

Working Paper No. 3
27 March 2003

ENGLISH ONLY

**UN STATISTICAL COMMISSION and
UN ECONOMIC COMMISSION FOR EUROPE**

**STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES (EUROSTAT)**

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE-Eurostat Work Session on Migration Statistics
organised in cooperation with the UN Statistics Division
(Geneva, 28-30 April 2003)

Session III – Invited paper

**ESTIMATING ILLEGAL IMMIGRANTS IN A CENSUS OR SURVEY:
THE THREE-CARD METHOD – AN INNOVATIVE TECHNIQUE**

Submitted by United States¹

¹ Paper prepared by Nancy Kingsbury, Judith Droitcour and Eric Larson, United States General Accounting Office. This paper (submitted March 2003) reflects the views of the authors; it does not necessarily reflect the views of the U.S. General Accounting Office. The authors may be contacted via e-mail at: kingsburyn@gao.gov, droitcourj@gao.gov, larsone@gao.gov

I. INTRODUCTION

1. In the United States, the size of the resident illegal population is debated (with current estimates ranging upwards from about 8 million). Little reliable information exists concerning the characteristics and behaviors of this increasingly large group, and as a result, the impacts of illegal immigration are difficult to document. Thus, policymakers facing decisions about immigration may not be adequately informed (U.S. Commission on Immigration Reform, 1994). Moreover, without reasonably reliable and valid estimates of the illegal immigrant population, it is difficult to accurately assess the coverage of our national decennial census; for example, if census figures indicate unexpected population increases, can this be traced to unexpectedly high levels of illegal immigration—or might this have resulted from shifts in coverage levels? (See Deardorff and Blumerman, 2001.)

2. The “three-card method” (GAO, 1999, 1998; Droitcour and Larson, 2002; Droitcour et al., 2001) is a relatively new survey approach designed to estimate both the size and characteristics of the resident illegal population. This approach builds on previous anonymous questioning techniques—such as randomized response (Greenberg, et al., 1969; Warner, 1965), aggregated response (Boruch and Cecil, 1979; Warner, 1971), and the item-count technique (Droitcour et al., 1991)—as well as demographic methods of residual estimation (Warren and Passel, 1987).

3. The present paper (1) explains the basic logic of the three-card method, (2) summarizes special features and results of preliminary testing, (3) describes how the “visa overstay” portion of the illegal population can be estimated, (4) outlines alternatives for fielding a three-card survey, (5) discusses variance costs, and (6) considers design-based strategies for reducing those costs.

II. BASIC LOGIC OF THE THREE-CARD METHOD

4. The three-card method is a useful alternative to threatening questions that ask the respondent to choose one of several mutually exclusive answer categories (e.g., immigration statuses)—only one of which is highly sensitive (the illegal or undocumented category). Using the three-card method to estimate illegal immigration involves:

- Selecting three random samples, each of which (1) represents the resident foreign-born population, (2) is composed of different respondents, (3) is asked the same question about immigration status, and (4) is shown an answer card with immigration status categories;
- Using three slightly different answer cards, one for each sample; and
- Arranging the (mutually exclusive) answer categories in a three-box answer format—with a slightly different arrangement on each answer card.

5. The three slightly different answer cards are designed so that each sample provides a different piece of less sensitive information—a different “piece of the puzzle.” The outlines of the missing sensitive piece are apparent when all other pieces are in place. No respondent is ever asked directly whether he or she is in the sensitive answer category. Information on the sensitive answer category is estimated indirectly for the entire population of interest, or for major subgroups. The three-box format is illustrated in Legal Status Card 1 on the following page (shown in a reduced size; actual cards are 8-1/2” by 11” or approximately 210mm by 197mm). Because the

cards were developed in Spanish, we provide an English translation of Legal Status Card 1, as follows:

- Box A of card 1: Legal permanent resident with a valid and official green card² issued to me by the U.S. government.
- Box B of card 1: U.S. citizen; student, work or tourist visa; undocumented (do not have my own valid official green card); and refugee or asylee³ (without a green card).
- Box C of card 1: Some other category not in Box A or Box B (specify).

