
TD – (TD/TD) (UCD) = CD	(since TD/TD = 1)
TD – (TD) (UCD/TD) = CD	(algebra: multiplying terms)
TD – P(TD) = CD	(since, by definition, P = UCD/TD)
TD (1 – P) = CD	(algebra: factoring)
TD = CD/(1 – P)	(algebra: dividing both sides by (1 – P))

C. To estimate overstays:

OV = EXD – TD	(from A above)
OV = EXD – (CD/(1 – P))	(substituting last line of B above)
OV = EXD – (CD/(1 – P´))	(where P´ is an estimate of system error[a])

EXD = number of expected departures
TD = number of true departures
OV = number of overstays
CD = number of counted departures
UCD = number of uncounted departures

[a]P´ is calculated based on index country data, in order to estimate P in other countries. P´ is calculated using the average value of UCD/TD for index countries plus one standard deviation. This adjustment seems justified because index countries may have lower values of P than other countries.

INS estimated overstays using the following incorrect formula:

$$OV = [(URDF/EXD) - P']EXD)$$

where OV refers to the overstay estimate for a sending country (visa category and mode of travel), URDF refers to the number of unrecorded departure forms for that country, EXD refers to the number of expected departures for that country, and P′ refers to the estimated rate of system error, based on the index countries data. (See Warren, 1990, p. 106.)

As shown in figure I.5, the incorrect INS formula reduces to

$$OV = EXD - CD - P'(TD + OV)$$

or

$$OV = EXD - CD - P'(TD) - P'(OV)$$

This result indicates that when INS subtracts out the number of uncounted departures attributable to system error, it actually subtracts out not only the proportion of true departures that went uncounted, but also—wrongly—the same proportion of overstays. Assuming P′ is an adequate estimate of the rate of system error, the effect of using the incorrect INS calculation formula shown here is to underestimate the number of overstays.[19]

Figure I.5: Explication of the Error in INS' Formula

OV = [(URDF/EXD) – P´](EXD)	(INS formula (Warren, 1990, p. 106))
OV = [((EXD – CD)/EXD) – P´](EXD)	(substituting EXD – CD for URDF, based on definition of URDF in figure I.2)
OV = [(EXD – CD)/EXD](EXD) – P´(EXD)	(algebra: multiplying EXD by the two terms inside parentheses)
OV = EXD – CD – P´(EXD)	(canceling out the EXDs in the first term on the right-hand side of the equation)
OV = EXD – CD – P´(TD + OV)	(substituting TD + OV for EXD, based on figure I.3)

OV = number of overstays
URDF = number of unrecorded departure forms
EXD = number of expected departures
P´= estimate of system error
CD = number of counted departures
TD = number of true departures

[19]To illustrate, suppose there are 1,000 expected departures for a country and 800 counted departures. Further suppose that the estimate of P (system error) is 0.10. Using GAO's formula, we have 1,000 – 800/(1 – 0.10) = 1,000 – (800/0.90) = 1,000 – 889 = 111 overstays. Using INS' incorrect formula produces a different result: 1,000 – 800 – 0.10(1,000) = 1,000 – 800 - 100 = 1,000 – 900 = 100 overstays.

GAO Method for Overstay Estimation

To obtain an estimate of overstays that takes advantage of INS data on the specific airlines that foreign visitors from different countries actually use, we developed the four-step approach described below. Three points regarding the assumptions made and not made by our method are crucial:

- First, our method does not assume that a single estimate of system error characterizes each country of the world, regardless of which airlines are used. It does assume that the separate level of system error estimated for each airline that serves one or more index countries (for example, American Airlines, Delta, or British Airways) is about the same for passengers from index countries and for passengers from each nonindex country.[20]
- Second, our method does assume that system error can be adequately described by the airline of entry to the United States.[21]
- Third, with respect to nonqualifying airlines, which do not serve index countries, our method does require that an assumption be made. Specifically, we used two alternative assumptions as follows: We developed one set of estimates based on the assumption that for visitors from a particular sending country (for example, India), the overstay rate is the same for (1) passengers entering the United States on qualifying airlines, such as American Airlines or Delta, which serve one or more index countries, and (2) passengers entering the United States on other (nonqualifying) airlines. We also developed a second set of estimates based on the very different assumption that—again for visitors from a particular sending country—the rate of system error is the same for (1) passengers entering the United States on qualifying airlines and (2) passengers entering on nonqualifying airlines. (Essentially, where it was not possible to use information on the specific airline, we turned to available information on the specific country.)

Three caveats are appropriate: First, there is variation in system error across countries—even for a given airline (refer to table I.2 on p. 25); thus, it is possible that, even for a specific airline, system error may be different for passengers from index countries and from nonindex countries (either

[20]Specifically, we assume that for each airline of entry, the weighted average of the percentage of uncounted departures for index countries' passengers (plus one standard deviation) represents system error for that airline worldwide.

[21]This assumption is required because it is not possible to know what airline a traveler's departure form was not counted on—or what airline he or she did not depart on.

individual nonindex countries or nonindex countries in general).[22] Second, there is not a perfect correlation between the airline upon which a passenger enters and departs from the United States; and in fact, logic suggests that for Mexico and Canada, air passengers may be more likely to depart by land than would air passengers from index countries (since these do not border the United States)—suggesting that system error may be higher for citizens of Mexico and Canada than the levels that are estimated using an index countries' methodology.[23] Third, overstay rates—as well as rates of system error—may differ for travelers from the same country who arrive on qualifying and nonqualifying airlines.[24]

Despite these caveats, we believe the assumptions we have made are more plausible than the global assumption that INS made concerning system error.[25] The four steps of the approach we devised are described below and illustrated in table I.3.

[22]This is partly because variables other than the specific airline, such as port of entry, may play a role in determining levels of system error. Also, travelers from certain countries may be more proactive in turning in their departure forms than travelers from other countries. Factors such as these represent potentially key sources of uncertainty in estimating overstays when using an index countries' methodology.

[23]As described in a later section of this appendix, there are no facilities for collecting forms for land departures; therefore, the rate of uncounted departures is much higher for land departures.

[24]See p. 10 for a brief description of the interrelationship of the two alternative assumptions we make for nonqualifying airlines. Also presented there is an explanation of why, when taken together, these two alternative assumptions define the likely level of uncertainty attributable to the fact that not all passengers arrived on qualifying airlines.

[25]Recall that our assumptions are limited to passengers traveling on the same airline or to visitors from the same foreign country.

Table I.3: GAO's Airline-By-Airline Method for Estimating Overstays

	Visitors from index countries[b]	
Visitors arriving on qualifying airlines[a]	**Yes**	**No**
Yes	Step 1: Estimate the rate of uncounted departures (system error) for each qualifying airline, using data on visitors from index countries who entered the United States on qualifying airlines.	Step 2: Use step 1 estimates of uncounted departures (system error) together with the appropriate formula to estimate overstays among visitors from each nonindex country who arrived on each qualifying airline. Step 3: For each nonindex country: (a) combine the step 2 overstay estimates for qualifying airlines to obtain the overstay rate for visitors from that country who arrived on qualifying airlines; and (b) combine the step 1 estimates of system error for qualifying airlines to obtain the rate of system error for visitors from that country who arrived on qualifying airlines.
No	Not applicable.[c]	Step 4: For each nonindex country: (a) project the step 3a overstay rate to visitors from the same country who entered the United States on nonqualifying airlines, which produces one set of overstay estimates; and (b) project the step 3b rate of system error to visitors from the same country who entered on nonqualifying airlines, which produces a second set of overstay estimates.

[a]Conceptually, qualifying airlines are defined as those that regularly fly between at least one index country and the United States. Operationally, we defined qualifying airlines as those with at least 1,000 passengers from a single index country in a 6-month period.

[b]Visitors from index countries (such as Sweden and Switzerland) are thought to include virtually no true overstays. Data on visitors from these countries are used to estimate rates of uncounted departures, airline by airline. The nine countries for which we obtained overstay estimates are all nonindex countries. (We note, however, that INS does obtain overstay estimates for index as well as nonindex countries.)

[c]Data on visitors from index countries are not used to estimate rates of uncounted departures for nonqualifying airlines.

Step 1

Using data on index countries' air passengers, we estimated system error separately for each "qualifying airline." Conceptually, we defined a qualifying airline as one (such as American Airlines, Delta, British Airways, SAS, SwissAir, or Qantas) that regularly serves one or more index countries—and has sufficient passengers to allow a reasonably reliable estimate of system error for that airline. Operationally, we defined a qualifying airline as one that, during the 6-month period we examined, carried at least 1,000 passengers from a single index country to the United States.[26] In the sample data INS provided to us, we found that, according to our operational definition, there were 29 qualifying airlines.

To estimate system error for a specific qualifying airline, such as American Airlines, we used the index countries' data for that airline alone. For example, to estimate P_{AA}, we first calculated the proportion of uncounted departures (system error) for each index country—Australia ($P_{AUS/AA}$), Switzerland ($P_{SWZ/AA}$), Norway ($P_{NOR/AA}$), and so forth; then, we took the weighted average of these figures.[27]

To complete step 1, we also calculated the standard deviation of the proportion of uncounted departures across index countries for a specific airline. We then took, as the estimate of system error for a specific airline, the weighted average plus one standard deviation.

Step 1 was repeated for each of the 29 qualifying airlines.

Step 2

For a specific foreign country, such as Mexico, we estimated the number of overstays who had come into the United States on each qualifying airline. That is, we estimated overstays for American Airlines ($OV_{MEX/AA}$), Delta ($OV_{MEX/DEL}$), and so forth—performing the overstay calculations for each of the 29 qualifying airlines that transported Mexican visitors into the United States.

[26]For each qualifying airline, we used data from all index countries. Alternative operational definitions are, of course, possible. In selecting the level 1,000, we were mindful of the following considerations. First, the cutoff of 1,000 passengers ensured sufficient data to support an estimate of system error for each qualifying airline. Second, given the set of 11 index countries that INS had selected, the level of at least 1,000 passengers from a single index country ensured that each qualifying airline probably regularly served at least one index country, so that the data are not likely to be anomalous. The cutoff of 1,000 is appropriate, in part, because of the size of INS' original set of index countries (Sweden, Switzerland, and so forth). Were INS to select much larger index countries, a higher cutoff would probably be required to avoid including airlines that do not regularly serve index countries.

[27]In this way, an index country with few visitors using a particular airline to enter the United States (for example, few Australian visitors using SwissAir) would contribute only a small amount to the estimate of P for that airline. In other words, by using weighted data, we avoided the "small cell problem."

There are two points to note with respect to these procedures.

- First, within a particular sending country, the number of overstays who had arrived on each qualifying airline was estimated using the index countries' estimate of system error developed for that particular airline. For example, in estimating $OV_{MEX/AA}$, we used P'_{AA}, and in estimating $OV_{MEX/DEL}$, we used P'_{DEL}, and so forth.
- Second, in estimating $OV_{MEX/AA}$, $OV_{MEX/DEL}$, and so forth, we used the correct estimation formula specified in figure I.4:

$$OV = EXD - (CD/(1 - P'))$$

where P' is the estimate of system error. We applied the formula separately for passengers traveling on each qualifying airline within each sending country. For example, to estimate Mexican overstays who arrived in the United States via American Airlines, we used the formula:

$$OV_{MEX/AA} = EXD_{MEX/AA} - (CD_{MEX/AA}/(1 - P'_{AA}))$$

Step 2 was repeated separately for each qualifying airline within each country.

Step 3

Step 3 includes two alternative procedures (steps 3a and 3b), each of which is directed toward producing a separate set of estimates.

In step 3a, we totaled estimated overstays across qualifying airlines, within each country, and calculated the overstay rate for passengers on all qualifying airlines in that country. Specifically, we divided the estimated total number of overstays who had arrived on qualifying airlines by the total number of expected departures for those airlines—again, for a particular country.[28]

In step 3b, we calculated the rate of system error, within each country, for visitors who arrived on qualifying airlines.

Step 4

Step 4 also has two components (steps 4a and 4b); again, each component is directed toward obtaining a separate set of estimates.

[28]These procedures are the equivalent of calculating a separate overstay rate for passengers arriving in the United States on each qualifying airline, within each sending country, and then taking the weighted average of those rates across qualifying airlines.

In step 4a, we projected the step 3a overstay rate to visitors from the same country who entered the United States on nonqualifying airlines. In other words, we used the step 3a overstay rate as the estimated overstay rate, that is, R(OV), for the entire country.

Then, to obtain the estimated number of overstays for each country, we multiplied the country-specific overstay rate times the total number of expected departures for that country. (The total number of expected departures includes expected departures for visitors who had arrived in the United States on qualifying airlines and for those who had arrived on other, nonqualifying airlines.[29]) This produced one set of overstay estimates; expressed statistically, the number of overstays for the k^{th} sending country (air arrivals) is estimated as shown in figure I.6.

In step 4b, we performed analogous procedures using rates of system error rather than overstay rates. This produced a second set of alternative overstay estimates.

[29]We also performed a check at this point. Specifically, we compared the overstay estimate for each sending country to the number of unrecorded departure forms for that country. In all instances, the estimated number of overstays was less than the number of unrecorded departure forms. (Logically, the overstay estimate must be less than or equal to the number of unrecorded departure forms, since the number of unrecorded departure forms—which is known—includes both overstays and uncounted departures.)

Figure I.6: Statistical Expression of the New Overstay Estimator for Air Arrivals, Using Assumption 1

Estimator for the number of overstays from the kth sending country:

$$EST\ (OV_k) = [EST(R(OV)_k)](EXD_k)$$

where $EST(R(OV)_k)$ refers to the estimated overstay rate for the kth sending country and EXD_k refers to the number of expected departures for that country.

Estimator for the overstay rate for the kth sending country:

$$EST(R(OV)_k) = \sum(OV_{jk}) / \sum(EXD_{jk})$$

where OV_{jk} refers to the estimated number of overstays for the jth qualifying airline and the kth sending country, and the sum is taken across all qualifying airlines within the kth country; and where EXD_{jk} refers to the number of expected departures for the jth qualifying airline and the kth country, and again, the sum is taken across all qualifying airlines within the kth country.

Estimator of the number of overstays for the kth sending country and the jth qualifying airline:

$$EST\ (OV_{jk}) = EXD_{jk} - [CD_{jk} / (1 - P_j')]$$

where CD_{jk} refers to the counted departures for visitors from the kth sending country who arrived on the jth qualifying airline, P_j' refers to the estimated system error for the jth airline, and the formula is from figure 1.4.

Estimator for the system error for the jth qualifying airline:

$$P_j' = [\ \sum W_{ji} P_{ji}] + SD_{P_{ji}}$$

where P_{ji} is calculated as $URDF_{ji} / EXD_{ji}$, that is, as the number of unrecorded departure forms for the jth airline and the ith index country divided by the number of expected departures for the jth airline and the ith country, W_{ji} is $EXD_{ji} / \sum EXD_{ji}$, where the sum is taken across all index countries for the jth airline, and $SD_{P_{ji}}$ is calculated as the square root of the following expression: $\sum W_{ji} [P_{ji} - (\sum W_{ji} P_{ji})]^2$

Other Possible New Approaches

In response to our oral presentation of the new procedures, INS suggested some additional possibilities for improved approaches to estimating overstays. Expert reviewers of a draft of this report also made suggestions.

INS' Suggestions

One alternative approach would be a modification of the separate-airlines procedures we presented above. Specifically, INS suggested an approach in which it would identify additional index countries, thus expanding the number of "qualifying airlines." The object (and benefit) would be to increase the percentage of passengers from each sending country who arrive in the United States on qualifying airlines. In other words, the likely effect would be to substantially reduce the number of persons for whom it would be necessary to assume that the overstay rate (or, alternatively, the rate of system error) is the same as on the qualifying airlines. We believe that increasing the number of index countries could be a promising approach if appropriately implemented. But the issue of how to define specific criteria for index countries is a complex one and has not been addressed in this report.

Another approach that INS suggested would involve estimates of system error that adjust for different ports of entry (for example, Miami, New York, and so forth) as well as different airlines. If this approach proves practicable, it could also be used in conjunction with adjustments for different airlines. That is, airline and port of entry could both be taken into account.[30]

Experts' Suggestions

Other approaches or variations were suggested by expert reviewers, including ideas for small-scale empirical studies, special analyses, and complex estimation models.

Empirical Studies

Two experts suggested that it might be possible to conduct special studies designed to estimate system error empirically. Empirical studies might consist of relatively small-scale data collections, such as those used by the Bureau of the Census to estimate the undercount in the decennial census.[31]

[30]A log-linear analysis would allow simultaneous adjustments for airline and port of entry, as well as other factors. (For a discussion of log-linear analysis, see Y.M.M. Bishop, S.E. Fienberg, and P.W. Holland, Discrete Multivariate Analysis: Theory and Practice (Cambridge, Mass.: MIT Press, 1975).)

[31]One expert pointed out that although a complete census of cases (such as INS' Nonimmigrant Information System) can be useful administratively, it may not be optimal for research purposes. Sample surveys—with a higher per-case cost—can provide higher quality data than a census (for example, a lower rate of missing information). Because a relatively small number of cases are involved, the total cost of sample surveys can be low relative to the generalizable information that is provided.

Empirical studies may be especially important for certain nonindex countries such those where system error may be atypically high (for example, Mexico and Canada),[32] those where the majority of visitors fly to the United States on nonqualifying airlines, and so forth. There are three possibilities for empirical studies: (1) "planted travelers" studies, (2) follow-up studies of samples of visitors whose departure forms were not recorded, and (3) airport surveys of passengers flying out of the United States.

A "planted travelers" study would utilize persons hired by INS as "plants" who would exit the United States by air, using passports and departure forms supplied by INS.[33] The agency could subsequently check its data set to determine whether records of departure forms for these persons were included. Assuming that the planted travelers were instructed to be passive rather than proactive in turning in the forms, the results could be used to estimate a maximum rate of system error for the airlines or countries in question. Depending on the number of planted travelers and the design of the study, system error estimates could be obtained for selected countries, for the system as a whole, or both.

Another type of empirical study would be a "follow-up" study based on a random sample of perhaps 400 unrecorded departure forms from a specific country, airline, or other defined group. These unrecorded departure forms could be followed up in a number of ways, depending on the country in question. The initial follow-up in all cases could consist of additional checks of INS records.[34] Then, for countries that maintain records of reentering citizens or residents, it might be possible to determine that, in some instances, the country of origin's files contained a record of the traveler's return. It also might be feasible to attempt to contact individual travelers either in their home country (after their return home) or in the United States (if they overstayed).[35] Contacted persons

[32]As previously explained, visitors from Mexico and Canada who arrive by air may depart by land, and system error is higher for land departures. Also, since I-94 forms are not collected for visitors from other continents who are taking a short side trip to Mexico or Canada, collection of other visitors' I-94 forms may be less complete on flights from the United States to Mexico or Canada than on other departing flights.