Legal Status Card 1

A **Extranjero residente legal**
con una tarjeta válida y oficial
(también le llaman "green
card") expedida para su
persona por el gobierno de
los Estados Unidos

B **Ciudadano de
los Estados Unidos**

**Visa de estudiante,
de trabajo, o de turista**

Indocumentado
La persona que no tiene una
tarjeta válida y oficial (también
le llaman "green card")
expedida para su persona por
el gobierno de los Estados Unidos

Refugiado o asilado
Sin "green card"

C **Otra categoría**
Que no se encuentra en A o B
(especifique)

6. Legal Status Card 1 is used with sample 1. Respondents are asked to report only which box applies to them and are told that if it is Box B, we do not want to know which specific category applies to them.

7. The main purpose of interviews with sample 1 is to obtain a valid estimate of the percentage of foreign-born persons who have officially obtained green cards (the first "piece of the puzzle"). Sample 1 data also provide information on the characteristics of the "green card" population, based on other items in the questionnaire.


² In the United States, the term "green card" is used informally to refer to legal permanent resident status.





³ An asylee is a person who has been granted asylum.


8. Legal Status Card 2 rearranges the answer categories. Here, Box A contains the status of a naturalized U.S. citizen, whereas Box B now includes the legal permanent resident or green card status, along with the sensitive illegal category and other immigration statuses. This answer card is shown to respondents in sample 2 (completely different individuals than in sample 1).

9. Again, the respondent is asked to answer by picking one of the boxes. Sample 2 provides a direct estimate of the percentage of foreign-born persons who are naturalized citizens. Sample 2 also provides information on the characteristics of naturalized citizens.

Legal Status Card 2

A Ciudadano de los Estados Unidos 

B Extranjero residente legal 
 Visa de estudiante, de trabajo, o de turista 
 Indocumentado 
 Refugiado o asilado Sin "green card" 

C Otra categoría Que no se encuentra en A o B (especifique) 

10. Legal Status Card 3 (on the following page) again rearranges the answer categories. This time, Box A features refugees and persons granted asylum as well as those here legally with temporary visas. This card is shown to respondents in sample 3 (completely different individuals than those in samples 1 and 2). Respondents choosing Box A, can be asked which specific status applies to them. But as always, respondents are told that if they are in Box B, we do not want to know which specific status applies to them. The main purpose of sample 3 data is to provide a valid estimate of the percentages of foreign-born in Box A statuses. Again, Sample 3 also provides correlates and other information about persons in Box A statuses.

Legal Status Card 3

A

Visa de estudiante, de trabajo, o de turista
Válida actualmente, no expirado

Refugiado o asilado
Con aprobado oficial vivir en los Estados Unidos
Sin "green card"

B

Ciudadano de los Estados Unidos

Extranjero residente legal

Visa expirado o indocumentado

C

Otra categoría
Que no se encuentra en A o B (especifique)

11. In summary, each of the legal status categories—except for the sensitive category—alternatively appears in Box A. Assuming that the legal status categories are mutually exclusive and that these categories, together with Box C (“some other category...”), represent an exhaustive set of possible legal statuses, it is possible to obtain an indirect estimate of illegal immigrants via a subtraction procedure (which may be carried out separately for various major subgroups, thus allowing major characteristics and behavior groups to be estimated).⁴

12. To illustrate, we use the following hypothetical example for U.S.-resident, Mexican-born respondents.⁵ Suppose 39% of sample 1 Mexican-born respondents chose Box A of Card 1 (green card status); 20% of sample 2 Mexican-born respondents chose Box A of Card 2 (naturalized U.S. citizen); and 4% of sample 3 Mexican-born respondents chose Box A of Card 3 (refugees, asylees, and persons here legally with temporary visas). This would indicate that 63% of Mexican-born respondents are here legally in these major statuses ($39\% + 20\% + 4\% = 63\%$). Suppose that 1% also picked Box C (some other category). Subtracting these hypothetical estimates from 100 yields 36% ($100\% - 63\% - 1\% = 36\%$). Thus, our indirect estimate would be that 36% of Mexican-born residents are here illegally.⁶

⁴ The statistical expression of this indirect estimator has been reported elsewhere (GAO, 1999; 1998).