[33]The idea for the planted travelers study was suggested by a similar study that was aimed at assessing the undercount of the homeless in the decennial census. (See Kim Hopper, "Counting the Homeless: S-Night in New York," Evaluation Review, 16 (1992), pp. 376-88.)

[34]For example, one check could be whether the individual has applied for legal status in the United States; if so, he or she would be categorized as an overstay. Another would be whether he or she reentered the United States before the time of expected departure.

[35]A Department of State representative at the U.S. Embassy in Mexico City told us that the addresses of visa applicants are kept by the U.S. Department of State visa-issuing office for 1 year.

might be asked a few brief questions about trips in the past 2 years (airline used, route flown, month of trip) as well as personal data (date of birth). The answers to the questions could be used to establish that the correct person had, in fact, been contacted. Random subsampling of nonrespondents for further follow-up attempts might also be appropriate. Even though some unrecorded departure forms may remain "unsolved," the percentage in question would, in all likelihood, be greatly reduced.[36]

If a general follow-back study of, for example, Mexican visitors proved too difficult, it still might be possible to follow-back a preselected subgroup, such as Mexican visitors aged 50 or older. This would be extremely useful if analyses of index countries' data indicated either that (1) system error did not vary substantially across age groups, or (2) system error varied across age groups according to a reliable pattern.

Yet another type of empirical study would involve surveying passengers on selected flights departing the United States. For example, it might be possible for INS inspectors to check passports immediately before passengers board. This would reveal the number of departure forms that were not collected by airline personnel. Or, a complete listing of departing visitors could be made and subsequently compared to records in INS' data system. Still other types of empirical studies might be developed. But the key point is that the option of fielding small-scale empirical studies has the potential to provide actual data on system error in nonindex countries.[37]

Empirical studies might make it possible to avoid difficult assumptions about system error in nonindex countries and about nonqualifying airlines. At a minimum, the data that these studies could provide would greatly reduce the effects of possible bias on the results.

Special Analyses

The expert reviewers also pointed to various ways to utilize existing INS data. One analytic approach would be to look for a variable that is not related to system error—but that might be related to overstays.[38] Such a variable could form the basis of a new index. Suppose, for example, that

[36]See National Research Council (U.S.) Panel on Census Requirements in the Year 2000 and Beyond, Modernizing the U.S. Census (Washington, D.C.: National Academy Press, 1995).

[37]While this discussion has focused on studies in selected countries, a global survey might also be possible. If an empirically based worldwide estimate of system error could be obtained, this would provide a basis for an empirically based all-countries estimate.

[38]In building models to explain or predict system error, one would look for variables—such as the specific airline—that are related to system error; but another approach would be to check index countries' data for variables for which system error does not vary.

system error does not vary across age.[39] Then, if it is possible to identify an age group with virtually no overstays (one possible candidate would be persons over 50), persons from that age group could be used as the index for their own country and airline of arrival. And still assuming that system error does not vary across age groups, even if there were no particular age group for which it could be assumed that overstays were virtually nonexistent, a similar approach could nonetheless be used if additional relevant data could be obtained. That is, if an empirical estimate of system error could be obtained for at least one age group (for example, persons over 50) in a small-scale empirical study, this estimate could be applied to other age groups. (Refer back to the discussion of empirical studies above.) Of course the variable, age, is used here for illustrative purposes.

Other analytic approaches were also suggested. For example, when frequent travelers reenter the United States, it is sometimes evident that their previous departure forms were not collected. (An analysis of these cases might indicate a rate of uncounted departures that would be considered a minimal rate if one believes that frequent international travelers are more knowledgeable and proactive about turning in their forms.[40]) Estimates could be obtained separately for different countries, airlines, or other subsets of visitors.

Sensitivity analyses were also suggested as a way of examining a variety of analytic options. For example, as one expert emphasized, negative estimates for specific airlines or countries occur when the rate of unrecorded departure forms is lower than the estimated rate of system error. Both we and INS have reported these negative estimates as zero overstays. However, one reviewer suggested that because a random component may influence these estimates, the overall estimates should be calculated without converting negative numbers to zero—thus reducing the number of overstays estimated when airlines and countries are combined. Whether issues like this are relatively minor or might have a substantial impact could be explored in sensitivity analyses.

[39]In fact, this is an empirical question that can be examined through index country data.

[40]Another possibility might be to examine data on recorded departure forms for which no arrival form could be found; these data might be used to estimate the number of arrival forms that are somehow lost or unrecorded even if turned in. If optical scanning is used in the future to record the arrival form number as well as the departure form number, the estimate of lost or misrecorded arrival forms could be used to estimate the corresponding number of lost or misrecorded departure forms. Balancing equations, analogous to those commonly used in making demographic estimates of population changes over time, might be useful in quantifying different components of system error. (Missing arrival forms might also be estimated by counting "breaks" in the consecutively numbered arrival forms that are recorded in INS' database.)

Estimation Models

In the absence of empirical studies such as those discussed above, a more complex estimation model might improve on the airline-by-airline method. Indeed, instead of using only adjustments for airline of arrival (as we did) or airline plus port of entry (a possibility that INS suggested), experts pointed out that it is possible to adjust for other background variables (for example, age and sex). One reviewer indicated that an optimal model would adjust for as many measured variables as are relevant.[41] Another said that the model should include enough variables to sufficiently account for variation in system error across countries. (As previously discussed, there is variation in system error across countries—even for a specific airline; refer to table I.2 on p. 25.)

Overstay Estimates for Land Entries

Within the INS data set being considered here, overstays who entered the United States by land and sea are much less numerous than those who entered by air. However, a sizable number of Mexican visitors enter by land—and for Mexicans, the estimated number of overstays entering by land is not negligible.[42] System error is much higher for land entries than for air arrivals because INS does not have facilities for collecting the I-94 forms when visitors exit by land. Nevertheless, some I-94 forms are collected for persons who arrived by land because (1) some visitors mail the forms in after returning home or turn the forms in to a U.S. embassy; (2) other visitors reenter the United States, and the INS inspector collects the old departure form at that time; and (3) still other visitors who arrived by land depart by air.

Estimates of overstays who arrived by land, like the air arrival estimates, rely on estimates of system error based on unrecorded departure forms for visitors from index countries. Again, INS' assumption is that a single estimate of system error, based on the experience of visitors from Australia, Switzerland, and so forth, applies to every sending country. For the land entries that are captured by NIIS (comprised chiefly of Mexicans), this seems especially implausible to us. (Specifically, INS assumes that the experience of visitors from countries like Australia and Switzerland who entered the United States by land is the same or at least similar to that of Mexicans who entered the United States by land.)

[41]See Roderick Little and Donald Rubin, Statistical Analysis With Missing Data (New York: Wiley, 1987), and Donald B. Rubin, Multiple Imputation for Nonresponse in Surveys (New York: Wiley, 1987).

[42]For example, during the 6-month period from October 1990 to March 1991, I-94 entry forms were processed for over 50,000 Mexican tourists arriving by land—and INS estimated that more than 5,000 of these persons became overstays. Applying the correct computational formula would raise this figure substantially.

The key issues in estimating overstays among land entries concern (1) whether the system error estimates for foreign visitors from index countries can fairly be applied to Mexican visitors arriving by land, and more generally, (2) whether the lack of facilities to collect departure forms when visitors exit by land means that overstay estimation for land arrivals is ill-advised, given current methods and data.

One possibly troubling situation is that a number of tourists from index countries may enter the United States from Mexico by land and then fly home to Sweden, Switzerland, and so forth. Such visitors may comprise a relatively large proportion of the number of index country visitors who arrive by land. The possible result is that index country visitors who arrive in the United States by land may—in some fairly large proportion—depart by air. If, by contrast, Mexicans who enter by land generally also depart by land, the use of the index country visitors' rate of system error would be inappropriate.[43] Thus, the overstay estimates for land arrivals are potentially misleading. We realize that this argument is somewhat speculative, but it highlights the possible need for improved procedures for estimating overstays who did not arrive by air—and illustrates the kinds of problems that might be assessed.

[43]Since the departure record has data fields for the carrier (and flight number), it may be possible to empirically estimate the number of land arrivals who were recorded as having departed by air.

Comments From the Immigration and Naturalization Service

Note: GAO comments supplementing those in the report text appear at the end of this appendix.

U.S. Department of Justice
Immigration and Naturalization Service

Office of the Commissioner

425 I Street NW.
Washington, DC 20536

JUN 22 1995

Mr. Robert L. York
Director, Program Evaluation in Human Service Areas
Program Evaluation and Methodology Division
General Accounting Office
Washington, D.C. 20548

Dear Mr. York:

This letter responds to your May 10 request to the Attorney General for comments on the draft report **ILLEGAL IMMIGRATION: INS Overstay Estimation Methods Need Improvement.** I am pleased that my staff in the INS Office of Policy and Planning has cooperated extensively during your review and contributed invaluable resources and suggestions to your methodological ideas. I place high priority on INS' capacity to make rigorous estimates of the illegal immigration population and plan on remaining leaders in promoting a broad discussion on how to improve these efforts.

As you may know, we have been working on various methods to improve estimation procedures, including your specific suggestion to utilize information for each airline. We conclude from our current analyses that revised estimation procedures validate previous INS overall estimates of overstays. For "All country" totals, revised INS procedures that incorporate the GAO approach yield a total of 104,199 overstays. The estimate is very close to our original total of 107,412 for tourists arriving by air during the October 1990 to March 1991 period.

We also believe that your suggestion for using information for each airline improves INS estimates for overstays from any particular country. Our belief, however, results only after developing new procedures to make estimates for the substantial proportion of nonimmigrant arrivals that could not be estimated reliably using the data and methodology contained in your report. I hope the more technical comments that follow prompt continued cooperation among our staff.

As stated in your report, the central problem in estimating the number of nonimmigrant overstays is how to divide the apparent overstays (non-matched I-94 arrival/departure forms) into system error (mostly due to incomplete collection of I-94 departure forms by the airlines) and actual overstays. In the 1980's, the INS developed techniques designed to estimate the average rate of system error for all countries. In that estimation procedure, the average rate of error is subtracted from the rate of apparent overstay for each country; the residual is the rate of actual overstay. Using that method, if the average rate of system error were estimated accurately, the total number of overstays would be correct, even though the estimate for any country would be in error to the extent

that that country had a higher or lower rate of system error than the overall average. The revised INS estimates indicate that, at least for this set of data, our original method produced reasonably accurate estimates of the average level of system error. The potential contribution of the GAO proposal is that it enables us to be accurate for individual countries as well as for total overstays.

The GAO's suggestion for incorporating nonimmigrant data for selected airlines into our procedure for estimating overstays arriving by air will be useful for refining our estimates for each country. However, our review of your proposals, along with the data that we compiled in the process of deriving our revised estimates, demonstrate that the estimation procedure described in the GAO draft report is only partially adequate. The GAO alternative estimates displayed in Table 1 rely on a critical underlying assumption that has little empirical or theoretical support. With the information now available, we believe your assumption is incorrect and unnecessary. The estimates presented in your Table 1 require further review and revision and would wisely be carefully re-examined before release to prevent unnecessary confusion.

The GAO draft report includes two procedures for deriving alternative estimates. The results that GAO considers the "better estimates" are displayed in Table 1 of the report. The second GAO procedure, also incorporating airline data in the estimation, produces a total estimate for the nine selected countries that is more than 50 percent higher than the first. Neither procedure produces tenable outcomes because both are partially the result of using unsubstantiated assumptions, as described below.

GAO first derived estimates of system error for the 29 airlines that flew 1,000 or more passengers from any of the 12 "index" countries, that is, countries that were determined by INS to have extremely few actual overstays. For nonimmigrants who arrived on the 29 airlines, estimates of overstays were derived using the level of system error computed for each airline. The limitation of the GAO approach presented in the draft report is that GAO had no basis for estimating overstays of the substantial number of nonimmigrants who arrived on other airlines. For some important countries (Mexico, for example), a sizeable number of their citizens fly on airlines for which GAO could not estimate system error. For every country, GAO assumed that the overstay rates for those who flew on airlines, for which estimates could not be derived, were identical to the overstay rates of nonimmigrants who arrived on airlines for which estimates could be made.

In the case of Mexico, for example, GAO assumed that Mexicans who arrived on Mexicana de Aviacion and Aeromexico had the same low overstay rate as Mexicans who arrived on Air France, Braniff, Air Micronesia, and other relatively more international carriers. The same type of assumption had to be made to estimate some proportion of the overstays from every country. The revised set of procedures that we developed in response to the GAO review generated empirical data that confirmed our initial skepticism about the validity of the assumptions underlying a significant part of the GAO estimates.

Our revised estimation procedure and the new set of estimates were derived over the past few weeks. Because of the unaccountably short timeframe established by GAO for completing this project, we have not had time to document fully our revised methodology. Still, we have summarized the most significant aspects below. As part of our continuing efforts to broaden discussion of these procedures, we plan to release more information after we have reviewed the results, worked out the final details of the new techniques, and conducted sensitivity tests to evaluate various components of the new process.

Our first step was to review the basis for selecting "index" countries to determine whether a statistically defensible method could be developed to expand the list of index countries so that we could estimate system error for a larger number of airlines. When the original list of index countries was selected in the 1980's, the 12 countries were selected because they had extremely low numbers of: apprehensions; applications under IRCA; backlogs for immigrant visas; rates of apparent overstay; and numbers of undocumented immigrants counted in the 1980 census. The original selection process, which was based on determining which countries had the lowest numbers (rather than the lowest rates) of indicated overstays, limited the number of countries selected and produced some index countries that have relatively small numbers of nonimmigrant arrivals. For example, Suriname was selected because it had extremely low numbers on all of the indicators, but Japan was not selected, even though its rates were minuscule, because it did not have extremely low numbers on all five criteria.

In selecting the new list of index countries, we first converted the indicators of overstaying into rates, with the denominator being expected departures of arrivals by air from each country during the October 1990 to March 1991 period. The selection was based on rates computed from two relatively direct indicators of overstaying, apprehensions and applications under IRCA, and one less direct indicator, backlogs for immigrant visas. The 93 countries that had 1,000 or more (air B2) arrivals during the period were eligible for selection. For each country, the three sets of rates by country were weighted equally and summed, the countries were ranked on these totals, and countries with the lowest values were selected to be index countries. The revised list includes: Austria, Belgium, Denmark, Finland, France, Iceland, Luxembourg, Netherlands, Norway, Sweden, Switzerland, Germany, U.K., Japan, Australia, New Zealand, and Cayman Islands.

Ideally, none of the selected index countries would have any overstays. It is likely, however, that a relatively small number of nonimmigrants overstay from each of the countries selected. The estimated number of overstays will be understated to the extent that nonimmigrants from the index countries actually do overstay. Thus our revised figures could be underestimated slightly. Note that it is not possible to overestimate the number of overstays by improperly selecting index countries.

After selecting the new index countries, we followed the GAO recommended procedure, using data on apparent overstays for passengers from index countries to estimate system error for qualifying airlines. Estimates of system error were computed for airlines that flew 1,000 or more

total passengers of index countries; 89 airlines qualified. Passengers from the new set of index countries made up 57 percent of all arrivals, and the 89 qualifying airlines flew 95 percent of all tourists arriving by air during the period.

The estimates, using our revised procedures, are shown in the two attached tables. Our Table 1 replicates your Table 1 and adds new columns with information from our revised estimates. Note that in Table 1 the numbers in the column labeled "Original INS estimates" were used to derive INS' estimates of the resident undocumented population for October 1992. Those estimates are shown in our tables because the numbers that were used to estimate the population are the ones that should be evaluated.

The numbers added to Table 1 show that, for the nine countries selected by GAO, the INS' revised estimates incorporating GAO's recommended approach produced total numbers that are close to the total of the original INS estimates. As would be expected, the probable gain in precision occurs in the distribution by country of origin. The similarities and differences are more evident in the complete list of countries shown in Table 2.

Table 2 shows revised and original numbers and rates of overstays, for every country, for tourists arriving by air during the October 1990 to March 1991 period. Although we have not had time to evaluate fully the two sets of estimates shown in Table 2, we offer the following tentative observations.

1. The "All country" totals are strikingly similar. The totals might have been even closer except that the data base used to make the revised calculations apparently did not include a few hundred thousand arrivals who were included in the data used to compute the original INS estimates.

2. The overall correlation between the two detailed sets of estimates is very high (r=.95). Note, for example, the similarity of the estimates for the African countries. The general pattern of overstay rates will hardly change. Reduced overstay rates for Ireland, Italy, and the Bahamas, and increased rates for China, Hong Kong, and Korea might generate the most interest. The figures for Micronesia and Palau have to be checked.

3. As the result of incorporating the data by airline, as suggested by GAO, there appear to be notable gains in accuracy for individual countries (although such an evaluation has to be somewhat subjective because there are no objective reference points). Except for countries with extremely few arrivals, the revised methodology produced estimated overstays for every country, including many countries estimated by INS to have zero overstays, such as Denmark, Germany, Korea, Singapore, Australia, New Zealand, Palau, and Argentina. A very limited analysis of the distribution, by country, of estimated rates of system error

indicates that the relatively large differences shown in column 4 of table 2 are more likely to represent gains in accuracy than to be spurious results.

To summarize, the work that we have done in response to the GAO review of our methods validates our total estimate of overstays for this time period. It also suggests that GAO's proposed changes will improve our estimates for individual countries. The information generated in computing our revised estimates should be useful for revising the numbers in the draft report (table 1) that have been superseded by the estimates that we derived by incorporating and expanding your proposed changes in the methodology. We will provide you with details of the revised estimation procedures as soon as we have completed our verification and review of the numbers.

Thank you for the opportunity to review the draft report. We will carefully consider any other proposals to improve our estimates of the undocumented immigrant population. If you have questions or need additional information, please contact Robert Warren at (202) 514-5205.