⁵ While these are hypothetical data, we based these figures on existing data, estimates, or projections to the extent possible. Because of possible inaccurate self-reports in current surveys, there is some uncertainty regarding the percentage of Mexican residents of the U.S. who are naturalized U.S. citizens (see Passel and Clark, 1998).

⁶ We note that, in general, this approach to indirect estimation would not be appropriate if the sensitive category is extremely rare.

III. SPECIAL FEATURES AND PRELIMINARY TESTING

13. “Trainer cards” are a special feature of the method. Before showing a respondent the legal status card assigned to his or her sample, the interviewer asks questions on less sensitive topics using answer cards with three boxes. The purpose is to train respondents in the use of the three-box format—and to familiarize them with the notion that “if you’re in Box B, we don’t want to know which specific category applies to you.” Thus, a respondent presented with a legal status card will be pre-set to understand that the interviewer is not going to “zero in” on the sensitive category in Box B. Two three-box trainer cards (not shown here) are illustrated in GAO, 1999.

14. Another special feature is the ability to ask follow-up questions of respondents who choose Box A. (This is because, presumably, those with illegal status have chosen Box B.) The follow-up questions can help validate the respondent’s choice of Box A.

15. Preliminary testing of the Legal Status Card 1 was conducted with foreign-born Hispanics residing in the United States; these initial interviewees were selected to over-represent illegal immigrants. Three phases of testing included: 27 initial pretests conducted by bilingual GAO staff; 81 interviews with farm workers conducted by a contractor; and 8 cognitive interviews⁷ conducted by a pretest interviewer, after training by a cognitive expert. Overall, most respondents appeared to comprehend the three-box format for answering and to accept it at face value. Of the 116 persons interviewed using Legal Status Card 1:

- None refused to answer the legal status question (i.e., everyone “picked a box”).
- About 70% (81 out of 116) picked Box B, which contained the sensitive category.⁸

IV. ESTIMATING “VISA OVERSTAYS”

16. By asking an additional direct question, the three-card method can be used to estimate “visa overstays” and their characteristics (no additional cards are needed). The additional direct question would be asked a bit differently for Samples 1 and 2 than for Sample 3. For respondents in Samples 1 and 2, the additional direct question would simply be whether they entered the U.S. on a temporary visa when they first came here to stay. For respondents in Sample 3, the additional direct question could be asked using Legal Status Card 3 (because on that card, the temporary-visa status is in Box A and, once Box A has been chosen, the respondent can be asked for the specific category). It is thus possible to identify in each of the three samples, those persons who first came on a temporary visa.

17. Then—using data on this subgroup only—it is possible to use the basic 3-card data and the subtraction technique illustrated earlier to estimate illegals in this subgroup. Logically, such persons would be visa overstays.

V. ALTERNATIVES FOR FIELDING A THREE-CARD SURVEY

18. At present, administration of three-card questions is limited to in-person interviews (because answer cards must be shown to respondents⁹). The main options are:

- i) To add (or “piggy-back”) the required questions onto a census or a very large existing survey, e.g., one that includes tens of thousands of foreign-born respondents.

⁷ See Tourangeau (1984).

⁸ For more detailed information on the testing and results, see GAO (1999).

⁹ Potentially, a web-survey or a mail-survey format could be devised.

ii) To field a follow-on three-card survey of foreign-born respondents who participated in an initial survey or census.¹⁰

iii) To conduct a special survey or add the three-card questions to a smaller survey (e.g., 12,000 foreign born respondents).

19. Unless the three-card questions can be administered to a very large number of respondents, strategies are needed to reduce variance costs—particularly, the variance costs associated with indirect estimation of the sensitive category.

VI. VARIANCE COSTS

20. Anonymous estimation typically carries variance costs. For the three-card method, variance costs derive, first, from the need to use three separate samples. The estimate of the less sensitive category featured in Box A of card 1 (legal permanent residents) is based on Sample 1 respondents only. Obviously, the number of Sample 1 respondents is lower than the total number of respondents in all three samples. Thus, the variance of that estimate is higher than it would have been if respondents in all three samples had contributed to that estimate. The same is true of each less sensitive category estimated with data from Sample 2 only or from Sample 3 only. The variance is further increased for the indirect or anonymous estimate of the sensitive category (which never appears in Box A), because the indirect estimate is obtained by a linear combination of the three direct estimates. Assuming that the three samples are independent, the variance of the indirect estimate is the sum of the variances of the direct estimates.

21. Immigration status estimates for subgroups (e.g., percent of foreign-born males who are here on temporary visas) have a higher variance than corresponding estimates for the total population. (Fewer respondents are in the subgroup than in the total population.)¹¹

VII. VARIANCE REDUCTION STRATEGIES

22. Design strategies for efficient estimation of the sensitive category include (1) modifying procedures for selecting the three samples or (2) developing an alternative conceptualization of the cards.

23. Procedures for selecting the three samples can be modified by:

- Using a principle of “optimal allocation”¹² to choose relative sizes for samples 1, 2, and 3, based on the relative sizes of the variances of the expected Box A parameters for cards 1, 2, and 3.
- Using information about the respondents in the initial sample (i.e., information they reported in the initial survey) to stratify when selecting samples 1, 2, and 3.¹³

24. We are currently exploring the impact of these approaches on variance reduction, through a set of simulations. Stratification seems a promising approach because existing information

¹⁰ Depending on the size of the initial survey and the costs of the follow-back, all foreign-born respondents or a subset of them might be selected for the three-card samples.

¹¹ Still higher variances would obtain for an estimate of the *percentage of those in the sensitive category who fall into a particular subgroup*. Such an estimate would consist of the ratio of two indirect estimates: The numerator would be the estimate of illegals falling in the designated category (e.g., Mexican-born persons illegally residing in the United States) and the denominator would be the estimate of total foreign-born residing in the United States illegally.

¹² See Cochran (1977, pp. 96ff).

¹³ This is obvious for a follow-on survey but might also be done for add-on questions if laptop computers are used to conduct the interview and to select the three samples.

suggests that immigration status is correlated with demographic information typically collected in a census or survey. (We refer here to data on the characteristics of foreign-born U.S. citizens from the U.S. Current Population Survey (CPS), March 2000 supplement, and to information on the characteristics of other statuses based on three Census Bureau publications.¹⁴)

25. The strategy involving an alternative conceptualization of the cards is simply to use two cards with two samples (instead of three cards with three samples). Consider the following example:

- On revised Legal Status Card 1, Box A would include two categories—green-card holders and persons here with temporary visas; Box B would include U.S. citizen, illegals, and refugees/asylees.¹⁵
- On revised Legal Status Card 2, Box A would include two categories—U.S. citizens and refugees/asylees; Box B would include green-card holders, illegals, and persons here with temporary visas.

26. The two-card approach might not be advised for some groups of foreign-born respondents, because it could be too threatening for those who are here illegally. (With only two cards, there are fewer “less sensitive” categories in Box B; thus two cards provide less “cover” for respondents in the illegal category.) However, the two-card approach might be used effectively with low-threat groups, such as foreign-born persons married to a U.S.-born person.¹⁶ High-threat versus low-threat groups could be determined as part of the survey, if a computer-assisted approach to interviewing is used. Alternatively, this targeted approach could be used in a follow-on survey, as discussed in an earlier section of this paper.