Sincerely,

Doris Meissner
Commissioner

Enclosures

Table 1. GAO Draft Report, Table 1, with Data Added by INS

[----------GAO table 1----------]

Continent and sending country	Expected departures	GAO overstay estimates		INS overstay estimates 1/		Original INS 2/ estimates	Revised INS estimates	
		Percent	Number	Percent	Number		Number	Pct arriving on qualifying airlines
North America								
Bahamas	112,717	0.8	909	6.1	6,889	6,235	4,003	95%
Haiti	20,382	9.3	1,903	15.7	3,208	3,090	3,066	80%
Mexico	399,355	1.8	7,097	4.7	18,697	16,379	17,144	97%
Europe								
France	220,815	1	2,149	1.8	3,921	2,639	1,438	99%
Germany	462,508	0.3	1,542	0	0	0	683	99%
Poland	17,042	22.2	3,787	34.7	5,919	5,820	6,012	32%
Soviet Union	22,220	10.4	2,309	9.8	2,181	2,052	1,776	22%
Asia								
India	21,878	7.7	1,683	9.8	2,103	1976	1,769	97%
Philippines	29,905	11.2	3,336	13.2	3,946	3,772	4,510	68%
Total, 9 countries	1,306,682	1.9	24,715	3.6	46,864	41,963	40,402	N/A
Total, All Countries	5,654,345	N/A	N/A	N/A	N/A	107,423	104,199	95%

1/ As coumputed by GAO.

2/ Used to derive estimates of the total resident undocumented population in October 1992. The updated list of "index" countries includes: Austria, Belgium, Denmark, Finland, France, Iceland, Luxembourg, Netherlands, Norway, Sweden, Switzerland, Germany, U.K., Japan, Australia, New Zealand, and Cayman Islands. These countries were selected because they had the lowest on each of the following: (1) apprehensions, FY 92/ expected departures; (2) visa backlogs, January 1992/expected departures; and (3) IRCA applications/expected departures. In the revised estimates, the 17 index countries had a combined overstay rate of 0.17% (See table 2).

Note: For correction and clarification of INS' footnotes in this table, see our footnote 7 on p. 58.

Table 2. Estimated Differences Between Original INS Estimates of Overstays and Revised Figures that Incorporate GAO's Alternative Estimation Procedure

[Tourists arriving by air, Oct '90 to March '91]

Country of citizenship	Arrivals expected to depart (1)	Estimated overstays: Original INS 1/ (2)	Estimated overstays: Revised per GAO (3)	Changes > 500, rounded to 100s (4)	Overstay rates: Original (5)	Overstay rates: Revised (6)
All countries	5,654,345	107,412	104,199	-3,200	1.9%	1.8%
Europe	2,360,920	25,815	19,155	-6,700	1.1%	0.8%
Albania	76	29	31	-	38.2%	41.4%
Andorra	115	9	9	-	7.8%	7.8%
Austria	38,916	0	52	-	0.0%	0.1%
Belgium	38,289	0	58	-	0.0%	0.2%
Bulgaria	1,965	863	950	-	43.9%	48.4%
Czechoslovakia	4,269	233	185	-	5.5%	4.3%
Denmark	34,527	0	184	-	0.0%	0.5%
Estonia	18	0	1	-	0.0%	5.6%
Finland	49,660	0	39	-	0.0%	0.1%
France	220,815	2,639	1,438	-1,200	1.2%	0.7%
Germany	462,508	0	683	700	0.0%	0.1%
Gibraltar	49	1	4	-	2.0%	8.2%
Greece	17,798	0	121	-	0.0%	0.7%
Hungary	5,666	190	196	-	3.4%	3.5%
Iceland	5,259	0	9	-	0.0%	0.2%
Ireland	38,364	755	320	-	2.0%	0.8%
Italy	144,227	5,155	1,936	-3,200	3.6%	1.3%
Latvia	37	1	4	-	2.7%	10.8%
Liechtenstein	283	0	9	-	0.0%	3.2%
Lithuania	27	5	7	-	18.5%	25.9%
Luxembourg	2,629	0	15	-	0.0%	0.6%
Malta	798	6	20	-	0.8%	2.5%
Monaco	129	0	5	-	0.0%	3.9%
Netherlands	94,510	76	491	-	0.1%	0.5%
Norway	29,496	0	79	-	0.0%	0.3%
Poland	17,042	5,820	6,012	-	34.2%	35.3%
Portugal	12,391	547	296	-	4.4%	2.4%
Romania	4,877	1,172	1,095	-	24.0%	22.4%
San Marino	32	0	1	-	0.0%	3.1%
Soviet Union	22,220	2,052	1,776	-	9.2%	8.0%
Spain	78,105	169	106	-	0.2%	0.1%
Sweden	113,720	0	444	-	0.0%	0.4%
Switzerland	96,176	0	293	-	0.0%	0.3%
United Kingdom	814,019	5,394	1,573	-3,800	0.7%	0.2%
Yugoslavia	11,908	699	710	-	5.9%	6.0%

[more] Footnote at end of table. - Less than 500.

Table 2. Estimated Differences Between Original INS Estimates of Overstays and Revised Figures that Incorporate GAO's Alternative Estimation Procedure [continued]

[Tourists arriving by air, Oct '90 to March '91]

Country of citizenship	Arrivals expected to depart (1)	Estimated overstays: Original INS 1/ (2)	Estimated overstays: Revised per GAO (3)	Changes > 500, rounded to 100s (4)	Overstay rates: Original (5)	Overstay rates: Revised (6)
Asia	1,655,506	17,382	22,910	5,500	1.0%	1.4%
Afghanistan	430	95	100	-	22.1%	23.4%
Bahrain	218	9	11	-	4.1%	4.8%
Bangladesh	1,498	511	530	-	34.1%	35.4%
Brunei	112	0	4	-	0.0%	3.6%
Burma	95	11	26	-	11.6%	27.4%
Cambodia	154	1	8	-	0.6%	5.2%
China, Mainland	32,927	237	1,138	900	0.7%	3.5%
Cyprus	1,034	0	14	-	0.0%	1.4%
Hong Kong	28,651	281	1,052	800	1.0%	3.7%
India	21,878	1,976	1,769	-	9.0%	8.1%
Indonesia	9,221	381	631	-	4.1%	6.8%
Iran	8,032	1,031	1,059	-	12.8%	13.2%
Iraq	220	66	74	-	30.0%	33.4%
Israel	48,254	2,095	1,390	-700	4.3%	2.9%
Japan	1,278,200	0	220	-	0.0%	0.0%
Jordan	4,745	1,232	1,252	-	26.0%	26.4%
Korea	75,904	0	2,101	2,100	0.0%	2.8%
Kuwait	801	58	56	-	7.2%	7.0%
Laos	439	195	205	-	44.4%	46.7%
Lebanon	5,316	980	1,007	-	18.4%	18.9%
Macau	135	1	9	-	0.7%	6.7%
Malaysia	11,731	584	880	-	5.0%	7.5%
Maldives	4	1	1	-	25.0%	25.0%
Nepal	250	29	32	-	11.6%	12.6%
Oman	131	2	4	-	1.5%	2.9%
Pakistan	10,842	2,885	3,032	-	26.6%	28.0%
Philippines	29,905	3,772	4,510	700	12.6%	15.1%
Qatar	138	0	1	-	0.0%	0.7%
Saudi Arabia	4,636	65	106	-	1.4%	2.3%
Singapore	12,412	0	220	-	0.0%	1.8%
Sri Lanka	1,389	248	266	-	17.9%	19.1%
Syria	1,744	283	290	-	16.2%	16.6%
Taiwan	44,353	0	240	-	0.0%	0.5%
Thailand	9,349	10	286	-	0.1%	3.1%
Turkey	8,973	196	213	-	2.2%	2.4%
United Arab Em.	338	6	10	-	1.8%	3.0%
Vietnam	810	66	82	-	8.1%	10.1%
Yemen	237	75	83	-	31.6%	35.2%

[more] Footnote at end of table. - Less than 500.

3.

Table 2. Estimated Differences Between Original INS Estimates of Overstays and Revised Figures that Incorporate GAO's Alternative Estimation Procedure [continued]

[Tourists arriving by air, Oct '90 to March '91]

Country of citizenship	Arrivals expected to depart (1)	Estimated overstays: Original INS 1/ (2)	Estimated overstays: Revised per GAO (3)	Changes > 500, rounded to 100s (4)	Overstay rates: Original (5)	Overstay rates: Revised (6)
Africa	33,149	5,079	5,449	-	15.3%	16.4%
Algeria	681	58	63	-	8.5%	9.2%
Angola	66	6	9	-	9.1%	13.2%
Benin	84	9	11	-	10.7%	13.3%
Botswana	96	1	6	-	1.0%	6.3%
Burundi	17	4	5	-	23.5%	29.4%
Cameroon	313	37	37	-	11.8%	11.9%
Cape Verde	93	15	22	-	16.1%	23.7%
Central Afr Rep	14	0	0	-	0.0%	0.0%
Chad	13	0	1	-	0.0%	7.7%
Comoros	2	2	2	-	100.0%	100.0%
Congo	65	7	7	-	10.8%	11.0%
Djibouti	228	176	191	-	77.2%	84.0%
Egypt	4,448	632	605	-	14.2%	13.6%
Ethiopia	1,762	543	579	-	30.8%	32.8%
Gabon	99	6	9	-	6.1%	9.1%
Gambia	256	91	102	-	35.5%	40.0%
Ghana	1,141	275	300	-	24.1%	26.3%
Guinea	176	71	78	-	40.3%	44.4%
Guinea-Bissau	37	14	17	-	37.8%	45.9%
Ivory Coast	463	43	52	-	9.3%	11.2%
Kenya	900	89	79	-	9.9%	8.8%
Lesotho	27	1	2	-	3.7%	7.4%
Liberia	761	508	552	-	66.8%	72.5%
Libya	88	5	9	-	5.7%	9.8%
Madagascar	49	2	4	-	4.1%	8.2%
Malawi	82	7	11	-	8.5%	13.0%
Mali	130	43	49	-	33.1%	37.5%
Mauritania	9	0	1	-	0.0%	11.1%
Mauritius	201	15	20	-	7.5%	10.1%
Morocco	2,069	97	98	-	4.7%	4.7%
Mozambique	46	2	5	-	4.3%	10.9%
Namibia	43	4	7	-	9.3%	16.3%
Niger	39	7	10	-	17.9%	25.6%
Nigeria	4,138	676	713	-	16.3%	17.2%
Rwanda	22	5	7	-	22.7%	31.8%
Senegal	537	86	95	-	16.0%	17.6%
Seychelles	60	0	5	-	0.0%	8.3%
Sierra Leone	736	457	496	-	62.1%	67.4%

[more] Footnote at end of table. - Less than 500.