REFERENCES

Boruch, Robert F., and Joe S. Cecil. Assuring the Confidentiality of Social Research Data. Philadelphia: University of Pennsylvania Press, 1979.

Cassidy, Rachel, and Lucinda Pearson. Evaluating Components of International Migration: Legal Temporary Migrants. U.S. Census Bureau Population Division. Washington, D.C.: U.S. Census Bureau, 2001.

Cochran, William G. Sampling Techniques. 3rd ed. New York: John Wiley & Sons, 1977.

Deardorff, Kevin E., and Lisa M. Blumerman. Evaluating Components of International Migration: Estimates of the Foreign-Born Population by Migrant Status in 2000. U.S. Census Bureau Population Division, Working Paper #58. Washington, D.C.: U.S. Census Bureau, 2001.

Droitcour, Judith et al., “The Item Count Technique as a Method of Indirect Questioning: A Review of Its Development and A Case Study Application” in Paul Biemer et al., Measurement Error in Surveys. Wiley, 1991.

Droitcour, Judith A., and Eric M. Larson, “An Innovative Technique for Asking Sensitive Questions: The Three Card Method,” Bulletin de Methodologie Sociologique (Paris), No. 75:5-23, July, 2002.

Droitcour, Judith A., Eric M. Larson and Fritz J. Scheuren, “The Three Card Method: Estimating Sensitive Survey Items--With Permanent Anonymity of Response,” 2001 Proceedings of the American Statistical Association, Social Statistics Section [CD-ROM], Alexandria, VA: American Statistical Association, 2001.

¹⁴ Cassidy and Pearson, (2001); Deardorff and Blumerman (2001); and Perry et al. (2001).

¹⁵ As previously noted, respondents choosing Box A can be asked which specific status applies to them. But as always, respondents are told that if they are in Box B, we do not want to know which specific status applies to them.

¹⁶ There would be very few illegal and deportable persons in this group.

GAO. See U.S. General Accounting Office, below.

Greenberg, Bernard, et al., "The Unrelated Questions Randomized Response Model Theoretical Framework," Journal of the American Statistical Association, 64:520-539, 1969.

Passel, Jeffrey S., and Rebecca L. Clark, "Immigrants in New York: Their Legal Status, Incomes, and Taxes." Washington, D.C.: The Urban Institute, 1998.
(www.urban.org/immig/immny.html)

Perry, Marc, Barbara Vandervate, Lea Auman, and Kathy Morris. Evaluating Components of International Migration: Legal Migrants. U.S. Census Bureau Population Division Working Paper #59. Washington, D.C.: U.S. Census Bureau, 2001.

Tourangeau, Roger, "Cognitive Sciences and Survey Methods," in Thomas Jabine et al. (eds.). Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines, Washington, D.C.: National Academy of Sciences, 1984.

U.S. Commission on Immigration Reform, U.S. Immigration Policy: Restoring Credibility, Washington, D.C.: U.S. Government Printing Office, 1994.

U.S. General Accounting Office (GAO), Immigration Statistics: Information Gaps, Quality Issues Limit Utility of Federal Data to Policymakers. GAO/GGD-98-164. Washington, D.C.: U.S. General Accounting Office, 1998.

U.S. General Accounting Office (GAO), Survey Methodology: An Innovative Technique for Estimating Sensitive Survey Items. GAO/GGD-00-30. Washington, D.C.: U.S. General Accounting Office, 1999.

Warner, Stanley, "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias," Journal of the American Statistical Association, 66:63-69, 1965.

Warner, Stanley, "The Linear Randomized Response Model," Journal of the American Statistical Association, 66:884-888, 1971.

Warren, Robert, and Jeffrey S. Passel, "A Count of the Uncountable: Estimates of Undocumented Aliens Counted in the 1980 Census," Demography, 24: 375-93, 1987.