Table 2. Estimated Differences Between Original INS Estimates of Overstays and Revised Figures that Incorporate GAO's Alternative Estimation Procedure [continued]

[Tourists arriving by air, Oct '90 to March '91]

Country of citizenship	Arrivals expected to depart (1)	Estimated overstays: Original INS 1/ (2)	Estimated overstays: Revised per GAO (3)	Changes > 500, rounded to 100s (4)	Overstay rates: Original (5)	Overstay rates: Revised (6)
Africa [continued]						
Somalia	700	505	545	-	72.1%	77.9%
South Africa	9,162	249	247	-	2.7%	2.7%
St. Helena	12	0	1	-	0.0%	8.3%
Sudan	394	106	118	-	26.9%	30.0%
Swaziland	32	0	2	-	0.0%	6.3%
Tanzania	301	27	32	-	9.0%	10.7%
Togo	143	23	27	-	16.1%	18.8%
Tunisia	639	18	25	-	2.8%	3.9%
Uganda	210	68	72	-	32.4%	34.4%
Zaire	273	55	58	-	20.1%	21.1%
Zambia	361	0	16	-	0.0%	4.6%
Zimbabwe	931	34	46	-	3.7%	5.0%
Oceania	208,121	574	1,419	800	0.3%	0.7%
American Samoa	7	0	0	-	0.0%	0.0%
Australia	151,021	0	106	-	0.0%	0.1%
Cook Islands	23	0	0	-	0.0%	0.0%
Fiji	1,280	188	214	-	14.7%	16.7%
French Poly.	203	0	4	-	0.0%	2.0%
Kiribati	90	0	3	-	0.0%	3.3%
Marshall Is.	7	0	2	-	0.0%	28.6%
Micronesia, FS	55	0	49	-	0.0%	89.1%
Nauru	269	0	5	-	0.0%	1.9%
New Caledonia	54	0	1	-	0.0%	1.9%
New Zealand	51,259	0	110	-	0.0%	0.2%
Niue	1	0	0	-	0.0%	0.0%
Pacific Is, TT	47	8	10	-	17.0%	21.3%
Palau	1,994	0	464	-	0.0%	23.3%
Papua New Guinea	124	4	8	-	3.2%	6.4%
Solomon Islands	24	0	2	-	0.0%	8.3%
Tonga	910	207	241	-	22.7%	26.5%
Tuvalu	8	0	1	-	0.0%	12.5%
Vanuatu	27	2	4	-	7.4%	14.8%
Wallis & Fut Is	1	0	0	-	0.0%	0.0%
Western Samoa	717	165	195	-	23.0%	27.2%

[more] Footnote at end of table. - Less than 500.

5.

Table 2. Estimated Differences Between Original INS Estimates of Overstays and Revised Figures that Incorporate GAO's Alternative Estimation Procedure [continued]

[Tourists arriving by air, Oct '90 to March '91]

Country of citizenship	Arrivals expected to depart (1)	Estimated overstays: Original INS 1/ (2)	Estimated overstays: Revised per GAO (3)	Changes > 500, rounded to 100s (4)	Overstay rates: Original (5)	Overstay rates: Revised (6)
N. America	874,722	48,740	44,361	-4,400	5.6%	5.1%
Anguilla	583	36	37	-	6.2%	6.3%
Antigua-Barbuda	5,181	55	36	-	1.1%	0.7%
Aruba	2,466	0	21	-	0.0%	0.9%
Bahamas	112,717	6,235	4,003	-2,200	5.5%	3.6%
Barbados	11,716	37	31	-	0.3%	0.3%
Belize	3,127	223	226	-	7.1%	7.2%
Bermuda	623	125	131	-	20.1%	21.1%
British Virg Is	661	27	25	-	4.1%	3.8%
Canada	4,645	1,741	1,834	-	37.5%	39.5%
Cayman Islands	8,139	0	13	-	0.0%	0.2%
Costa Rica	28,813	589	451	-	2.0%	1.6%
Cuba	21,968	3,749	3,050	-700	17.1%	13.9%
Dominica	28,048	1,242	1,162	-	4.4%	4.1%
Dominican Rep	23,647	1,400	1,348	-	5.9%	5.7%
El Salvador	20,328	1,609	1,205	-	7.9%	5.9%
Greenland	7	1	2	-	14.3%	28.6%
Grenada	1,291	140	138	-	10.8%	10.7%
Guadeloupe	208	6	10	-	2.9%	4.6%
Guatemala	38,945	3,189	2,928	-	8.2%	7.5%
Haiti	20,382	3,090	3,066	-	15.2%	15.0%
Honduras	20,543	2,620	2,719	-	12.8%	13.2%
Jamaica	45,826	2,755	2,130	-600	6.0%	4.6%
Martinique	236	0	6	-	0.0%	2.5%
Mexico	399,355	16,379	17,144	800	4.1%	4.3%
Montserrat	300	17	18	-	5.7%	6.0%
Neth. Ant.	6,471	0	23	-	0.0%	0.4%
Nicaragua	15,212	891	729	-	5.9%	4.8%
Panama	20,710	748	598	-	3.6%	2.9%
St. Kitts-Nevis	2,057	129	101	-	6.3%	4.9%
St. Lucia	2,374	231	211	-	9.7%	8.9%
St. Pierre & M.	5	0	0	-	0.0%	0.0%
St. Vincent & Gr	997	163	149	-	16.3%	15.0%
Trinidad & Tob.	24,633	1,273	789	-	5.2%	3.2%
Turks & C Is	2,508	40	26	-	1.6%	1.1%

[more] Footnote at end of table. - Less than 500.

6.

Table 2. Estimated Differences Between Original INS Estimates of Overstays and Revised Figures that Incorporate GAOs Alternative Estimation Procedure [continued]

[Tourists arriving by air, Oct '90 to March '91]

Country of citizenship	Arrivals expected to depart (1)	Estimated overstays: Original INS 1/ (2)	Estimated overstays: Revised per GAO (3)	Changes > 500, rounded to 100s (4)	Overstay rates: Original (5)	Overstay rates: Revised (6)
S. America	489,510	7,601	9,199	1,600	1.6%	1.9%
Argentina	93,007	0	146	-	0.0%	0.2%
Bolivia	5,741	783	863	-	13.6%	15.0%
Brazil	149,824	341	716	-	0.2%	0.5%
Chile	22,275	462	470	-	2.1%	2.1%
Colombia	46,747	2,858	3,001	-	6.1%	6.4%
Ecuador	22,905	1,783	1,918	-	7.8%	8.4%
Falkland Islands	10	0	0	-	0.0%	0.0%
French Guiana	2	0	0	-	0.0%	0.0%
Guyana	2,554	288	279	-	11.3%	10.9%
Paraguay	3,181	79	117	-	2.5%	3.7%
Peru	43,364	820	1,124	-	1.9%	2.6%
Suriname	4,963	0	65	-	0.0%	1.3%
Uruguay	7,596	127	64	-	1.7%	0.8%
Venezuela	87,341	60	436	-	0.1%	0.5%
Stateless	1,691	212	221	-	12.5%	13.1%
Unknown	30,726	2,009	1,486	-500	6.5%	4.8%

- Less than 500.

1/ These estimates were used in computing the INS estimates of the resident undocumented immigrant population as of October 1992. The total number of air B2 arrivals in the original data set was just under 6.1 million, compared to 5.8 million total arrivals in the data set that produced the figures shown in column 3.

The following are GAO's comments on the June 22, 1995, letter from the Immigration and Naturalization Service.

GAO Comments

INS told us that as a result of this report, the agency has changed its procedures for estimating overstays. Specifically, for foreign visitors who arrive by air, INS has discontinued use of the old "global" method and has adopted the airline-by-airline method. INS also told us that the error in its computation formula has been corrected. However, INS' written comments concluded that while the airline-by-airline method may improve estimates for individual countries, it produces a worldwide estimate that is "strikingly similar" to its old worldwide estimate.

We agree that by adopting an airline-by-airline approach, INS has taken an important step forward; however, three major issues remain concerning INS' new methods. These are discussed in the following three sections; a final section (pp. 61-62) responds to INS' concerns about the estimates we presented in table 1.

Does the Airline-By-Airline Method Make a Difference in the Worldwide Estimate?

The first issue concerns INS' conclusion that changing to an airline-by-airline method makes essentially no difference in the worldwide estimate. Alternative INS calculations show that depending on the index countries used, the airline-by-airline method produces a worldwide estimate that is 17- to 25-percent lower than the global-method estimate. As discussed below, (1) the index countries' methodology is potentially sensitive to the inclusion of even one inappropriate index country; and (2) INS' global-method estimate and its airline-by-airline estimate are based on different sets of index countries—and each set may include an inappropriate index country. Specifically, we believe that Saudi Arabia and Japan may be inappropriate index countries.

Table II.1 shows the impact of removing Saudi Arabia and Japan from the index countries INS used. Of the two global-method estimates shown in the first row of the table, we believe that 126,000 (the estimate calculated without Kuwait and Saudi Arabia as index countries) is the more valid figure.[1] The reason is that the index countries' methodology assumes that each index country has virtually no overstays. INS originally selected 12 countries that met this standard, but during the Gulf war, overstays apparently increased for two of INS' original 12 index countries: Kuwait

[1]The global-method estimates totaling 126,000, which we presented in our table 1 (on p. 8), were calculated for us by INS without Kuwait and without Saudi Arabia as index countries.

and Saudi Arabia.[2] Given the time frame involved in the estimates examined here (October 1990 to March 1991), we believe INS would ideally have removed not only Kuwait but also Saudi Arabia from its original set of 12 index countries.[3]

Table II.1: Worldwide Overstay Estimates, October 1990 to March 1991

INS method	Estimates in INS' written comments	Other estimates calculated by INS
Global method	107,000[a]	126,000[b]
Airline-by-airline method	104,000[c]	96,000[d]

[a]Based on 12 original index countries minus Kuwait.
[b]Based on 12 original index countries minus Kuwait and Saudi Arabia.
[c]Based on 17 new index countries.
[d]Based on 17 new index countries minus Japan.

Of the two airline-by-airline estimates shown in the second row of table II.1, we believe that 96,000 (the estimate calculated without Japan as an index country) may be the more valid figure. To calculate its airline-by-airline estimates, INS selected a new set of 17 index countries. Kuwait and Saudi Arabia, among other countries, were dropped, but several large countries were added, including Japan. We realize that INS' goal in selecting these countries was to expand the set of qualifying airlines. We agree with this goal, but we question the inclusion of Japan.

Japanese travelers have lower rates of system error than any other index country on almost every airline for which a comparison can be made.[4] Since index countries are used to estimate system error in nonindex countries, our concern is that the Japan data may not fairly reflect system error for passengers from nonindex countries. (In other words, using an atypical country as an index country can skew results—even if there are

[2]For the time period in question, both Saudi Arabia and Kuwait apparently failed to meet one of INS' stated criteria for selecting index countries for use with its global method: low rates of unrecorded departure forms. (The Saudi rate—9.1 percent—was higher than that of any other index country and higher than the worldwide average.)

[3]To assess whether trend figures would support this view, we examined INS' worldwide overstay estimate for a 6-month period not affected by the Gulf war—October 1986 to March 1987—and compared it to alternative estimates for the time period examined in the report. (We chose the 1986-87 time period because a published INS estimate was available; see Warren, 1990). We found that worldwide, of all tourists who arrived here by air, the percentage INS estimated to be overstays was the same—2.2 percent—in both time periods provided that both Saudi Arabia and Kuwait were dropped as index countries as of the Gulf war. But when INS kept Saudi Arabia as an index country, the estimate decreased to 1.9 percent during the Gulf war time period. That is, consistent results were obtained when, as of the Gulf war time period, both Saudi Arabia and Kuwait were dropped.

[4]For example, on Northwest Orient, the rate of system error is 2.3 percent for Japanese passengers versus 4.9 percent to 13.5 percent for other passengers; and on Air Micronesia, the rate is 3.6 percent for Japanese passengers versus 6.0 to 8.3 percent for other index countries' passengers. (See table I.2 in appendix I.)

no overstays from that country.) INS has said in the past that, for this reason, Japan would not be appropriate for use as an index country.[5] After the agency sent us its written comments, INS recalculated "test estimates" using its airline-by-airline procedures without Japan as an index country.[6]

As summarized in table II.1, INS calculations show that had Saudi Arabia been removed from the original set of index countries, INS' worldwide global-method estimate for the Gulf war period examined here would have been 126,000 rather than 107,000.[7] If Japan had not been included in INS' new, larger set of index countries, INS' worldwide airline-by-airline estimate would have totaled 96,000 rather than 104,000. Substituting one or both of these estimates for those that INS used would yield a very different conclusion: Specifically, the worldwide airline-by-airline estimate would be 17- to 25-percent lower than the worldwide global-method estimate. (That is, 104,000 is 17 percent lower than 126,000, and 96,000 is 25 percent lower than 126,000.)

In light of the foregoing discussion, we believe that INS' conclusion that the airline-by-airline method does not change the worldwide estimate may be an artifact of (1) retaining an index country that is inappropriate for the Gulf war time period and (2) selecting a new index country that may be atypical and, thus, inappropriate. Therefore, we also believe that, before drawing general conclusions about the effect of the airline-by-airline method, other worldwide estimates that are not affected by these possible problems should also be considered.[8]

[5]When introducing an early version of its overstay estimation strategy, INS stated that Japan's rates of system error were "extremely low relative to all other countries of the world" and "clearly do not represent overall system error." (Warren, 1986, p. 653.)

[6]INS describes these new results as preliminary figures.

[7]INS' table 1 (on p. 49) erroneously identifies the column labeled "INS overstay estimates" as computed by GAO. These estimates (which total 126,000 worldwide) were calculated for us by INS—without Kuwait and without Saudi Arabia as index countries. (To further clarify INS' table 1, we note that the next column is labeled "Original INS estimates." These figures, which total 107,000 worldwide, differ from those in the column labeled "INS overstay estimates" only in that Saudi Arabia was included as an index country. Also, to clarify a note in INS' table 1, INS' new set of 17 index countries was used only in calculating its airline-by-airline estimates, which are shown in the column labeled "Revised INS estimates" and total 104,000 worldwide).

[8]Among the comparisons we believe would be relevant are estimates for periods not affected by the Gulf war—for example, 1982 through 1988. Comparisons for these years might indicate a difference of approximately 17 percent, even with no change in index countries. Such comparisons would be particularly relevant because INS' earlier paper on the size of the resident illegal immigrant population (Warren, 1994) covers visitors who arrived from 1982 through 1991. The majority of these years were not affected by the Gulf war.

INS' Airline-By-Airline Procedures Need Further Study

The second issue concerning INS' new estimates arises because certain aspects of INS' airline-by-airline estimation procedures (as we understand them at this time) differ from our own. We have some concerns about these differences, and we believe that further study is needed before INS presents its new overstay estimates as definitive. As just discussed with respect to INS' worldwide estimates, INS' selection of Japan as one of its new index countries may be inappropriate.[9] This points to a need for INS to continue studying its criteria for selecting index countries.[10] As described below, we also have concerns about other aspects of INS' airline-by-airline procedures. These include (1) INS' definition of a qualifying airline and (2) the way that INS derives estimates for those airlines that do not meet that definition.

INS' definition of a qualifying airline is considerably more liberal than the definition we used.[11] Unlike our definition, INS' includes "home-country" airlines that do not fly between the United States and any continent where an index country is located. We agree that expanding the number of qualifying airlines is an appropriate goal. However, stretching the airline-by-airline method (to cover such airlines as A L Argentina, Lan Chile, Avianca, Mexicana de Aviacion, and Aero Mexico) means relying on estimates that are driven by two categories of index countries' travelers that we believe may be atypical.

The first category consists of European and Japanese members of the international business community who reside in Latin America and visit the United States as tourists. Our concern is that this category may consist mainly of experienced international travelers who might be proactive in turning in departure forms. And if they visit the United States frequently (especially, if they reenter this country within 9 months of the expiration of their initial visa), then any uncollected departure forms—which should still be stapled in their passports—would be collected by INS inspectors

[9]While the earlier discussion concerned worldwide estimates, we note here that using Japan as an index country can have a dramatic impact on overstay estimates for certain sending countries. Notably, by removing Japan as an index country, the overstay estimate for Korea decreases from 2,100 to 1,300.

[10]Until INS can adequately study the appropriateness of including Japan as an index country, it would seem more prudent to present a range of airline-by-airline estimates (calculated with and without Japan as an index country) rather than present a single set of estimates.

[11]Whereas 29 airlines met our definition, INS stated that 89 airlines met its definition. We defined a qualifying airline as one that, during the 6-month period we examined, flew at least 1,000 citizens of a single index country to the United States. INS defined a qualifying airline as one that flew 1,000 or more total passengers who are citizens of index countries to the United States. For this reason and also because INS' new set includes 17 index countries (some of which are quite large), INS' definition is clearly more inclusive than the one that we used.

and transferred to the INS database. Relying heavily on this category of index countries' travelers might, therefore, result in an underestimate of system error for the home-country airlines in question and a consequent overestimation of overstays.

The second category consists mainly of European and Japanese tourists on multilegged trips. Our concern is that relying heavily on this category of index countries' passengers might result in estimates of system error for the wrong airlines. The reason is that the airline-by-airline approach necessarily assumes that the airline of arrival is the same as the airline of departure.[12] But, when compared to passengers with a simple round-trip itinerary, "world travelers" on trips that cover three continents (for example, Europe, the United States, and South America) seem more likely to depart the United States on a different airline from the one on which they arrived. In fact, according to one expert, few airlines are licensed for routes both (1) between Europe or Japan and the United States and (2) between Latin American countries and the United States. This raises the possibility that INS' estimates of system error for airlines like Avianca may, in fact, reflect rates for other airlines (Lufthansa, Japan Air Lines, and so forth).

Turning to INS' estimation of overstays among passengers arriving on airlines it classifies as nonqualifying, we are concerned that, whereas we used two alternative assumptions, INS used only one (our assumption 2, as described in table 1 on p. 8). Overall, of course, only 5 percent of tourist air arrivals use airlines that do not qualify under INS' liberal definition.[13] But that 5 percent may include citizens of countries marked by high overstay rates, such as Poland. Use of assumption 1 versus assumption 2 could make a difference of roughly 2,000 overstays for Poland. We believe a range of estimates based on two alternative assumptions might be more meaningful. The need for such a range would be most apparent for sensitivity analyses that omit certain large index countries (such as Japan) or that use a stricter definition of qualifying airlines.

Additional Improvements May Be Possible

The third issue is that some uncertainties are common to our estimates and INS'—and that improvements aimed at reducing these uncertainties may be possible. For example, passengers from seemingly typical index countries may not, in general, represent the true level of system error for

[12]This assumption is required because it is not possible to know on which airline a traveler's departure form was not counted—or on which airline he or she did not depart.

[13]At least this is true when INS includes Japan as an index country.

passengers from nonindex countries—even when traveling on the same airline.[14] Another example is that both we and INS converted negative overstay estimates, which occur for several countries and airlines, to zero overstay estimates. However, because a random component may influence these estimates, it might be preferable to calculate combined-airline and combined-countries estimates without converting the negative numbers to zero. (See p. 41.) The effect of this change would be to reduce the number of estimated overstays.[15]

The special analyses, empirical studies, and more complex models suggested by experts who reviewed this report could shed light on these issues. (See pp. 38-42.) But INS has not yet had an opportunity to consider them.[16]

INS' Concerns About Our Estimates

Our table 1 estimates for nine sending countries include a set of estimates in which assumption 1 was used for passengers on nonqualifying airlines. That is, one set of our estimates is based on the assumption that the same overstay rate characterizes (1) passengers on airlines for which a separate estimate of system error is possible (our qualifying airlines) and (2) passengers from the same country who used other airlines. In its written comments, INS indicated it has empirical evidence that our assumption 1 is incorrect, but did not present that evidence.

We know of no direct way to check the assumption of equal overstay rates for qualifying and nonqualifying airlines.[17] Thus, we presented a second alternative set of estimates for the same countries, based on a very different assumption (assumption 2 in table 1).[18] INS also questioned

[14]The consistently low rate of unrecorded departure forms for Japanese passengers highlights this possibility.

[15]INS' new 104,000 worldwide estimate includes roughly 5,800 overstays from index countries (which supposedly have no overstays). This may be a result of converting negative numbers to zeros rather than including them with positive numbers.

[16]INS produced airline-by-airline estimates during the 30 days provided for comment on a draft of this report, although we had not suggested any time frame for INS to develop new estimation procedures or new estimates.

[17]We did examine three indirect indicators, including factors that might affect a traveler's choice between a qualifying and a nonqualifying airline. Because a difference in fares might be related, we compared current fares for qualifying and nonqualifying airlines flying a number of international routes, and we found no consistent differences. Because some of the major qualifying airlines are based in the United States, we confirmed that these airlines (for example, Delta, United, and American) typically provide bilingual flight attendants. Finally, we found that INS inspectors do not use the airline of arrival in profiles aimed at identifying illegal aliens.

[18]In the draft of this report, the estimates based on assumption 2 were reported in the text rather than in table 1.

assumption 2.[19] However, when assumptions 1 and 2 are used as the basis for alternative estimates, there is a specific logic that we believe justifies the use of these assumptions. (See p. 10.)

[19]Nevertheless, INS did use assumption 2 to derive overstay estimates for the 5 percent of air passengers who arrived on airlines it defined as nonqualifying.

List of Experts

The following experts in the fields of statistics and immigration research reviewed a draft of this report and provided us with their comments.

Frank D. Bean, Ph.D., Population Research Center, University of Texas at Austin

Kenneth H. Hill, Ph.D., Department of Population Dynamics, Johns Hopkins University, Baltimore, Maryland

Demetrios G. Papademetriou, Ph.D., Carnegie Endowment for International Peace, Washington, D.C.

Jeffrey S. Passel, Ph.D., Urban Institute, Washington, D.C.

Donald B. Rubin, Ph.D., Department of Statistics, Harvard University, Cambridge, Massachusetts

Michael S. Teitelbaum, Ph.D., Alfred P. Sloan Foundation, New York, New York

Karen Woodrow-Lafield, Ph.D., Center for Social and Demographic Studies, State University of New York at Albany

Major Contributors to This Report

Program Evaluation and Methodology Division

Robert L. York, Issue Area Director
Eric M. Larson, Project Manager
Donald M. Keller, Adviser
Patrick G. Grasso, Adviser
Harry Conley, Adviser
Venkareddy Chennareddy, Referencer

Bibliography

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland. Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass.: MIT Press, 1975.

Hill, Kenneth. "Discussion." In Bureau of the Census Second Annual Research Conference, March 23-26, 1986: Proceedings, U.S. Bureau of the Census, 1986, pp. 665-70.

Hopper, Kim. "Counting the Homeless: S-Night in New York." Evaluation Review, 16 (1992), pp. 376-88.

Levine, Daniel B., Kenneth Hill, and Robert Warren (eds.), Immigration Statistics: A Story of Neglect. Washington, D.C.: National Academy Press, 1985.

Little, Roderick, and Donald Rubin. Statistical Analysis with Missing Data. New York: Wiley, 1987.

National Research Council (U.S.) Panel on Census Requirements in the Year 2000 and Beyond. Modernizing the U.S. Census. Washington, D.C.: National Academy Press, 1995.

Norwood, Janet. Organizing to Count: Change in the Federal Statistical System. Washington, D.C.: Urban Institute, 1995.

Norwood, Janet. "The Consumer Price Index, the Deficit, and Politics." Policy Bites, No. 22. Washington, D.C.: Urban Institute, Mar. 1995.

Papademetriou, Demetrios G. "Undocumented Immigration in the United States: Some Thoughts About Research Challenges, Impacts, and Recent Policy Initiatives." Studi Emigrazione/Etudes Migrations, Vol. XXV, No. 91-92 (Sept.-Dec. 1988), pp. 591-617.

Passel, Jeffrey S. "Illegal Immigration: How Big A Problem?" Paper presented at the "Conference on Latin American Migration: The Foreign Policy Dimension" sponsored by the U.S. Department of State. Washington, D.C.: Mar. 17, 1995.

Rubin, Donald B. Multiple Imputation for Nonresponse in Surveys. New York: Wiley, 1987.

Warren, Robert. "Estimates of the Undocumented Immigrant Population Residing in the United States, by Country of Origin and State of Residence: October 1992." Revised version of a paper presented at "California Immigration 1994," a seminar sponsored by the California Research Bureau, 1994.

Warren, Robert. "Annual Estimates of Nonimmigrant Overstays in the United States: 1985 to 1988." In Frank D. Bean, Barry Edmonston, and Jeffrey S. Passel, Undocumented Migration to the United States: IRCA and the Experience of the 1980s. Washington, D.C.: Urban Institute Press, 1990, pp. 77-110.

Warren, Robert. "The INS Nonimmigrant Information System: Assessing the Statistics on Nondeparture." In Bureau of the Census Second Annual Research Conference, March 23-26, 1986: Proceedings. U.S. Bureau of the Census, 1986, pp. 650-64.

U.S. Immigration and Naturalization Service, Statistical Yearbook of the Immigration and Naturalization Service, 1990. Washington, D.C.: U.S. Government Printing Office, 1991.

Related GAO Products

Illegal Aliens: National Net Cost Estimates Vary Widely (GAO/HEHS-95-133, July 25, 1995).

Illegal Aliens: Assessing Estimates of Financial Burden on California (GAO/HEHS-95-22, Nov. 28, 1994).

Tax Administration: Earned Income Credit—Data on Noncompliance and Illegal Alien Recipients (GAO/GGD-95-27, Oct. 25, 1994).

Benefits for Illegal Aliens: Some Program Costs Increasing, But Total Costs Unknown (GAO/T-HRD-93-33, Sept. 29, 1993).

Illegal Aliens: Despite Data Limitations, Current Methods Provide Better Population Estimates (GAO/PEMD-93-25, Aug. 5, 1993).

Immigration Enforcement: Problems in Controlling the Flow of Illegal Aliens (GAO/T-GGD-93-18, Mar. 30, 1993).

Immigration and the Labor Market: Nonimmigrant Alien Workers in the United States (GAO/PEMD-92-17, Apr. 28, 1992).

Immigration: An Issue Analysis of an Emerging Problem (GAO, Sept. 1985).

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000 or by using fax number (301) 258-4066, or TDD (301) 413-0006.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (301) 258-4097 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

PRINTED ON RECYCLED PAPER

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested