

Evaluation Report Covering Place of Birth, U.S. Citizenship Status, and Year of Arrival

FINAL REPORT

Philip Harris,
Renuka Bhaskar, and
Claire Shook-Finucane
Population Division
Leah Ericson
Decennial Statistical
Studies Division

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

Intentionally Blank

TABLE OF CONTENTS

| | |
|---|-----|
| EXECUTIVE SUMMARY | ii |
| 1. BACKGROUND | 1 |
| 1.1 Motivation for the 2006 ACS Content Test..... | 1 |
| 1.2 Motivation for Proposed International Migration Question Changes..... | 1 |
| 1.3 Previous Testing or Analysis | 4 |
| Place of Birth | 4 |
| U.S. Citizenship Status | 5 |
| Year of Arrival..... | 5 |
| 2. RESEARCH QUESTIONS AND SELECTION CRITERIA..... | 6 |
| 2.1 Place of Birth | 6 |
| 2.2 U.S. Citizenship Status | 7 |
| 2.3 Year of Arrival..... | 7 |
| 3. METHODOLOGY | 9 |
| 3.1 Data Collection Methods | 9 |
| 3.2 Sample Design | 10 |
| 3.3 Methodology Specific to the Research Questions | 11 |
| 4. LIMITATIONS..... | 13 |
| 4.1 General Content Test and Content Follow-Up Limitations..... | 13 |
| 4.2 Limitations Specific to Place of Birth, U.S. Citizenship Status and Year of Arrival | 14 |
| 5. RESULTS | 15 |
| 5.1 Response to the Content Test and Content Follow-Up..... | 15 |
| 5.2 Responses to the Place of Birth Question..... | 15 |
| 5.3 Responses to the U.S. Citizenship Status Question | 18 |
| 5.4 Responses to the Year of Arrival Question..... | 19 |
| 6. SUMMARY OF EMPIRICAL RESULTS..... | 25 |
| References..... | 27 |
| Appendix A: Information Page..... | A-1 |
| Appendix B: Westat Cognitive Testing Report | B-1 |

EXECUTIVE SUMMARY

Objective

In January through March of 2006, the American Community Survey (ACS) conducted the first test of new and modified content since the ACS reached full implementation levels of data collection. The results of the testing provided below will determine the content for the 2008 ACS. The proposed changes to the place of birth, U.S. citizenship status, and “year of arrival” questions on the ACS are intended to correct inaccuracies with the current questionnaire items and are necessary to meet congressionally funded initiatives.¹

Collecting information on detailed place of birth will improve Congressionally-funded activities such as the intercensal population estimates program.² It will also improve the understanding of regional variations of source countries of the foreign born and aid in editing country of birth by providing more detail to use for editing. Collecting year of naturalization will 1) aid in verifying the accuracy of the reported rates of naturalization, shown by some research to be overstated with the current U.S. citizenship status question (Van Hook et al., 2004) and 2) serve as the first official benchmark to compare with Office of Immigration Statistics (OIS) administrative records on naturalization. It also will assist in editing U.S. citizenship status by providing more detail. Collecting more accurate information on year of arrival is imperative to enumerating the foreign born (such as international migration estimates) and checking the consistency of responses for year of naturalization. The current year of entry question does not measure duration of residence in the United States as well as possible because it does not ask respondents if they have entered the United States to live, work, or study more than once. Therefore, the year reported might not be the first time that they entered the United States.

Methodology

The 2006 ACS Content Test section related to the foreign born compared two versions each of the place of birth, U.S. citizenship status, and year of entry questions.³ The control version replicated the current ACS question. The test version modified the place of birth question by including a write-in line for respondents born outside the United States to provide their city, town, or village of birth. The test version also modified the U.S. citizenship status question by including a write-in line for naturalized foreign-born respondents to provide their year of naturalization. Finally, the test version modified the year of arrival question by asking

¹ The term year of arrival is used to indicate the proposed approach to the question. In the past, the term year of entry has been preferred and here the attempt is to introduce the term year of arrival.

² Detailed place of birth refers to the “city/town/village” of birth questionnaire item.

³ The place of birth question is also administered to the native population and the U.S. citizenship status and year of entry questions are also administered to persons born in Puerto Rico and the U.S. Island Areas.

respondents to this question when they *first* came to live in the United States, if they had come to live more than once and, if so, when was their most recent year of arrival.⁴

The test version of the U.S. citizenship status question also included the response option “Yes, born abroad of U.S. citizen parent or parents” in place of “Yes, born abroad of American parent or parents.” The test version of the year of arrival question also included revisions to the question language to clarify movement to the United States (adding the phrase “*not including holidays, vacations, or other brief visits*”).

The Content Follow-Up (CFU) reinterview portion of the testing was designed to measure consistency of response. For year of arrival, the CFU specifically examines whether respondents reported the same number of arrivals and the same year or years of arrival.

Research Questions and Results

Below are key research results for each question concentrating on response rates and overall effect on key statistics.

Place of Birth Summary

1. Does asking for a lower level of geography (city/town/village) for place of birth impact nonresponse rates for this item?

- No. Asking for a lower level of geography than state or country did not have a statistically significant impact on the response rate for state or country of birth.

2. Does asking for a lower level of geography (city/town/village) for place of birth impact the distribution of country of birth?

- No. Asking for a lower level of geography than country did not have a statistically significant impact on the distribution of country of birth responses for the targeted countries of birth of the foreign born.⁵

3. What proportion of the foreign-born responses can be uniquely coded to a region within the foreign country of birth?

- Overall, approximately 72 percent to 83 percent of foreign-born responses could be coded to a unique region within the targeted countries of birth.

4. How much additional time is required to code cases at a regional level rather than a country level only? (Note that this research question is for informational purposes only)

⁴ The control version only asked respondents to indicate when they came to live, work, or study in the United States.

⁵ The targeted countries were the four largest source countries for U.S. foreign born: Mexico, the Philippines, China, and India.

- An additional two to three weeks were required to manually code the detailed place of birth responses in the 2006 ACS Content Test. However, this coding system is not a production-ready system at this time.

U.S. Citizenship Status Summary

1. Do the following changes to the U.S. citizenship status question reduce the estimate of naturalized citizens?

- a) **Adding a write-in field for year of naturalization.**
- b) **Changing the born abroad category to specifically reference U.S. citizen parents.**
- Changing the U.S. citizenship status question to include year of naturalization and referencing U.S. citizen parents did not statistically reduce the proportion of respondents reporting that they were naturalized.

2. Do the changes to the U.S. citizenship status question impact item nonresponse rates for the question?

- No. Changing the U.S. citizenship status question to include year of naturalization and referencing U.S. citizen parents did not increase the nonresponse rate to the U.S. citizenship status item.

Year of Arrival Summary

1. Which of the two versions has less systematic response error associated with the response of entering the United States once or more-than-once?

- There was no difference in the systematic response error between the two versions. However, respondents in the test version were very inconsistent when reporting whether they arrived once or more than once.

2. Which of the two versions has the lowest item nonresponse rates?

- The control and test version had roughly equivalent item nonresponse rates, although nonresponse was high for both.

3. Do the year of entry values provided in the control version reflect a first year of arrival, most recent year of arrival, or something else? (Note that this research question is for informational purposes only)

- The year provided in the control version more often matched the first year of arrival given in the CFU than the last year of arrival. However, it often represented neither first nor last year of arrival.

Summary of the Test Results for the International Migration Questions⁶

Asking for a city, town, or village of birth did not significantly impact the response rate to the place of birth question, nor did it impact the distribution of country of birth responses for targeted countries of birth for the foreign born. Furthermore, the detailed place of birth question adequately allowed for coding to a higher level of geography. While the manual coding of detailed place of birth took approximately two to three weeks, the system used for the Content Test is currently not feasible for full production.

Changing the U.S. citizenship status question to include year of naturalization and referencing U.S. citizen parents neither reduced nor negatively impacted the proportion of respondents reporting that they were naturalized. Additionally, the changes did not impact the response rate for U. S. citizenship status and the vast majority of naturalized respondents were able to provide a year of naturalization.

Year of arrival reporting exhibited mixed results. Overall, the two versions did not differ significantly in terms of reliability. However, the accuracy of number of arrivals in the test version was poor, with a large increase in the proportion of respondents indicating single versus multiple arrivals when reinterviewed. For both the test and control, the reporting of the year arrived was marginally consistent when examining whether the respondent reported the exact same year upon reinterview. However, the reporting consistency increased when the range of acceptability increased to a two or four year window. Response to Year of Most Recent Arrival was low amongst respondents with more than one arrival. Additionally, for those control respondents in the CFU with more than one arrival, the year that they originally provided more often reflected their first arrival than their most recent arrival, although it quite often represented neither.

⁶ The international migration questions referred to in this report are the place of birth, U.S. citizenship status, and year of arrival/year of entry questions.

1. BACKGROUND

1.1 Motivation for the 2006 ACS Content Test

In January through March of 2006, the American Community Survey (ACS) conducted the first test of new and modified content since the ACS reached full implementation levels of data collection. The results of that testing will determine the content for the 2008 ACS. The year 2008 marks the first year of a three year aggregated data product that includes data from the same year as the 2010 decennial census (2008 - 2010). Similarly, 2008 is the midpoint year for the first five-year data product that includes data from 2010 (2006-2010). Given the significance of the year 2008, the ACS committed to a research program during 2006 that will result in final content determination in time for the 2008 ACS. This research is the 2006 ACS Content Test.

Through the Office of Management and Budget (OMB) Interagency Committee on the ACS, the Census Bureau included subject matter experts and key data users from other federal agencies in identifying questions for inclusion in the Content Test. In general the Content Test evaluated alternatives for questions which showed some indication of a problem, for example, high missing data rates, estimates which differed systematically from other sources of the same information, or high simple response variance as measured in the Census 2000 Content Reinterview survey (CRS). In addition, the Content Test also included testing of three new topics proposed by other federal agencies for inclusion in the ACS.

To meet the primary objective of the 2006 ACS Content Test, analysts evaluated changes to question wording, response categories, instructions, or examples relative to the current version of the questions. Additionally, the Content Test design reflected two secondary objectives. One of the secondary objectives addressed form design alternatives for the basic demographic section of the form. The second addressed the content of the questionnaire mailing package. Results indicated no interaction between either of the two secondary objectives and the first objective addressing changes made to questions. Thus, this report will only address testing specific to the first objective - testing of alternative questions, response categories, etc. Specifically, this report discusses the place of birth, U.S. citizenship status, and year of arrival questions.

1.2 Motivation for Proposed International Migration Question Changes

Overview

In the last few years, the U.S. Government Accountability Office (GAO), the National Academy of Sciences, and other experts have clearly stated the need for improvements to current immigration items and for the introduction of new immigration items on the ACS, among other national surveys (Edmonston, 1996; GAO 1998a and 1998b; Redstone and Massey, 2003). The multiyear appropriations received by the Census Bureau beginning in FY2003 cited the testing and implementation of improved survey items on international migration as key performance measures.

The data on the foreign born serve a critical operational need to the Census Bureau's annual population estimates program as they are the basis for the net international migration component,

which drives much of the annual population change.⁷ The proposed changes will benefit the ACS by improving the data for the current population estimates used as controls through the increased accuracy of place of birth, U.S. citizenship status, and year of arrival. The proposed changes will also benefit the ACS by aiding the Census Bureau's survey edit specifications regarding data on place of birth, U.S. citizenship status, and year of arrival because the changes should both lower the amount of edits done to the data and provide additional data on which to base the edits.

U.S. Citizenship Status

An addition of a question on date of naturalization (in addition to a question on duration of permanent resident) has been identified as a key needed improvement by the National Academies (Edmonston, 1996), a panel of social scientists who gathered in early 2003 at the Migration Policy Institute (Grieco, 2003), and researchers at the Urban Institute (Fix et. al., 2003). Fix, et. al. provide support for the need to have data on year of naturalization by indicating that the characteristics of naturalized citizens (such as age and country of birth) vary substantially depending on time of naturalization.

Year of Arrival

Census Bureau analysts and outside experts have identified two key problems with the current "year of entry" item: 1) the current item allows immigrants to report only one entry into United States, even if they have entered on more than one occasion; and 2) the current question leaves open for broad interpretation the notion of "coming to live" in the United States. Federal administrative data and other research have clearly demonstrated that immigration patterns into the United States have changed over the last few decades, with temporary and circular migration becoming more prevalent. This changing reality is not reflected in the current year of arrival question.

In its 1998 evaluation of the progress made by the Census Bureau (and the former Immigration and Naturalization Service, INS) in implementing the 1996 National Academies recommendations on international migration statistics, the GAO concluded that current immigration statistics "do not accurately reflect the number and characteristics of immigrant populations, or how immigrants fare after entry into the United States (GAO, 1998a)."

Several experts in the field of migration have identified problems with the Census Bureau's year of arrival data. For example, Redstone and Massey (2003) note:

We find that the usual census procedure for estimating immigrant experience systematically misses the mark in 41% of cases [in the decennial census], that [an] underestimate is much more likely than [an] overestimate, and that the size

⁷ According to "Annual Estimates of the Components of Population Change by Race and Hispanic or Latino Origin for the United States: July 1, 2003 to July 1, 2004" produced by the U.S. Census Bureau 43 percent of the increase in population between July 1, 2003 and July 1, 2004 was due to the increase in net international migration.

of the error is just under four years, on average. Both the likelihood and size of the error are determined by when the first trip was taken and the number of trips that have occurred between then and the survey date. The erroneous estimation of U.S. experience by the census question on year of arrival . . . (renders) virtually any comparative analysis of immigrant assimilation based on U.S. census data impossible to interpret. (Redstone and Massey, 2003)⁸

Although Myers (2004) generally concludes that the current decennial census question on year of entry is “more reliable and less error prone than other data commonly relied upon,” he does find inconsistencies, especially for recent immigrants (i.e., those who indicated a year of entry within the last five years). For example, Myers—comparing the Census question on year of entry with the Census question on residence five years ago—states that according to his analysis of Census 2000 PUMS data, nearly one-third (31.6 percent) of recent immigrants who reported a year of arrival in the United States later than 1995 indicated that they resided in the United States on April 1, 1995—this rises to 36.5 percent for Hispanic immigrants. Myers also highlights the importance of having reliable and valid year-of-entry data as they are at the center of the political debate on immigrant assimilation, especially the assimilation of recent immigrants—the group for which, he states, the year of entry question appears to yield the greatest inconsistencies in reporting.

Numerous other experts across disciplines have continued to directly or indirectly highlight data gaps and data quality issues with international migration data (National Research Council, 2004; Massey and Capoferro, 2004; National Research Council, 2001; National Research Council, 1996).

Other researchers external to the Census Bureau, such as Camarota (2003), have found issues with the current year of entry item in the ACS. During his evaluation of the year of entry data and his interviews with ACS field representatives (FRs), he found that both respondents and FRs were confused by what information the year of entry question was asking for and how to report multiple arrivals to the United States.

The Census Bureau’s analysis of the Census 2000 CRS data shows that the year of entry item has been moderately inconsistent for the two most recent censuses (U.S. Census Bureau 2004). Additionally, Schmidley and Robinson’s (1998 and 2003) analyses of CPS data found that some respondents claimed to be both naturalized U.S. citizens and recent immigrants (less than 5 years since entry); they suggested that multiple entries into the United States might have caused data inconsistencies as respondents may have provided their most recent entry year but actually had a first entry some time ago.

Finally, Wellens (1993) conducted a cognitive evaluation of CPS “nativity” (i.e., place of birth and U.S. citizenship status) data and found that respondents often had difficulty interpreting the year of entry item.

⁸ The authors compared census data with New Immigrant Survey Pilot [NISIP] data.

Thus, in summary, there are numerous studies that have questioned the validity of the U.S. Census Bureau's existing year of entry question. It obviously does not account for multiple entries; it has exhibited moderately high inconsistency when analyzed via reinterview; and respondents have consistently reported difficulties interpreting the question in cognitive studies.

1.3 Previous Testing or Analysis

Cognitive interviews are integral to the process of changing survey questions. This type of research is conducted to verify that the potential questions will be readily understood by the public in order to reduce response error. Two rounds of cognitive testing were conducted on the proposed place of birth, U.S. citizenship status, and year of arrival questions. Westat directed the first set of cognitive interviews and Saber Systems conducted the second set.⁹ A "think aloud" approach was used and respondents were asked to verbalize their thoughts on the meaning of the questions and their responses to them. After the respondents concluded the international migration questions, the interviewer asked additional probing questions regarding the items. A summary of the results and recommendations from the cognitive interviews follows.

Place of Birth

Originally, the detailed place of birth questionnaire item asked for the "state/province/region" of the respondent's country of birth. However, Westat found that respondents often answered the "state/province/region" questionnaire item incorrectly and would instead respond with the city, town, or village of birth. Westat concluded that respondents often did not have knowledge of the administrative subdivision of their country of birth. "Administrative subdivisions below the country level and above the city level do not seem to be as salient for many people from other countries, as compared to the United States (Kerwin, 2005)." Therefore, Westat recommended that the "Census Bureau consider asking for the city or town in which foreign-born individuals were born, rather than the state, province, or region (Kerwin, 2005)." See Appendix B for the complete Westat cognitive testing report.

Sabre Systems tested "city/town/village" of birth and found that the "respondents did not have difficulty understanding and providing valid detailed place of birth data (Poros and Orum, 2005)."

⁹ The results of the first round of cognitive interviews, conducted by Westat, resulted in the modification of the proposed question items for the ACS. These revisions were approved on March 29, 2005. However, the structure of the year of arrival (formerly year of entry) question would not have allowed for the calculation of nonresponse rates for the year of most recent arrival questionnaire item.

The Immigration Statistics Staff, the Decennial Statistical Studies Division (DSSD), and Westat collaborated on creating two alternate versions of a revised year of arrival question. Sabre Systems cognitively tested these two versions of the year of arrival question, along with minor edits to the place of birth and citizenship questions. The findings from the second round of cognitive testing conclusively found that one version of the proposed year of arrival question was easier for the respondents to answer than the other.

While the question itself asks for city, town, or village, the goal is to code data to state, province, or regional levels.

U.S. Citizenship Status

The proposed U.S. citizenship status question included the addition of the year of naturalization questionnaire item and the changing of the response option “Yes, born abroad of American parent or parents” to “Yes, born abroad of U.S. citizen parent or parents.” Westat found that respondents “were able to provide the year of naturalization for themselves and family members (Kerwin, 2005).” See Appendix B for the complete Westat cognitive testing report.

Sabre Systems agreed with Westat’s findings that naturalization is a salient event that carries little recall response error. Sabre found that “respondents understood and gave valid answers to the question on U.S. citizenship status, including the year of naturalization, both for themselves and others in their households (Poros and Orum, 2005).” They also recommended including the response option “No, not a U.S. citizen” which was needed for parallel response structure.

Year of Arrival

The proposed year of arrival question included questionnaire items asking for year of first arrival and year of most recent arrival. It also included revisions to the question language to clarify movement to the United States. Westat found that there was very little problem with respondent recall of the information; however, issues existed in respondents’ ability to provide the information the question was asking for. Westat studied the use of the phrase “*not including holidays, vacations, or other brief visits*” and found that it successfully clarified the question.

Sabre Systems tested two versions of a revised year of arrival question. While the respondents understood both versions, the version that was selected as the ACS year of arrival question for the content test had a less complex response structure and therefore was found to be conclusively easier for respondents to answer.

2. RESEARCH QUESTIONS AND SELECTION CRITERIA

The following research questions and hypotheses are being used to evaluate the proposed changes to the place of birth, U.S. citizenship status, and year of arrival questions on the ACS. In addition, the question language and a summary of the research questions, evaluation measures, and selection criteria can be found in Appendix A.

2.1 Place of Birth: Research Questions and Selection Criteria

2.1.1 Does asking for a lower level of geography (city/town/village) for place of birth impact nonresponse rates for this item?

Selection Criterion: The item nonresponse rates for the test version will be equal to or less than the rates for the control version.

2.1.2 Does asking for a lower level of geography (city/town/village) for place of birth impact the distribution of country of birth?

Selection Criterion: The distribution of place of birth at the country level (the level of geography currently collected) will be roughly equivalent between the test and control versions.

2.1.3 What proportion of the foreign-born responses can be uniquely coded to a region within the foreign country of birth?

Selection Criterion: At least 70 percent of the responses from respondents reporting birth outside of the United States will code to a distinct region for each of the targeted countries for analysis.

2.1.4 How much additional time is required to code cases at a regional level rather than a country level only? (Note that this research question is for informational purposes only)

Selection Criterion: Time and staffing requirements will not be excessively more burdensome for the test question than the control question.

2.2 U.S. Citizenship Status: Research Questions and Selection Criteria

2.2.1 Do the following changes to the U.S. citizenship status question reduce the estimate of naturalized citizens?

- *adding a write-in field for year of naturalization*
- *changing the born abroad category to specifically reference U.S. citizen parents*

Selection Criterion: Estimates of naturalized U.S. citizens from the test version will be equal to or less than the estimates from the control version.

2.2.2 Do the changes to the U.S. citizenship status question impact item nonresponse rates for the question?

Selection Criterion: The item nonresponse rate for the test version will be equal to or less than the item nonresponse rate for the control version.

2.3 Year of Arrival: Research Questions and Selection Criteria

The Content Follow-up Survey (CFU) was used to provide data to measure the consistency of responses for the year of arrival question. The Index of Inconsistency will measure consistency between the original survey and the follow up.

2.3.1 Which of the two versions has less systematic response error associated with the response of entering the United States once or more-than-once?

Selection Criterion: The net difference rate will be lower in the test version than the control version (at a decade level).

2.3.2 Which of the two versions has the lowest item nonresponse rates?

Selection Criterion: The item nonresponse rate for the test version will be equal to or less than the item nonresponse rate in the control version.

2.3.3 Do the year of entry values provided in the control version reflect a first year of arrival, most recent year of arrival, or something else? (Note that this research question is for informational purposes only)

Selection Criterion: If the test version is accepted, this will help us provide information to users regarding how to interpret the data from previous years relative to how the data are collected in the test version.

3. METHODOLOGY

3.1 Data Collection Methods

3.1.1 The 2006 ACS Content Test data collection

The 2006 ACS Content Test consisted of a national sample of approximately 62,900 residential addresses in the contiguous United States. (The sample universe did not include Puerto Rico, Alaska and Hawaii). To meet the primary test objective of evaluating question wording changes, approximately half of the sample addresses were assigned to a test group (31,450) and the other half to a control group (31,450). For the topics already covered in the ACS, the test group included the proposed alternative versions of the questions, and the control group included the current version of the questions as asked on the ACS. Both the test and control questionnaires included three new topics not currently on the ACS. Both test and control included the three new topics to keep context and questionnaire length consistent between the two versions.

The ACS Content Test used a similar data collection methodology as the current ACS, though cost and time constraints resulted in some deviations. Initially, the ACS collects data by mail from sampled households, following a mailing strategy geared at maximizing mail response (i.e., a pre-notice letter, an initial questionnaire packet, a reminder postcard, and a replacement questionnaire packet). The Content Test implemented the same methodology, mailing each piece on the same dates as the corresponding panel in the ACS. However, the Content Test did not provide a toll-free number on the printed questionnaires for respondents to call if they had questions, as the ACS does. The decision to exclude this service in the Content Test primarily reflects resource issues in developing the materials needed to train and implement the operation for a one-time test. However, excluding this telephone assistance allows us to collect data that reflects the respondent's interpretation and response without the aid of trained Census Bureau interviewer.

The ACS follows-up with mail nonrespondents first by Computer Assisted Telephone Interviewing (CATI) if a phone number is available, or by Computer Assisted Personal-visit Interviewing (CAPI) if the unit cannot be reached by mail or phone. For cost purposes, the ACS subsamples the mail and telephone nonrespondents for CAPI interviewing. In comparison, the Content Test went directly to CAPI data collection for mail nonrespondents, dropping the CATI data collection phase in an effort to address competing time and resource constraints for the field data collection staff. While skipping the CATI phase changes the data collection methods as compared to the ACS, eliminating CATI allowed us to meet the field data collection constraints while also maintaining the entire mail nonrespondent universe for possible CAPI follow-up. Using CATI alone for follow-up would have excluded households for whom we do not have a phone number.

The ACS also implements an edit procedure on returned mail questionnaires, identifying units for follow-up who provided incomplete information on the form, or who reported more than five people living at the address. (The ACS questionnaire only has space to collect data for five people.) This is called the Failed Edit Follow Up operation (FEFU). The ACS calls all households identified as part of the FEFU edit to collect the remaining information via a CATI

operation. The Content Test excluded this follow-up operation in favor of a content reinterview, called the Content Follow-Up (CFU). The CFU also contacts households via CATI but the CFU serves as a method to measure response error, providing critical evaluative information. The CFU operation included all households who responded by mail or CAPI and for whom we had a phone number. More information about the CFU operation follows below.

The Content Test mailed questionnaires to sampled households around December 28, 2005, coinciding with the mailing for the ACS January 2006 panel. The Content Test used an English-only mail form but the automated instruments (both CAPI and CFU) included both English and Spanish translations. Beginning February 2006, a sample of households that did not respond by mail was visited by Census Bureau field representatives in attempt to collect the data. The CAPI operations ended March 2, 2006.

3.1.2 Content Follow-Up data collection

The CFU reinterview, conducted by the Census Bureau's three telephone centers, provided a method for measuring response error. About 2 weeks after receiving the returned questionnaire or completed CAPI interview, the responding unit entered the CFU operation. Telephone staff completed the CFU interviews between January 17 and March 17, 2006. At the first contact with a household, interviewers asked to speak with the original respondent. If that person was not available, interviewers scheduled a callback at a time when the household member was expected to be home. If at the second contact we could not reach the original respondent, interviewers completed the interview with another adult household member.

The CFU reinterview did not replicate the full ACS interview. Rather, the CFU used the roster and basic demographic information from the original interview and only asked questions specific to the analytical needs of the Content Test. Reinterview questions were of two general formats: the same question as asked in the original interview (in some cases, modified slightly for a CATI interview), or a different set of questions providing more detail than the question(s) asked in the original interview for the same topic. For topics in which the CFU asked the same question as the original interview, the CFU asked the test or control version of the question based on the original treatment. For these cases, the goal was to measure the reliability of the answers – how often we obtained the same answer in the CFU as we did in the original mail or CAPI data collection. For topics using a different question or set of questions than the original interview, we asked the same detailed series of questions regardless of the original treatment condition. Generally, these questions were more numerous than what we could ask in the ACS. In some cases the questions came from another existing survey, for example, for labor force, we asked the labor force questions from the Current Population Survey questions. In other cases the CFU asked additional probing questions based on prior testing results, such as for health insurance. For these topics, the goal was to measure how close the original answers were to the more detailed CFU answers.

3.2 Sample Design

The sample design for the ACS Content Test consisted of a multi-stage design, with the first stage following the Census 2000 Supplementary Survey (C2SS) design for the selection of

Primary Selection Units (PSUs) defined as counties or groups of counties. The first stage selection of PSUs resulted in 413 PSUs or approximately 900 counties being selected.

Within sampled PSUs, households were stratified into high and low response strata based on tract level mail response rates to the Census 2000 long form and a stratified systematic sample of households was selected. The strata were defined such that the high response stratum contained 75 percent of the housing units that reside in tracts with the highest mail response rate. The balance of the tracts was assigned to the low response stratum. To achieve similar expected number of mail returns for the high and low response strata, 55 percent of the sample was allocated to the low response strata and 45 percent to the high response strata.

A two-stage sampling technique was used to help contain field costs for CAPI data collection. The initial sample of PSUs was sorted by percentage of foreign-born population since the majority of that target population responds via CAPI. At least one item undergoing testing in the content test required an adequate sample of this population. The 20 PSUs with the highest percentage of foreign-born population were included with certainty and the remaining PSUs were sampled at a rate of 1 in 3. For the second stage, mail nonresponding households were sampled at a rate of 1 in 2 within the top 20 PSUs and at a sampling rate of 2 in 3 within the remaining PSUs. The final design designated 151 PSUs be included in the CAPI workload.

In the majority of PSUs, we assigned cases to both the control and test groups. To maintain field data collection costs and efficiencies, PSUs with an expected CAPI workload of less than 10 sampled addresses had all of their work assigned to only one treatment (either control or test). The PSUs were allocated to the two groups such that the aggregated PSU characteristics between the two groups are similar for employment, foreign born, high school graduates, disabled, poverty status, tenure, and Hispanic origin.

There was no sampling for CFU. A CFU interview was attempted for all responding households to the Content Test for which we had a phone number.

3.3 Methodology Specific to the Research Questions

An automated coding process was used for the place of birth item. Codes were assigned to U.S. state and foreign country responses.¹⁰ The items were assigned high, medium, and low confidence, with high and medium confidence items being autocoded at the U.S. Census Bureau headquarters. Expert coders at the National Processing Center (NPC) coded items with low confidence. Expert coders at headquarters manually coded the detailed place of birth section of the item. A code was assigned to the city, town, or village responses that were related to the foreign country of birth. Coded information from the NPC file was used in conjunction with write-in responses to assign the appropriate code to the city, town, or village responses. The completion of the coding took approximately two to three weeks.

¹⁰ This included coding for place of birth responses of Puerto Rico, the U.S. Island Areas, and the District of Columbia.

The CFU for the control panel question on year of arrival was designed to be similar to the test panel question. In the original treatment, the control respondents were only given the option to report one entry. One of the goals of the CFU for the year of arrival item was to understand how control respondents in the original treatment would have responded had they been given the option to report more than one arrival. Therefore, the interviewers in the CFU presented respondents with a similar year of arrival question that was used for the original test panel. This method of reinterview allowed for an assessment of the year most respondents provided (first, most recent, or neither) when asked what year they came to live in the United States.

4. LIMITATIONS

4.1 General Content Test and Content Follow-Up Limitations

As noted in section 3.1, Data Collection Methods, the Content Test maintained the same general mail data collection methodology as the ACS, but differed in the mail nonresponse follow-up operations. In general, the deviations do not impact the validity of the results, and in many cases increased the effectiveness of the testing. However, some aspects of the Content Test implementation should be considered in evaluating the data.

- As noted, the Content Test did not include CATI data collection in order to meet field data collection constraints. While the design of the Content Test allowed all sampled housing units an opportunity to participate even without CATI, questions administered differently over the phone did not get the benefit of a full CATI operation (though some of the CAPI interviews actually do occur by phone). However, since only ten percent of ACS data is collected by CATI and CATI interviewers are trained to help respondents understand question intent and response categories, overall ACS data quality should not suffer when questions are implemented using CATI.
- Though the test design required that field interviewers work only control or only test cases, interviewers in both conditions worked regular ACS production interviews at the same time they completed the Content Test cases. By design the control instrument very closely replicated the ACS production instrument, only differing in the addition of the three newly proposed topics. As a result, interviewers in the test condition had to learn and use two very different instruments, while control interviewers used basically the same instrument between their Content Test cases and ACS production. Thus, test interviewers experienced more challenges in completing their overall caseload. Interviewer debriefing suggested that test interviewers had some difficulty dealing with the two very different instruments simultaneously which may have some impact on the administration of the test version.
- On the first day of CFU interviewing, we discovered a usability problem with the CFU instrument. Left unaddressed, the usability problem could have potentially impacted comparisons between the Content Test and CFU responses when looking specifically at gross difference rate or simple response variance calculations. However, we immediately implemented two steps to mitigate any data problems -- a special instruction sheet to remind interviewers about how to avoid the potential problem and a procedure to report any problems to headquarters for repair. Interviewers followed the instructions and reported 90 cases to us. Post-collection processing corrected all reported errors, though it is possible that some cases went unreported.

4.2 Limitations Specific to Place of Birth, U. S. Citizenship Status, and Year of Arrival

- The Content Test did not provide Telephone Questionnaire Assistance (TQA) or Telephone Edit Follow-Up (TEFU) and it did not use CATI. As the foreign-born population may not have been as acquainted with the ACS questions or language, it likely affected the responses to the Content Test. The TEFU is especially useful for questionnaires with evidence of potential confusion, or that have extensive missing data. Additionally, if respondents only filled out the short form items there would be no way to follow-up on missing data.
- The ACS Content Test sample design was stratified by population from the Census 2000 Supplementary Sample (C2SS), with a base size of 250,000 or higher. This design may not be able to correctly identify non-traditional settlement areas, or areas that have quickly increased in population since C2SS which may lead to an underreporting by the foreign-born population.
- Research has shown that language barriers “represent a significant impediment to collecting high-quality data on the foreign born (Camarota, 2003).” Therefore, having the Content Test mail form only available in English and the CATI/CAPI only available in English and Spanish may have affected the data collected on the foreign-born population as a whole.
- The foreign-born population is typically undercovered, especially certain sub-groups of the population. There are many reasons for this undercoverage, including living arrangements that are difficult to enumerate and certain sub-groups of the foreign-born population that avoid enumeration (Camarota, 2003). Research has shown that ACS data on “seasonal” populations may be more difficult to utilize, therefore it may be harder to collect data on the foreign-born population who participate in seasonal migration patterns (Gage, nd).

5. RESULTS

5.1 Response to the Content Test and Content Follow-Up

Control and test treatment groups obtained equivalent response rates overall, and for each mode of collection. Similarly, response to the Content Test is comparable to response for the production ACS.

The table below gives the weighted response rates for each data collection operation and a test of differences between the control and test groups. The overall response rate reflects the final response to the initial data collection (mail and CAPI only). There are no significant differences between response rates for the control and test groups. Note that the denominator for each calculation includes only eligible cases for each mode.

Table 1. Content Test Response Rates, Control vs. Test

| Response Rate | Total (%) | Control (%) | Test (%) | Difference (%) | Margin of Error (%) | Significant |
|-----------------------|-----------|-------------|----------|----------------|---------------------|-------------|
| Overall response rate | 95.7 | 95.8 | 95.5 | -0.3 | ± 0.9 | No |
| Mail response rate | 51.3 | 51.5 | 51.2 | -0.3 | ± 2.2 | No |
| CAPI response rate | 92.4 | 92.6 | 92.1 | -0.4 | ± 1.7 | No |
| CFU response rate | 76.2 | 75.9 | 76.4 | 0.5 | ± 1.6 | No |

5.2 Responses to the Place of Birth Question

5.2.1 Does asking for a lower level of geography (city/town/village) for place of birth impact item nonresponse rates for this item?

There was no statistically significant difference between the control and test panels in terms of nonresponse to the state or country of birth item (7.5 percent and 6.9 percent, respectively (Table 2)).¹¹ Additionally, only 16 percent of respondents who provided a foreign country of birth did not provide a detailed place of birth in the test version.

¹¹ All of the population universes for tables 2 through 11 are restricted to those records in which a data defined person could be established. The population universe in Table 2 for the state or country of birth item is all records. The population universe in Table 2 for the detailed place of birth item is all test records indicating a place of birth outside of the United States.

Although a statistically higher proportion of test cases were completely missing a state or country (6.4 percent to 5.2 percent), many of these cases contained country of birth information in the detailed place of birth field.

Table 2. Place of Birth Nonresponse Rates, Control vs. Test

| | State or country of birth | | | | | Detailed place of birth | |
|--|---------------------------|----------|----------------|---------------------|-------------|-------------------------|---------------------|
| | Control (%) | Test (%) | Difference (%) | Margin of Error (%) | Significant | Test (%) | Margin of Error (%) |
| Completely Missing Write-In Value | 5.2 | 6.4 | 1.2 | ±0.6 | Yes | 10.2 | ±2.1 |
| Completely Missing or Un-codeable Write-In Value | 7.5 | 6.9 | -0.6 | ±0.6 | No | 16.0 | ±2.4 |

5.2.2 Does asking for a lower level of geography (city/town/village) for place of birth impact the distribution of country of birth?

According to the chi-square statistic, there was no statistically significant difference between control and test panels for the targeted countries of birth in the distribution of codeable responses to the country of birth item (Table 3).¹² The chi-square statistic measures the difference in the control and test distributions for a given question. If the statistic is significantly large, the distributions are not the same. Significantly large here would be a p-value of .at least .01. In this case the statistic is not significantly large. Therefore, we can conclude that asking for a lower level of geography does not impact the distribution of country of birth.

Table 3. Place of Birth: Distribution of Codeable Responses

| | Country of birth | | | | | Detailed place of birth |
|----------------------|------------------|----------|----------------|---------------------|-------------|-------------------------|
| | Control (%) | Test (%) | Difference (%) | Margin of Error (%) | Significant | Test (%) |
| Mexico | 29.7 | 26.9 | -2.8 | ±4.3 | No | 27.5 |
| China | 3.8 | 3.0 | -0.8 | ±1.0 | No | 3.0 |
| Philippines | 3.2 | 3.2 | -0.0 | ±0.9 | No | 2.9 |
| India | 3.1 | 2.7 | -0.4 | ±1.1 | No | 2.5 |
| All other countries* | 60.1 | 64.2 | 4.1 | ±4.4 | No | 64.1 |
| Total: | 100.0 | 100.0 | | | | 100.0 |

Chi-Squared Value: 4.54 p = 0.34

*Includes Puerto Rico and the U.S. Island Areas.

Table 3 also contains the high level distribution of codeable responses to the detailed place of birth question. This distribution represents city, town, or village responses, grouped by country.

¹² The population universe in Table 3 for the country of birth item is all records indicating a place of birth outside of the United States. The population universe in Table 3 for the detailed place of birth item is all test records indicating a place of birth outside of the United States.

Here it is important to note that the distribution of codeable responses to detailed place of birth in the test panel are similar to the distribution of codeable responses to country of birth (the level of geography that the ACS currently collects the data) for both the test and control panels.

5.2.3 What proportion of the foreign-born responses can be uniquely coded to a region within the foreign country of birth?

Overall, 81 percent of foreign-born detailed place of birth responses were codeable to a unique region (Table 4).¹³ For the targeted foreign-born countries of birth, the proportion ranged from 72 percent to 83 percent. For the Philippines and India, the proportion of detailed place of birth responses codeable to a unique region was not statistically different than 70 percent. However, for these countries an additional 6 to 9 percent of cases could have been coded, but were not coded to a unique area. In a production environment with increased resources and staff specialization it may have been possible to code these cases to a more specific region. Therefore the selection criterion that more than 70 percent of the detailed place of birth responses should be codeable to a unique region for each targeted country of birth was met.

Table 4. Place of Birth: Proportion of Codeable Detailed Place of Birth Responses (Test Panel Only)

| | Mexico | China | Philippines | India | All Other Countries* | Total Foreign | No Country |
|------------------------------------|--------|-------|-------------|-------|----------------------|---------------|------------|
| 1 Codeable Region | 82.5 | 80.3 | 72.2 | 75.1 | 81.3 | 81.1 | 0.0 |
| 2 or More Codeable Regions | 2.3 | 5.6 | 9.2 | 5.7 | 3.2 | 3.3 | 12.3 |
| Region Write-In Missing or Unclear | 15.2 | 14.1 | 18.6 | 19.2 | 15.6 | 15.6 | 87.7 |
| Total | 100% | 100% | 100% | 100% | 100% | 100% | 100% |

*Includes Puerto Rico and the U.S. Island Areas.

5.2.4 How much additional time is required to code cases at a regional level rather than a country level only? (Note that this research question is for informational purposes only)

The completion of the coding for the detailed place of birth write-in field took approximately two to three weeks. The detailed place of birth responses were manually coded using four coders. This type of coding was appropriate and feasible given the sample size of the content test. However, in full production, an automated coding system would need to be implemented.

Coding of detailed places would require additional resources including a system of automated coding, training for coders, development of manuals, and an additional step of coding for places that were initially deemed uncodeable. A system similar to one used to code the ancestry item would be optimal but is not currently feasible.

¹³ The population universe in Table 4 is all test records for which a place of birth outside of the United States was coded. The column “no country” represents records in which a detailed place of birth was reported but no country was reported. This is included for informational purposes only.

Additional challenges for implementation of a coding system are 1) the volume of detailed places requiring an assigned code 2) the similarity or replication of names for many places and 3) the lack of a unified geographic structure across source countries and the United States.

5.3 Responses to the U.S. Citizenship Status question

5.3.1 *Do the following changes to the U.S. citizenship status question reduce the estimate of naturalized citizens?*

- *adding a write-in field for year of naturalization*
- *changing the born abroad category to specifically reference U.S. citizen parents*

In the control panel, the proportion of respondents to the U.S. citizenship status question who were naturalized citizens was 5.2 percent, compared with 5.0 percent of responses in the test version. This difference, however, was not statistically significant (Table 5).¹⁴ While the introduction of the year of naturalization write-in field did not significantly reduce the estimate of naturalized citizens, it did not negatively impact the U.S. citizenship status question and can provide operational advantages. Specifically, one advantage is that having a year of naturalization allows for improved editing of U.S. citizenship status. For example, in family households in which a parent has a reported year of naturalization and the U.S. citizenship status of his or her child born outside the United States is unknown, the year of naturalization of the parent (in combination with place of birth of the child) can help determine whether to assign the child a naturalized status (by derivation from the parent) or a status of born abroad to a U.S. citizen parent. If the year of naturalization were not known, that child would most likely be assigned a naturalized status because it is unclear whether the child was born before or after the parent naturalized.

Table 5. U.S. Citizenship Status: Distribution of Responses

| | Control (%) | Test (%) | Difference (%) | Margin of Error (%) | Significant |
|--|-------------|----------|----------------|---------------------|-------------|
| Born in the United States | 86.6 | 87.5 | 0.9 | ±1.3 | No |
| Born in Puerto Rico, Guam, U.S. Virgin Islands, or Northern Marianas | 0.6 | 0.5 | -0.1 | ±0.2 | No |
| Born abroad of U.S. citizen parent (s) | 0.6 | 0.8 | 0.1 | ±0.2 | No |
| Citizen by naturalization | 5.2 | 5.1 | -0.2 | ±0.6 | No |
| Not a U.S. citizen | 6.9 | 6.2 | -0.7 | ±0.9 | No |
| Total | 100.0 | 100.0 | | | |

Chi-Squared Value: 4.45 p = 0.35

¹⁴ The population universe in Table 5 is all records.

5.3.2 Do the changes to the citizenship question impact item nonresponse rates for the question?

There were no statistically significant differences in the item nonresponse rates. About 3 percent of responses to the U.S. citizenship status question were completely missing for both control and test panels (Table 6).¹⁵ In terms of year of naturalization, only 10 percent of respondents who indicated being a naturalized U.S. citizen did not report a year of naturalization. Analysis of consistency measures for year of naturalization also showed good reporting as about 88 percent of respondents reported a year of naturalization in the CFU that was within 3 years of the original year given in the Content Test.

Table 6. U.S. Citizenship Status: Nonresponse rates

| | Control (%) | Test (%) | Difference (%) | Margin of Error (%) | Significant |
|-------------------------|-------------|----------|----------------|---------------------|-------------|
| U.S. Citizenship Status | 3.0 | 2.7 | -0.3 | ±4.0 | No |
| Year of Naturalization | N/A | 9.8 | N/A | ±2.0 | N/A |

5.4 Responses to the Year of Arrival Question

5.4.1 Which of the two versions has less systematic response error associated with the response of entering the U.S. once or more-than-once?

Overall, the indicators of consistency for the year of arrival item showed mixed results. The selection criterion for this research question had to be adapted because it did not provide enough insight into how well the test and control items worked. Conceptually, the year of arrival evaluation measures were designed to determine 1) if, and how accurately, people reported single or multiple arrivals (test version only) and 2) if, and how accurately, people reported their year(s) of arrival. The original selection criterion associated with this question only addressed the second of these two concerns. Therefore, the following section discusses the results for both of these concerns.

We present information on the number of arrivals first (Table 7).¹⁶ When examining the section of the test question in which respondents were to indicate whether they had entered once or more than once (“Have you come to live. . .MORE THAN ONCE?”), the index of inconsistency was quite high. The index of inconsistency (IOI) is the percentage of the variance that is due to

¹⁵ The population universe in Table 6 for the U.S. citizenship status item is all records. For the year of naturalization item the universe is all test records indicating a naturalized U.S. citizenship status.

¹⁶ The population universe in Table 7 is all test records indicating both a place of birth outside of the United States and a response to the item asking for more than one arrival (“Have you come to live. . .MORE THAN ONCE?”) in both the original Content Test version and the CFU.

simple response variance for the given response category, and it is a measure of reliability or consistency. IOI values of less than 20 percent indicate high reliability (H), 20 to 50 percent indicate a moderate level of reliability (M), and over 50 percent indicates low reliability (L). About four in ten respondents indicated only one arrival in the original survey, whereas in the CFU, that ratio increased dramatically to about nine in ten. This difference is represented with a very high IOI of 93 percent. This large amount of inconsistency indicates that respondents had difficulty understanding the concept of the question. Since the CFU distribution is drastically different than the original distribution, it is likely that this conceptual error was corrected when respondents spoke to a trained CFU interviewer.

Table 7. Number of Arrivals, Test to CFU

| | Test | Follow-up | Net Difference Rate | | Index of Inconsistency | | |
|----------------|------|-----------|---------------------|---------------------|------------------------|---------------------|-------|
| | | | Estimate | Margin of Error (%) | Estimate | Margin of Error (%) | Level |
| Once | 39.4 | 87.2 | -47.7 | 4.6 | 93.4 | ±2.9 | H |
| More Than Once | 60.6 | 12.8 | 47.7 | 4.6 | 93.4 | ±2.9 | H |
| Total | 100% | 100% | | | | | |

Next, we present the results for the reported year of arrival in both the test and control versions of the question. For this analysis the response given in the control version of the question is compared to the year of only or most recent arrival in the CFU. In the test version, the year of only or most recent arrival provided of the question will be compared to the year of only or most recent arrival in the CFU. A net difference rate will be used to determine how consistently respondents report these years.

The net difference rate (NDR) is used when we assume that the CFU interview, which asks more questions and collects more detailed data about a topic, provides a better measure than the control or test versions of a question. The NDR reflects the net change between the original response and the response given for the more detailed CFU questions. In other words, since we assume the CFU provides better data, the NDR indicates to what extent the test or control version of a question over- or underestimates the topic (or category) of interest. Relative to the CFU estimate, an NDR with a negative value indicates an underestimate and a positive value indicates an overestimate. An NDR that does not statistically differ from “0” indicates that the question asked in the original test or control interview produces results similar to the more detailed question set asked in CFU. In other words, the question should not result in a systematic over- or underestimate of the topic (or category) of interest.

For the purpose of this evaluation, we compared the NDR calculated for the test group to that of the control group to assess which version of the question resulted in more systematic error, regardless of whether the error reflected an over- or underestimate. Thus, the analysis also provides the “Absolute Net Difference” rate by taking the absolute value of the NDR for control group and the test group.

Table 8. Year of Arrival: Reinterview Comparison (Control Panel to Test Panel)

| | Net Difference Rate | | | | | Index of Inconsistency | | | | | |
|----------------------------|---------------------|------------------|------------------|--------------|--------|------------------------|------------------|----------|--------------|--------|-------|
| | Control vs. CFU (%) | Test vs. CFU (%) | Diff T - C (%) | Marg Err (%) | Signif | Control vs. CFU (%) | Test vs. CFU (%) | Diff (%) | Marg Err (%) | Signif | Level |
| Entered U.S. 2000 or later | -1.9 | 2.5 | 0.6 | 4.0 | No | 12.7 | 16.8 | 4.1 | 11.2 | No | L |
| Entered U.S. 1990 to 1999 | 0.5 | -0.1 | -0.4 | 1.8 | No | 16.5 | 12.2 | -4.3 | 5.3 | No | L |
| Entered U.S. 1980 to 1989 | 0.8 | -1.3 | 0.5 | 1.6 | No | 15.2 | 11.3 | -3.9 | 6.1 | No | L |
| Entered U.S. before 1980 | 0.6 | -1.1 | 0.5 | 1.7 | No | 7.4 | 7.3 | -0.1 | 5.5 | No | L |

Table 8 compares the test and control NDR and IOI for the year of arrival question at a decade level. Census tabulated data generally groups the individual years of entry into decades or periods (e.g. Before 1980, 1980 to 1989, etc.).¹⁷ Therefore, consistency based upon this broader categorical definition is used to measure the accuracy of the tabular data currently accessible to the public. At a decade level, consistency seems to be high and there are no significant differences between versions. However, it should be noted that the comparison between the year given in the original interview and the year given in the CFU was made regardless of a difference in the number of arrivals that the respondent reported in each version.

Although these questions provide consistent results at a decade or period level, it is important to investigate whether respondents consistently report the same year of arrival. To determine the accuracy of the exact year of arrival reported, a comparison was made between test and control to see whether the year that respondents reported in the original treatment matched the year that they reported in the CFU (Table 9).¹⁸ A year-to-year or exact match occurs if the reported year in the original Content Test version is exactly the same as the reported year in the CFU. A “plus or minus” match occurs if the reported year in the original Content Test version matches the year reported in the CFU within the range indicated (i.e. within two years or within four years).

¹⁷ The population universe in Table 8 is all records indicating a place of birth outside of the United States and a year of arrival (only or most recent) in both the original Content Test versions and the CFU.

¹⁸ The population universe in Table 9 for the control panel is all control records indicating a place of birth outside of the United States and a year of arrival in the original Content Test and all control records that provided a year of only or most recent arrival in CFU. The universe for the test panel (Only or Most recent arrival) is all test records indicating both a place of birth outside of the United States and a year of only or most recent arrival in the original Content Test and the CFU. For First arrival, the universe is all test records indicating: a place of birth outside of the United States; more than one arrival in the original Content Test and the CFU, and; a year of first arrival in the original Content Test and the CFU.

Table 9. Year of Arrival: Reinterview Comparison, Control and Test to CFU

| | Test Panel: | | | | | | | | |
|---------------|----------------------------------|----------------|----------------|-----------------------------|----------------|----------------|---------------|----------------|----------------|
| | Control Panel: 1 “year of entry” | | | Only or Most recent arrival | | | First arrival | | |
| | Exact Match | ± 2 Year Match | ± 4 Year Match | Exact Match | ± 2 Year Match | ± 4 Year Match | Exact Match | ± 2 Year Match | ± 4 Year Match |
| 2000 or later | 70.3 | 94.1 | 97.0 | 68.0 | 80.5 | 81.6 | 63.8 | 95.9 | 95.9 |
| 1990 to 1999 | 63.5 | 83.7 | 89.3 | 74.0 | 90.0 | 93.6 | 63.1 | 74.9 | 85.1 |
| 1980 to 1989 | 63.5 | 84.1 | 87.1 | 66.1 | 84.1 | 88.8 | 54.7 | 83.2 | 86.7 |
| Before 1980 | 60.2 | 84.7 | 88.1 | 65.7 | 87.7 | 90.5 | 73.1 | 84.3 | 88.0 |

When reporting their year or years of arrival, test and control respondents were both marginally consistent in providing the exact same year that they originally reported (Table 9). Roughly 60 to 70 percent of years given in the original control version matched exactly to those given in the CFU. Of the only or most recent (in the case where there is more than one year of arrival) years of arrival provided in the original test version, approximately 66 to 74 percent matched the year of only or most recent arrival provided in the CFU. For the respondents who arrived more than once, the year they provided as their first year of arrival in the original test version matched the year of first arrival in the CFU about 55 to 73 percent of the time. Because the current ACS year of entry question is intended to capture most recent arrival, only arrival and most recent arrival were combined in Table 9. This allows us to compare what data is currently available (the control item providing one year of arrival) with the data that would be available if the test question were accepted (only and most recent arrival). Information on the year of first arrival is provided for informational purposes only.

Broadening the definition of consistency yielded better results. About 84 percent or more of the reported years of arrival in the reinterview were within two years of the year originally reported in the control version. About 81 percent or more of the reported years of only or most recent arrival in the reinterview were within two years of the only or most recent years of arrival provided in the test version. These proportions increased even more when the window of acceptability was widened to four years.

5.4.2 Which of the two versions has the lowest item nonresponse rates?

When comparing control panel year of entry response rates with test panel year of first or most recent arrival response rates, no statistically significant differences were found between the estimates. In the control panel, 22 percent of respondents did not provide a year of arrival, compared with 23 percent of respondents in the test panel (Table 10).¹⁹ Although these

¹⁹ The population universe in Table 10 for the control panel is all control records indicating a place of birth outside of the United States. For the test panel “Arrived More than Once?” item, the universe is all test records indicating a place of birth outside of the United States. For the test panel Year of most recent arrival item, the universe is all test records indicating both a place of birth outside of the United States and more than one arrival.

estimates were not different, both are unacceptably high according to ACS standards. In addition, in the test version, about 23 percent of respondents did not respond at all to the second part of the question (“Has this person come to live, work, or study in the United States MORE THAN ONCE?”). Of those respondents who did answer the second part of the question and indicated that they arrived more than once, an alarming 83 percent did not provide a year of most recent arrival. This high nonresponse rate for reporting a year of most recent arrival suggests confusion about the intent of the “Arrived more than once?” item by persons who arrived only once. This confusion was also evident in the sizeable proportion of respondents who reported multiple arrivals but indicated the same first and most recent arrival year (about 27 percent). While it is plausible that respondents could have entered more than once in the same year, this proportion appears high. Since the year of most recent arrival is the year of interest for this question, it is clear that the test version would not be acceptable.

Table 10. Year of Arrival: Non-response rates

| | Control (%) | Test (%) | Difference (%) | Margin of Error | Significant |
|-----------------------------|-------------|----------|----------------|-----------------|-------------|
| Year of Arrival | 21.7 | 23.1 | 1.4 | 2.8 | No |
| Arrive More than Once? | N/A | 22.6 | N/A | 3.2 | N/A |
| Year of Most Recent Arrival | N/A | 83.3 | N/A | 2.3 | N/A |

5.4.3 Do the year of entry values provided in the control version reflect a first year of arrival, most recent year of arrival or something else?

For this research question, one way in which we can determine the year to which respondents were referring in the control version is to look solely at control version respondents who indicated in the CFU that they arrived more than once. As mentioned earlier, the CFU indicated that about 13 percent had arrived more than once.

When compared to the years of first and most recent arrival provided in the CFU, the year given in the original control version matched both years about 18 percent of the time (Table 11).²⁰ An example of this would be if a respondent reported arriving in 2003 in the original Content Test control version (where they were only allowed to provide one year of arrival), and then in the CFU indicated that they arrived more than once and reported 2003 for both first and most recent arrival years.

The arrival year respondents provided in the control version of the question exactly matched the most recent year of arrival provided in the CFU about 19 percent of the time (about 21 percent of the time the two years provided were in the same decade). It matched the first year of arrival in

²⁰ The population universe in Table 11 is all control records indicating birth outside of the United States, a year of entry in the original Content Test version, and first and most recent arrival years in the CFU.

the CFU about 31 percent of the time. Thus, 33 percent of the time it represents some other year (not first or most recent). These results show that the year of arrival given in the control version does not seem to consistently represent either first or most recent year of arrival for those who entered the United States more than once. This should be taken into consideration when developing instructional materials and providing training for data collection, which are currently designed to capture year of most recent arrival.

**Table 11. Year of Arrival: Reinterview Comparison
(Control to CFU)**

| | Entered More than Once | |
|-------------------------------------|------------------------|-------------------------|
| | Year-to-Year Match (%) | Decade/Period Match (%) |
| Matches Both | 18.0 | 44.6 |
| Matches Year of First Arrival | 30.6 | 25.1 |
| Matches Year of Most Recent Arrival | 18.8 | 21.3 |
| No Match | 32.6 | 9.1 |
| Total | 100.0 | 100.0 |

6. SUMMARY OF EMPIRICAL RESULTS

6.1 Place of Birth

Asking for a lower level of geography than state or country did not impact the state or country of birth response rate nor did it impact the distribution of targeted foreign-born country of birth responses. Furthermore, the detailed place of birth question adequately allowed for coding to a higher level of geography. While the manual coding of detailed place of birth took approximately two to three weeks, the system used for the Content Test is currently not feasible for full production.

Selection Criteria Met

- The state or country nonresponse rate for the test version will be equal to or less than the rate for the control version.
- The distribution of country of birth responses will be roughly equivalent between the test and control versions.
- At least 70 percent of the responses for the non-U.S. born will be codeable to a distinct region for each targeted country.

Conclusion

Although empirical evidence suggests the test version performed better, this version lacks operational feasibility at this time.

6.2 U.S. Citizenship

Changing the U.S. citizenship status question to include year of naturalization and referencing U.S. citizen parents neither reduced nor negatively impacted the proportion of respondents reporting that they were naturalized. Additionally, the changes did not reduce the response rate for citizenship status and the year of naturalization response rate was acceptable.

Selection Criteria Met

- The item nonresponse rate for the test panel will be equal to or less than the item nonresponse rate for the control.
- The estimate of naturalized citizens from the test panel will be equal to or less than the control estimate.

Conclusion

Although these criteria were met, the empirical evidence does not conclusively support one panel over the other. However, the year of naturalization write-in in the test panel provides operational advantages, as it would substantially aid editing procedures.

6.3 Year of Arrival

The control and test version had roughly equivalent item nonresponse rates for year of arrival, although nonresponse was high for both. Overall, there were no differences between test and control in terms of consistency at a decade level. However, the accuracy of number of arrivals for the test version was poor, with large differences in the proportion of respondents indicating multiple arrivals in the original interview versus single arrivals in the CFU. For both the test and control, the reporting of the year arrived was marginally consistent when considering a year-to-year match, but good when broadening the definition of consistency to include within two or four years of the original response provided.

Examining the reinterviews for the control, respondents determined in the CFU to have had more than one arrival more often treatment had reported their first year of arrival than their most recent year of arrival in the original. However, the original year reported quite often did not match either their year of first arrival nor their most recent year of arrival. In some instances, the original year provided was not even in the same decade as the years reported in the CFU for first arrival or last arrival. Therefore, it is unclear what year respondents with more than one arrival are providing most often in the control version of the question.

Selection Criterion Met

- The item nonresponse rate for the test version will be equal to or less than the item nonresponse rate for the control version.²¹

Selection Criteria Not Met²²

- The number of arrivals was reported inconsistently.
- The period of entry was marginally consistent in providing a year-to-year match and fairly consistent when providing a period/decade match.

Conclusion

The empirical evidence suggests the control panel performed better than the test panel, although comparison to the CFU suggested concerns for the control as well.

²¹ However, there was high nonresponse for both versions.

²² This refers to the adapted selection criteria.

References

Asiala, M. and A. Navarro. (2006). "Experimental Design for the 2006 American Community Survey Content Test," American Statistical Association 2006 Proceedings of the Section on Survey Research Methods [CD-ROM].

Camarota, S.A. 2003. *Assessing the Quality of Data Collected on the Foreign Born: An Evaluation of the American Community Survey (ACS)*. U.S. Census Bureau. Unpublished.

Edmonston, B. 1996. *Statistics on U.S. immigration: An Assessment of Data Needs for Future Research*. Washington, D.C. National Academy Press.

Fix, M., J.S. Passel, and K. Sucher. 2003. *Trends in Naturalization* (Urban Institute).
<http://www.urban.org/url.cfm?ID=310847>

Gage, L. No date. *Comparison of Census 2000 and American Community Survey 1999-2001 Estimates San Francisco and Tulare Counties, California*. California Department of Finance. Research done under contract with the U.S. Census Bureau.

Grieco, E. 2003. *Census 2010 and the Foreign Born: Averting the Data Crisis*. Migration Policy Institute.
<http://www.migrationpolicy.org/pubs/MPIPPolicyBriefCensus.pdf>

Kerwin, J., S. Heltemes, S. Berry, D. Nelson, and M. Popovic. 2005. *Cognitive Testing of Proposed Items on International Migration for the American Community Survey*. Report prepared by Westat for the U.S. Census Bureau. Unpublished.

Massey, D.S. and C. Capoferro. 2004. "Measuring Undocumented Migration." *International Migration Review*, 38:1075-1102.

Myers, D. 2004. "Accuracy of Data Collected by the Census Question on Immigrants' Year of Arrival." (Working Paper PDRG04-01) University of Southern California.

National Research Council. 2004. *The 2000 Census: Counting Under Adversity*. Panel to Review the 2000 Census, C.F. Citro, D.L. Cork, and J.L. Norwood (eds.). Committee on National Statistics. Washington, D.C.: National Academy Press.

National Research Council. 2001. *The 2000 Census: Interim Assessment*. Panel to Review the 2000 Census, C.F. Citro, D.L. Cork, and J.L. Norwood (eds.). Committee on National Statistics. Washington, D.C.: National Academy Press.

National Research Council. 1996. *Improving Data on America's Aging Population: Summary of a Workshop*. Committee on National Statistics, Committee on Population, Commission on Behavioral and Social Sciences and Education; D. Carr, A. Pemmarazu, and D.P. Rice (eds.). Washington, D.C.: National Academy Press.

http://books.nap.edu/books/0309056330/html/18.html#page_middle

Poros, M. and A. Orum. 2005. *Report on Cognitive Testing of Proposed International Migration Questions for the American Community Survey*. Report prepared by Sabre Systems for the U.S. Census Bureau. Unpublished.

Redstone, I. and D. Massey. 2003. "Coming to Stay: An Analysis of the U.S. Census Question on Year of Arrival." Paper presentation at the Population Association of America annual meeting, Minneapolis, Minnesota, May 1-3, 2003.

Schmidley, A.D. and J.G. Robinson. 2003. *Measuring the Foreign-Born Population in the United States with the Current Population Survey: 1994-2002*. Population Division Working Paper Series, No. 73. U.S. Census Bureau, Washington D.C.

Schmidley, A.D. and J.G. Robinson. 1998. *How Well Does The Current Population Survey Measure The Foreign-Born Population In The United States?* Population Division Working Paper Series, No. 22. U.S. Census Bureau, Washington D.C.

U.S. Census Bureau. 2004. *Content and Data Quality in Census 2000*. Census 2000 Topic Report No. 12, Census 2000 Testing, Experimentation, and Evaluation Program. Washington, D.C.: U.S. Government Printing Office.

U.S. General Accounting Office, 1998a. *Immigration Statistics: Status of the Implementation of National Academy of Sciences' Recommendations* (GAO/GGD-98-119) 18 pp.

U.S. General Accounting Office. 1998b. *Immigration statistics: information gaps, quality issues limit utility of federal data to policymakers*. (GAO/GGD-98-164) 87 pp. Washington, D.C.
<http://www.gao.gov/archive/1998/gg98164.pdf>

United Nations. 1998. *Recommendations on Statistics of International Migration*.
http://unstats.un.org/unsd/publication/SeriesM/SeriesM_58rev1E.pdf

Van Hook, J., J. Passel, W. Zhang, and F. Bean. 2004. *Foreign-Born Emigration: New Approaches and Estimates Based on Matched CPS Files*. U.S. Census Bureau, unpublished.

Wellens, T. 1993. *Report on the Cognitive Evaluation of the Nativity Questions for the Current Population Survey*. Center for Survey Methods Research. U.S. Census Bureau, Washington DC.

Appendix A: Content Test Information Pages

CONTENT TEST INFORMATION PAGE

For

PLACE OF BIRTH (no CFU required)

Question Wording:

| Current ACS Wording | Content Test Wording |
|---|---|
| <p>Where was this person born?</p> <p><input type="checkbox"/> In the United States – <i>Print name of state.</i></p> <p style="margin-left: 40px;"><input style="width: 200px; height: 15px;" type="text"/></p> <p><input type="checkbox"/> Outside the United States – <i>Print name of foreign country, or Puerto Rico, Guam, etc.</i></p> <p style="margin-left: 40px;"><input style="width: 200px; height: 15px;" type="text"/></p> | <p>Where was this person born?</p> <p><input type="checkbox"/> Inside the United States – <i>Print name of state.</i></p> <p style="margin-left: 40px;"><input style="width: 200px; height: 15px;" type="text"/></p> <p><input type="checkbox"/> Outside the United States – <i>Print name of foreign country, or Puerto Rico, Guam, etc.</i></p> <p style="margin-left: 40px;"><input style="width: 200px; height: 15px;" type="text"/></p> <p style="margin-left: 40px;"><i>Print name of city/town/village of foreign country.</i></p> <p style="margin-left: 40px;"><input style="width: 200px; height: 15px;" type="text"/></p> |

Research Questions & Evaluation Measures:

| No. | Research Questions | Evaluation Measures |
|-----|--|---|
| 1. | Does asking for a lower level of geography (city/town/village) for place of birth impact item nonresponse rates for this item? | Compare item nonresponse rates (at the country level) between test and control. Compare item non-response at the lower level of geography for the test version, as compared to higher level of geography for the control |
| 2. | Does asking for a lower level of geography (city/town/village) for place of birth impact the distribution of country of birth? | Compare distributions of country of birth between the test and control versions |
| 3. | What proportion of the foreign born responses can be uniquely coded to a region within the foreign country of birth? | Based on the information provided by respondents to both the country of birth and the city/town/village question, what proportion of responses code to a unique or specific region in the foreign country? What proportion cannot be coded to any region in the foreign country? What proportion could code to more than one region in the foreign country? |

| | | |
|----|---|--|
| | | <i>NOTE: This analysis will focus only on coding for the following countries: US, Mexico, China, Philippines, India, Other countries</i> |
| 4. | How much additional time is required to code cases at a regional level rather than a country level only? (Note that this research question is for informational purposes only) | Estimate the amount of time required to complete coding for an ACS production size sample, starting at the creation of the coding file through the receipt of the final codes between test and control versions. |

Selection Criteria:

| Research Q | Criteria |
|-------------------|---|
| 1. | The item nonresponse rates for the test version will be equal to or less than the rates for the control version. |
| 2. | The distribution of place of birth at the country level (higher level of geography) will be roughly equivalent between the test and control versions |
| 3. | At least 70% of the responses from respondents reporting birth outside of the U.S. will code to distinct region for each of the targeted countries for analysis. |

All criteria must be met in order to accept the test version.

CONTENT TEST INFORMATION PAGE
For
CITIZENSHIP (no CFU required)

Question Wording:

| Current ACS Wording | Content Test Wording |
|---|---|
| <p>Is this person a CITIZEN of the United States?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Yes, born in the United States→<i>SKIP to 10a</i> <input type="checkbox"/> Yes, born in Puerto Rico, Guam, the U.S. Virgin Islands, or Northern Marianas <input type="checkbox"/> Yes, born abroad of American parent or parents <input type="checkbox"/> Yes, U.S. citizen by naturalization <input type="checkbox"/> No, not a citizen of the United States | <p>Is this person a citizen of the United States?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Yes, born in the United States→<i>SKIP to 10a</i> <input type="checkbox"/> Yes, born in Puerto Rico, Guam, the U.S. Virgin Islands, or Northern Marianas <input type="checkbox"/> Yes, born abroad of a U.S. citizen parent or parents <input type="checkbox"/> Yes, a U.S. citizen by naturalization – <i>Print year of naturalization</i> <div style="border: 1px solid black; width: 100px; height: 20px; margin: 5px 0;"></div> <input type="checkbox"/> No, not a citizen of the United States |

Research Questions & Evaluation Measures:

| No. | Research Questions | Evaluation Measures |
|-----|--|--|
| 1. | Do the following changes to the citizenship question reduce the estimate of naturalized citizens? <ul style="list-style-type: none"> - adding a write-in field for year of naturalization - changing the born abroad category to specifically reference U.S. citizen parents | Compare estimates of naturalized citizens between the test and control panels. |
| 2. | Do the changes to the citizenship question impact item nonresponse rates for the question? | Compare item nonresponse rates between the test and control panels. |

Selection Criteria:

| Research Q | Criteria |
|------------|--|
| 1. | Estimates of naturalized citizens from the test version will be equal to or less than the estimates from the control version. |
| 2. | Item nonresponse rates for the citizenship question will be equal to or better than the item nonresponse rates for the control version. |

Each criteria must be met in order to select the test version.

CONTENT TEST INFORMATION PAGE
For
YEAR OF ENTRY (CFU required)

Question Wording:

| Current ACS Wording | Content Test Wording |
|---|---|
| <p>When did this person come to live in the United States? <i>Print numbers in boxes.</i></p> <p>Year</p> <div style="border: 1px solid black; width: 100px; height: 20px; margin-top: 10px;"></div> | <p>When did this person FIRST come to live, work or study in the United States? <i>Do not include holidays, short business trips, or other brief visits.</i></p> <p>Year of arrival</p> <div style="border: 1px solid black; width: 100px; height: 20px; margin-top: 10px;"></div> <p>Has this person come to live, work or study in the United States MORE THAN ONCE? <i>Do not include holidays, short business trips, or other brief visits.</i></p> <p><input type="checkbox"/> Yes – <i>print year of most recent arrival</i></p> <div style="border: 1px solid black; width: 100px; height: 20px; margin-left: 100px; margin-top: 5px;"></div> <p><input type="checkbox"/> No</p> |

Research Questions & Evaluation Measures:

| No. | Research Questions | Evaluation Measures |
|-----|---|--|
| 1. | Which of the two versions has less systematic response error associated with the response of entering the U.S. once or more-than-once? | Compare the net difference rate for year of entry between the two versions (<i>using decades as the unit of measure</i>) |
| 2. | Which of the two versions has the lowest item nonresponse rates | Compare the item nonresponse rates between the two versions. |
| 3. | Do the year of entry values provided in the control version reflect a first year of arrival, most recent year of arrival or something else? | For informational purposes only: If we accept the test version, this will help us provide information to users regarding how to interpret the data from previous years relative to how the data are collected in the test version. |

Selection Criteria:

| Research Q | Criteria |
|-------------------|--|
| 1. | The net difference rate will be lower in the test version than the control version (at a decade level). |
| 2. | The item nonresponse rate will be equal to or less than the control version |
| 3. | For informational purposes only |

Criteria for selecting the test version:

- The net difference rate must be lower in the test version, and
- The item nonresponse rates must be equal or lower in the test version.

Appendix B: Westat Cognitive Testing Report

Cognitive Testing of Proposed Items on International Migration for the American Community Survey

Authors:

Jeffrey Kerwin
Susan Heltemes
Heather Berry
Dawn Nelson
Martha Popovic

April 26, 2005

Prepared for:

U.S. Census Bureau
4700 Silver Hill Road
Washington, DC 20548

Prepared by:

WESTAT
1650 Research Boulevard
Rockville, Maryland 20850

TABLE OF CONTENTS

| | | <u>Page</u> |
|----------|--|-------------|
| | Executive Summary | ii |
| 1 | Background | 1 |
| 2 | Methods | 2 |
| | 2.1 Study Design | 2 |
| | 2.2 Recruitment | 2 |
| | 2.3. Interview Logistics and Procedures..... | 4 |
| 3 | Findings | 6 |
| | 3.1 Reporting State/Province/Region of Foreign Country | 6 |
| | 3.2. Reporting Citizenship and Year of Naturalization | 9 |
| | 3.3. Reporting Years of First and Most Recent Arrival in U.S. | 11 |
| 4 | Conclusions | 14 |

APPENDICES

| | | |
|----|--|----|
| A. | ACS Migration Items Tested..... | 15 |
| B. | Recruiting Screener | 19 |
| C. | Cognitive Interview Guide (mail version)..... | 24 |
| D. | Westat Recommendations for the ACS Migration Items | 31 |

Executive Summary

Background

The United States Census Bureau recently implemented the household component of the American Community Survey (ACS). Each month a paper questionnaire will be mailed to about 250,000 addresses. The Census Bureau will collect data via Computer Assisted Telephone Interviewing (CATI) and Computer Assisted Person Interviewing (CAPI) with those households not responding to the mail survey.

The Census Bureau is considering collecting new data for the ACS related to international migration. The proposed new questions include:

- Collecting information regarding state, province or region within the country of origin for persons emigrating to the United States;
- Collecting year of naturalization for immigrants who become naturalized citizens;
- Collecting the year of first arrival and the year of the most recent arrival for persons immigrating to the United States more than once.

The Census Bureau asked Westat to conduct cognitive interviews with individuals residing in households containing one or more persons born outside the United States. The purpose of the task was to examine the cognitive aspects of responding to questions on international migration, especially aspects that might contribute to response error in the proposed new questions.

Method

Westat conducted 40 cognitive interviews with individuals from the Washington, DC area utilizing both mail (self-administered) and telephone (interviewer administered) survey modes. Three groups of respondents were recruited for the study:

- Households with unrelated members;
- Multi-generational/extended family households;
- Simple family unit households.

Participants were recruited through a variety of measures including advertisements in newspapers, on the internet, “word of mouth” referrals from Westat staff, and flyers placed in various locations. The volunteers were screened for eligibility prior to being interviewed. The interviews were conducted at Westat’s headquarters in Rockville, MD, or in a hotel conference room in Washington, DC. Almost all of the interviews were audio-taped (with the participant’s permission) and participants were paid a \$40 incentive.

Findings and Recommendations

State/Province/Region of Foreign Country

A new item proposed for the ACS would ask for the “*state/province/region*” of the country of origin for each foreign born person in the household. Participants often answered incorrectly by instead giving the city or town where the person was born. Participants often did not have knowledge of the appropriate administrative subdivision. This was true even when participants were answering for themselves. Administrative subdivisions below the country level and above the city level do not seem to be as salient for many people from other countries, as compared to the United States. Westat recommends that the Census Bureau consider asking for the city or town in which foreign-born individuals were born, rather than the state, province, or region.

Citizenship and Year of Naturalization

The ACS determines whether foreign-born residents are citizens of the United States and if so, whether they are citizens due to having been born in a United States territory, born abroad to United States citizen parents, or naturalized. For naturalized citizens, a proposed new question would ask for the year of naturalization. Participants were able to provide the year of naturalization for themselves and family members, but had much more difficulty reporting this information for unrelated household members. Also, of the six participants who reported for someone born abroad to United States citizen parents, 3 reported incorrect category of citizenship. Westat recommends that the Census Bureau consider revising the response category “yes, born abroad of a U.S. citizen parent or parents” to “born of at least one U.S. citizen parent.”

Year of First and Most Recent Arrival

The ACS also collects information regarding the year each foreign-born person came to live in the United States. A proposed revision would ask when the foreign-born person “*come to live, work or study*” in the United States. Additionally, for foreign-born persons who have moved to the United States more than once, a proposal was made to differentiate the year of one’s first arrival from that of the most recent arrival.

Significant confusion arose concerning interpretation of the phrase “*entered the U.S. more than once*.” Participants frequently reported dates of vacations and short visits outside the United States. But we obtained reactions to possible alternative wordings of this question. One of these alternatives (“*if this person has entered the U.S. more than once (not including holiday, vacations or other brief visits)*”) appeared successful in clarifying the Census Bureau’s intent. Thus, Westat recommends the use of this alternative wording for collecting data on years of entry into the United States.

The full set of Westat recommendations for the ACS migration questions, for both the mail and telephone mode, is shown in Appendix D.

1. BACKGROUND

The U.S. Census Bureau recently completed full implementation of the household component of the American Community Survey (ACS). Through multi-mode data collection, it will provide annual estimates of important demographic, socio-demographic and housing characteristics for the nation, states, and counties. Each month, a paper questionnaire will be mailed to about 250,000 addresses. The Census Bureau will collect data through Computer Assisted Telephone Interviewing (CATI) and Computer Assisted Person Interviewing (CAPI) from households not responding to the mail survey.

The Census Bureau is considering proposals for new questions related to international migration, as well as revisions to the current questions on this topic. The proposed additions of interest include:

- Collecting information as to the state, province, or region of birth within the country of origin for persons who have emigrated to the U.S.;
- For immigrants who are naturalized citizens of the U.S., collecting information as to the year of naturalization;
- For persons who have emigrated to the U.S. more than once, collecting both the year of their first arrival, as well as the year of their most recent arrival in the U.S.

The Census Bureau asked Westat to conduct cognitive interviews with persons residing in households containing one or more persons born outside the U.S. The purpose of this task was to examine the cognitive aspects of responding to the questions on international migration, especially aspects that might contribute to response error in the proposed additions noted above. For example, we sought to determine whether respondents' interpretations of the questions and response categories match the Census Bureau's intent. In addition, we examined whether respondents have difficulties with the wording of the questions, or difficulties in recalling the information requested. Both mail and telephone versions of the migration questions were tested.

Appendix A presents the questions on international migration we tested in this study, for both the mail and telephone modes.

2. METHOD

We conducted cognitive interviews with 40 individuals from the Washington, DC area. In this chapter we discuss the study design, our process for recruiting study participants, and describe the participants. We also describe our procedures for conducting the cognitive interviews.

2.1 Study Design

The Census Bureau requested that we cognitive test the ACS items of interest in forms suitable for two survey modes: a) mail (self-administered), and b) telephone (interviewer-administered). The Census Bureau further requested that we test the items with three groups of respondents:

- Households with unrelated members;
- Multi-generational/extended family households;
- Simple family unit households.

Table 2-1 shows how respondents were distributed across the interview modes completed.

**Table 2-1.
Respondent Groups by Mode of Cognitive Interview**

| | Mail | Telephone | Total |
|---|-----------|-----------|-----------|
| Households with unrelated members | 10 | 6 | 16 |
| Multi-generational / extended family households | 7 | 4 | 11 |
| Simply family unit households | 8 | 5 | 13 |
| Total | 25 | 15 | 40 |

2.2 Recruitment

We recruited participants for this project through a variety of means. We placed an ad for research volunteers in the *Gazette*, a weekly newspaper serving Montgomery, Prince George’s, and Frederick counties of Maryland. We also placed ads in the *Washington City Paper* and *El Tiempo Latino* (a local Spanish-language newspaper). The ads specified a need for volunteers from households with one or more members born outside of the U.S. Additional recruiting was done over the Internet through ads posted to www.craigslist.com and several neighborhood email newsgroups, as well as flyers posted in various locations. Approximately 175 persons contacted Westat in response to these ads, and this group largely served as our pool of potential study participants. A small number of persons were recruited by “word of mouth” referrals from Westat staff (Westat employees were not eligible for participation) and study participants. Volunteers

were screened for eligibility on an “as needed” basis prior to being scheduled for interviews their household. The recruiting screener used for this study is included in Appendix B.

Twenty-four of the 40 participants were born outside of the U.S., with the remaining participants qualifying on the basis of other household members being foreign-born. English was a second language for 17 participants. All of the interviews were conducted in English, which may have influenced the results. Altogether, the test ACS questionnaire covered 135 household members, 75 of whom were born outside the U.S. The ACS items were asked about persons who had migrated to the U.S. from a wide range of countries and continents:

- Central/South America (10 interviews)
- Asia (10 interviews)
- North America (Mexico, Canada, Jamaica) (8 interviews)
- Europe (8 interviews)
- Mid-East (3 interviews)
- Africa (3 interviews)

Table 2-2 below presents a demographic summary of the 40 study participants.

**Table 2-2
Demographic Summary of Cognitive Interview Participants**

| Demographic Characteristics | Number of Participants |
|-----------------------------|------------------------|
| Gender | |
| Male | 17 |
| Female | 23 |
| Age | |
| 18-30 | 11 |
| 31-45 | 17 |
| 46-60 | 11 |
| Over 60 | 1 |
| Education Completed | |
| Less than high school | 1 |
| High school/GED | 4 |
| Some college | 13 |
| College degree | 10 |
| Graduate school | 12 |
| Race/Ethnicity | |
| White | 13 |
| Black/African-American | 11 |
| Hispanic/Latino | 8 |
| Asian | 8 |
| American Indian | 1 |

Table 2-2 shows few participants in the study had low levels of education, and over half were college-educated. This may have influenced the study results to some degree. We suspect the comprehension problems discussed later in this report would be even more common among persons of low education backgrounds.

2.3 Interview Logistics and Procedures

The interviews were conducted by Westat staff and adhered to a semi-structured interview guide consisting of the following three elements:

- **Interview introduction:** Here we explained the purpose of the project and assured participants that all information they share will be treated as confidential. Participants were also informed that the interviews were to be audio-taped, and (if applicable) that Census Bureau staff were observing behind a one-way mirror. Participants were also asked to sign consent forms.
- **Administration of the ACS:** Participants were asked to respond to a subset of the demographic questions contained in the ACS for each member of their household. The ACS questions on housing characteristics were not included in the interview. Participants were asked to respond to these questions through one of two survey modes:
 - **Mail.** These participants were provided a paper-and-pencil copy of the ACS specially prepared for the cognitive interviews. They were given a pen and asked to fill it out at their own pace, but reading along and “thinking aloud” as they did so. The cognitive interviewer closely observed and took notes of any apparent difficulties experienced by the respondent.
 - **Telephone.** The cognitive interviewer entered the room behind the one-way mirror and called into the room where the participant sat. The participant answered a phone placed on the table upon hearing it ring. The interviewer then administered the ACS items in much the same way as a Census telephone interviewer would (reading the questions, as worded).
- **Cognitive debriefing:** Here, the Westat interviewer followed up on any observed confusion or difficulties that participants may have experienced when answering the key ACS items of interest. For example, verbal cues such as hesitation or changing one’s answer, as well as nonverbal cues such as eye rolling can suggest evidence of respondents’ difficulty in understanding a question. The interviewer also probed as to how he or she interpreted the key questions. Some probing was done concurrently with administration of a questionnaire, but most occurred after administration of the ACS. To obtain

additional data on how respondents interpreted the ACS items, the debriefing also presented respondents with a series of hypothetical scenarios, asking them how they would respond to an ACS item if the situation applied to them.

The guide for conducting the cognitive interviews (mail version) is shown in Appendix C.

The majority of the cognitive interviews were conducted onsite at Westat's headquarters in Rockville, MD. These interviews were conducted in a qualitative interviewing suite equipped with a one-way mirror for observation purposes. Census Bureau staff observed many of the interviews. Some interviews were conducted at a downtown Washington, DC hotel conference room. Interviews typically lasted 30 to 45 minutes. Participants received an incentive payment of \$40.

3. FINDINGS

3.1 Reporting State/Province/Region of Foreign Country

A new item proposed for the ACS would ask for the “state/province/region” of the country of origin, for each foreign-born person in the household:

| |
|--|
| <p>Where was this person born?</p> <p><input type="checkbox"/> Inside the United States—<i>Print name of state.</i></p> <p>_____</p> <p><input type="checkbox"/> Outside the United States—<i>Print name of foreign country, or Puerto Rico, Guam, etc.</i></p> <p>_____</p> <p><i>Print name of specific state/province/region of foreign country.</i></p> <p>_____</p> <p style="text-align: center;">(mail version)</p> |
|--|

We asked participants in the cognitive interviews to tell us what this question means to them, or tell us in their own words what the question is asking. Virtually all participants appeared to grasp the intent of the question and the information being requested:

“What part of the country you’re from.”

“More specific geographic place within a country.”

“Like in [the] U.S., they have 50 states. In China, they have 30 provinces.”

Although the question was easily understood, participants often answered incorrectly. That is, they reported a city or town in which the person was born rather than an appropriate administrative subdivision, or reported no information more specific than the country of origin. Across the 75 foreign-born persons covered in the interviews, incorrect information for “state/province/region” was obtained for 30 (40%) of them.¹ The response error was seen to a higher degree in the mail mode (25 incorrect among 51

¹ Useful information for evaluating the accuracy of reporting was obtained at the website *statoids.com.*, operated by Gwillim Law. Mr. Law is author of the book *Administrative Subdivisions of Countries: A Comprehensive World Reference, 1900 through 1998*, published by McFarland and Company in 1999.

foreign-born persons) than for the telephone mode (5 incorrect among 24 foreign-born persons). However, the source of the problem had nothing to do with survey mode – rather, participants generally lacked *knowledge* regarding the administrative subdivisions where their foreign-born household members were born.

Not surprisingly, participants were most often correct in the reporting of an appropriate administrative subdivision when answering for themselves, and most often incorrect when reporting for unmarried partners or unrelated household members (see Table 3-1). Participants often mentioned that if they were actually filling out the ACS at home, they would ask the relevant person for the appropriate answer, so the degree of response error would presumably be less than was observed in the cognitive interviews.

**Table 3-1.
Accuracy of Reporting Administrative Subdivision
as a Function of Relationship to Respondent**

| | Self | Family members | Unrelated roommates and partners | Total |
|------------------|------|----------------|----------------------------------|-------|
| Correct | 19 | 21 | 5 | 45 |
| Incorrect | 6 | 15 | 9 | 30 |

Although most of our 25 foreign-born participants could report an appropriate administrative subdivision in their country of origin, it was somewhat surprising that six were unable to do so. Some of these persons had moved to the U.S. as young children with their parents, and thus have little knowledge of the countries where they were born. Interestingly, some even said that their country of origin was not divided into districts in any way relevant for the question:

“I don’t think that Romania has anything that’s recognized as such...the only thing I would have is a city, and that is not ‘state/province/region.’”

In fact, Romania is divided into counties. Other respondents who believed their country of origin to have no administrative subdivisions included immigrants from Haiti (divided into departments), Trinidad (counties), and Scotland (UK subdivisions).

It is also worth noting that among those who answered correctly, several reported a name that represents both a city and a higher level administrative subdivision. For example, in each of the seven cases below, the respondent reported a *city* in which someone was born, stating they were unable to provide a state, region, or province. Nevertheless, the answer turned out to be an appropriate administrative subdivision:

- Kingston, Jamaica (capital city of the Kingston parish);
- Lima, Peru (capital city of the Lima department);
- Tokyo, Japan (capital city of the Tokyo prefecture);

- Dundee, Scotland (capital city of Dundee, a primary UK subdivision)
- Columbo, Sri Lanka (a district of Sri Lanka)
- Trieste, Italy (a province of Italy)
- Madrid, Spain (a province of Spain)

A number of observations suggest that administrative subdivisions below the country level, yet above level of city or town, do not appear to be as salient and meaningful for persons in (or from) some countries, compared to the U.S. This may partially explain the relative lack of knowledge many people in this study expressed about relevant administrative subdivisions in their home countries. For example, the woman discussed above who immigrated from Romania indicated that when she writes her relatives in that country, the address only includes the street, city, and the country (i.e., no state, province, or region). She and others also mentioned that when they meet other people from their home countries, they ask each other only what city or town they are from. Another participant, when asked to explain why he entered a city on the form, replied:

“I just assumed you wanted to know where I was born, so I put down [a city] ...I don’t understand why anyone would want to know where I was born other than the city.”

It could be that once immigrants come to the U.S., the relevant administrative subdivisions lose some of their salience. For example, another participant initially refused to enter the province of China she was from when filling out the form, on the grounds that it was not important information:

“No one in the U.S. cares what province in China someone comes from.”

Recommendation: The findings suggest that respondents to the new proposed question will often report a city or town, rather than the appropriate administrative subdivision intended by the question. Apparently, it is much easier for many to report a city or town, due to a lack of knowledge about states, provinces, regions, etc. Thus, the Census Bureau should consider asking foreign-born respondents to specify the city or town:

- **Mail mode:** *Specify city or town of foreign country.*
- **Telephone mode:** *In what city or town (was NAME / where you) born?*

3.2 Reporting Citizenship and Year of Naturalization

The ACS determines whether foreign-born residents are citizens of the U.S., and if so, whether they are citizens due to having been born in a U.S. territory, born abroad to a citizen parent or parents, or naturalized. If a resident is reported to have become a citizen through the naturalization process, a proposed additional element of the question would ask for the *year of naturalization* (i.e., the year in which their U.S. citizenship was granted):

Is this person a citizen of the United States?

- Yes, born in the United States—*Skip to 10a*
- Yes, born in Puerto Rico, Guam, the U.S. Virgin Islands, or Northern Marianas
- Yes, born abroad of a U.S. citizen parent or parents
- Yes, U.S. citizen by naturalization—*Print year of naturalization*
- No, not a citizen of the United States

(mail version)

Participants understood what this question is asking, and almost everyone had an accurate interpretation of “naturalization,” whether or not it applied to anyone in their household:

“That’s when they receive their citizenship....having taken the test and everything.”

“You apply to get citizenship. The process is called naturalization.”

“You get it through your parents or you apply for it. I got it through my parents. Children automatically get it when their parents become citizens.”

A few respondents did not understand what was meant by “naturalization” - these were persons within households where probing revealed that no one had gone through the naturalization process, and thus a lack of understanding of the term did not affect their answer. There were two exceptions, however. One respondent (in the telephone mode) asked the interviewer what was meant by “naturalization” when asked this question. The interviewer explained it was “the process of becoming a citizen.” The respondent thus

classified a roommate into this category because he is in the country attending college, and she assumed that you have to be a citizen in order to attend college. Later probing revealed that the roommate had been in the country only a very short time and thus is very unlikely to be a naturalized citizen. Another participant (also in the telephone mode) who was naturalized as a young child in the late 1940s insisted the term applied to his parents, but not to him. He believed that his citizenship was obtained through “derivation” (i.e., derived from his parents) rather than naturalization. He even stated that the term “derivation” appears on his citizenship papers.

Reporting the year of naturalization was generally quite easy for participants, especially when reporting for oneself. The ceremony in which one is sworn in as a citizen appeared to be a very salient event and thus easy to place in time. Some discussed vivid memories of making a trip to a courthouse in Baltimore and participating in the ceremony. Recall was more difficult (and thus answers less certain) for those where naturalization had occurred as a child, or when reporting for other family members. Still, participants seemed quite confident that they could estimate the year within a 2-3 year span of time. Among the 29 naturalized citizens covered in the cognitive interviews, a year of naturalization could not be reported for 5 – all were household members unrelated to the participant.

Finally, it should be noted that out of six participants who reported for someone born abroad to a U.S. citizen parent, three reported incorrectly:

- Two persons (mail mode) marked both the “Yes, *born in the U.S.*” and “Yes, *born abroad of a U.S. citizen parent or parents*” response categories. When probed, they both discussed how they think of their children as being “American” and so checked the first plausible response category without carefully reading all of the response options. Both ultimately realized on their own that they should have marked only one category.
- Another participant (mail mode) left the question blank for his grandchild born abroad in Jamaica. The grandchild’s father is a U.S. citizen, but the mother is not. This participant was not sure if the “Yes, *born abroad of a U.S. citizen parent or parents*” applied to his grandchild. After a few minutes, he reasoned that this probably would be the appropriate response choice, but it appeared likely that he would have left the question unanswered had he been filling out the form on his own.

Recommendation: The findings suggest that respondents to the ACS should generally be able to accurately report a year of naturalization. And though we see little reason to think the response option “Yes, *born abroad of a U.S. citizen parent or parents*” poses a serious problem, perhaps it could be revised so as to be more clear in situations where one parent is a U.S. citizen, and the other parent is not. Consider:

- **Mail mode:** *Yes, born abroad to at least one U.S. citizen parent.*
- **Telephone mode:** *(Was Name/Were you) born abroad to at least one U.S. citizen parent, or did (he/she/you) become a U.S. citizen by naturalization?*

3.3 Reporting Years of First and Most Recent Arrival in U.S.

After obtaining country of origin and citizenship status, the ACS collects information regarding the year each foreign-born person came to live in the U.S. A proposed revision would ask when did the person “come to live, work, or study” in the U.S., rather than “come to live.” Furthermore, for persons who have moved to the U.S. multiple times, a proposal was made to differentiate the year of one’s first arrival from that of one’s most recent arrival:

When did this person come to live, work, or study in the United States?

a. Print year of first arrival.→ _____

b. If this person has entered the U.S. more than once, print year of most recent arrival.→ _____

(mail version)

Participants generally provided responses to these items rather easily. Lack of knowledge appeared to be less of an issue here compared to the two other questions we tested. Only a couple of persons could not confidently answer for others in the household. However, in both the mail and telephone modes we did observe significant confusion over interpretation of the phrase “*entered the U.S. more than once.*” The problem was effectively articulated one of the participants in the telephone mode:

“I have difficulty answering that question. What do you mean ‘entered the United States?’ I’ve traveled abroad, like on vacations, so then I indeed returned to the United States. So I’m not sure what you mean by that.”

Several participants reported a year for their “most recent arrival,” since they believed that they (or another household member) had entered the U.S. more than once. But almost all of them did so incorrectly, since their answer represented only returning to the U.S. from a brief overseas visit, rather than a period of time living elsewhere. For example, one participant reported that he moved to the U.S. in 1962, and that his most recent arrival was 1969. He reported that his wife came here in 1970, and that her most recent arrival was in 1985. But both he and his wife have never lived elsewhere since moving to the U.S. – his answers for their most recent arrivals were based on brief visits made to their home country. Furthermore, even some who understood the intent of the question noted it was difficult to understand – as one put it, “*it’s awkwardly phrased.*” One participant filling out the mail version was quite confused by item (b) above, ultimately interpreting it correctly and leaving it blank. But he was far from certain about his interpretation and thus wrote what he assumed would be a helpful note in the margin of the questionnaire:

“I have entered 15 times over a 34 year period.”

The problem discussed above became apparent very early in the cognitive interviews, and so for many of the remaining interviews we obtained participants’ reactions to two alternative versions of item (b) above:

1. If this person has entered the U.S. more than once to live, work, or study, print year of most recent arrival.
2. If this person has entered the U.S. more than once (not including holidays, vacations, or other brief visits) print year of most recent arrival.

In general, both alternatives were viewed as being clearer than the original item printed in the test ACS questionnaire. But the second alternative was often said to be most effective, in that it communicates the intent without leaving one to infer that visits should not be considered when answering. In fact, a few participants did not correctly comprehend the intent of the item at all until they were shown the second alternative.

Although most of the problems with this question concerned determining if persons had entered the U.S. more than once, we also observed some confusion in response to item (a) above. The item asks participants to report the “year of first arrival,” and one person reported a year in which only a brief visit to the U.S. took place. A couple of other participants briefly wondered if they should base their answers here on visits as well, though ultimately they answered correctly. One specifically noted that “arrival” is a potentially troublesome term in this question:

“‘Come to live’ and ‘arrival’ are different things – because you can arrive for a holiday.”

Recommendation: The interviews revealed that these items are easily misinterpreted. To avoid confusion in reporting the year of “first arrival” in the mail survey, consider deleting this phrase. In order to minimize comprehension problems with reporting when the “most recent arrival” occurred, consider using the second alternative we discussed above because it seems to clarify the intent. The revised versions of these items would be as follows:

- **Mail mode:**

When did this person come to live, work, or study in the United States?

- a. Print year.→ _____*
- b. If this person has come to the U.S. more than once (not including holidays, vacations, or other brief visits), print year of most recent arrival.→ _____*

- **Telephone mode:**

What year did (<NAME>/you) FIRST come to live, work, or study in the United States?

(Have you / Has NAME) come to the United States more than once, not including holidays, vacations, or other brief visits??

- Yes, continue*
- No, DK or REF, skip out*

In what year did (you / NAME) most recently come to the United States?

4. CONCLUSIONS

In this chapter we discuss the major conclusions we believe can be drawn from this study. Of course, the usual cautions for interpreting cognitive interview results apply here. The participants were volunteers paid an incentive to respond and share their reactions to a part of the ACS instrument. Although the research setting was artificial, the interview findings described in this report provide useful insights into potential respondent concerns and sources of confusion underlying response errors with the proposed ACS questions related to international migration. The proposed changes and additions to the ACS should also be closely examined and tested in the field.

The conclusions we would derive from this study are as follows:

- It would be very difficult for many persons to report an administrative subdivision of birth for persons born outside the U.S. that is consistent with the Census Bureau's intent. This is true even for foreign-born persons who are responding to the ACS for themselves. On the other hand, it should be relatively easy to report a city or town. In fact, the interviews suggest that many respondents may well choose to report a city or town rather than a state, province, or region. To maximize consistency of reporting across respondents, and to minimize respondent burden, we have recommended that the Census Bureau simply ask for the city or town where persons migrating to the U.S. were born.
- Reporting a year during which citizenship was obtained through naturalization should be fairly easy for most respondents to report accurately. This is a very salient event in the lives of people who become citizens through this process.
- Asking for the "year of first arrival" and "year of most recent arrival" can potentially confuse many respondents and result in a substantial degree of misreporting. While respondents can easily recall the year they moved to the U.S., they can also easily misinterpret these questions as asking about the first and most recent times they have crossed the U.S. border (e.g., for a brief visit). These questions must be worded very carefully to minimize this problem. Hopefully, Westat's recommendations for revision will prove useful, but they should be further evaluated in the field.

Westat's full recommendations for the ACS migration questions, for both the mail and telephone modes, are shown in Appendix D.

Appendix A
ACS Migration Questions Tested

--Telephone Mode--

9a. (Was <Name>/Were you) born in the United States?

- Yes, born in the United States
- No (If No, skip to 9c)

If DK or Ref, skip to Question 10a

9b. In what state (was <Name> / were you) born?

Enter name of state, then skip to 12

9c. In what country (was <Name>/were you) born?

Enter country.

9d. In what state, province, or region of <Country Name> (was <Name>/were you) born?

Enter state/province/region.

10a. (Is <Name>/Are you) a citizen of the United States?

Yes - if R REPORTED A COUNTRY WHERE BORN,, go to 10b, otherwise go to 10c
No - go to 11a

DK and Refuse go to 13

10b. (Was <Name>/Were you) born abroad of a U.S. citizen parent or parents, or did (he/she/you) become a U.S. citizen by naturalization?

Born abroad of a U.S. citizen parent(s) - go to 11a
U.S. citizen by naturalization - go to 10d
Neither - go to 11a

--Telephone Mode—
(continued)

Ask if DK or ref place of birth, and 10a=yes (citizen)

10c. You have indicated that (<Name> is/you are) a citizen, and that you don't know (his/her/your) exact place of birth, but perhaps you could give us other general information about (him/her/yourself)...

(Was <Name>/Were you) born in the U.S., born in Puerto Rico, Guam, U.S. Virgin Islands or Northern Marianas, born abroad of a U.S. citizen parent or parents, or did (he/she/you) become a naturalized citizen?

1. Born in the U.S.
2. Born in Puerto Rico, Guam, U.S. Virgin Islands, or Northern Marianas - go to 11a
3. Born abroad of a U.S. citizen parent(s) - go to 11a
4. Citizen by naturalization - go to 10d
5. None of the above - go to 12

10d. In what year did (<Name>/you) become a naturalized U.S. citizen?

Enter year of naturalization

Go to 11a

11a. What year did (<NAME>/you) FIRST come to live, work, or study in the United States?

Enter year of first entry

11b. (Have you / Has NAME) entered the United States more than once?

Yes, continue to 11c

No, DK or REF, skip to 12

11c. In what year did (you / NAME) MOST RECENTLY enter the United States?

Enter year of most recent entry

Appendix B
Recruiting Screener

**International Migration Cognitive Interviews
- Participant Screener -**

Name: _____ Phone: _____ ID#: _____

A few (days/weeks) ago, you called us in response to an advertisement for research volunteers. We are looking for people willing to meet with us (here at Westat or at a conference room near Gallery Place in downtown DC) for a study being conducted for the U.S. Census Bureau. It will take an hour or less, and we're paying people \$40 for participating. We're scheduling interviews for Wednesday, Jan. 19th through Wednesday, Feb. 16th..

In order to find out if you are eligible to participate in this study, I need to ask you a few questions.

1. First, have you ever worked for Westat? (since 2000?)

YES :__ : [IF SINCE 2000, TERMINATE]
NO :__ :

2. What is your age?

:____ : [IF UNDER 18, TERMINATE]

3. RECORD GENDER. IF NOT OBVIOUS, ASK: Are you male or female?

MALE :__ :
FEMALE :__ :

4. What is the highest level of education you have completed?

LESS THAN HIGH SCHOOL :__ :
HIGH SCHOOL/GED :__ :
SOME COLLEGE/VOC. ED :__ :
COLLEGE DEGREE :__ :
GRADUATE SCHOOL :__ :

5. What is your race or ethnic background?

WHITE :____:
BLACK/AFRICAN AMER. :____:
HISPANIC/LATINO :____:
ASIAN :____:
SOMETHING ELSE :____:

6. How many people live in your household? → _____

7. Was anyone in your household born outside the United States? (this includes persons born to U.S. citizens who were living abroad) [PUERTO RICO AND OTHER U.S. TERRITORIES COUNT AS OUTSIDE THE U.S.]

YES, one person :____:
YES, two or more :____:
NO :____: [TERMINATE]

8. Where (was this person/were these persons) born?

9. (Was this person/Were any these persons) born to U.S. citizens who were living abroad?

YES :____:
NO :____:

INTERVIEWER: IF THIS PERSON HAS DIFFICULTY UNDERSTANDING ENGLISH, PLEASE TERMINATE. THE INTERVIEWS MUST BE CONDUCTED IN ENGLISH.

10. [IF MORE THAN ONE HH MEMBER] Can you tell me how the persons who live here are related to one another? (for example, as unrelated roommates, or husband, wife, and children, and so)

CLASSIFY HOUSEHOLD INTO ONE OF THE FOLLOWING CATEGORIES:

CATEGORY 1 (Recruit 17 persons)

- two or more unrelated roommates;
- at least one person unrelated to others who may be related or married

CATEGORY 2 (Recruit 17 persons)

NOTE: Most of these should include some persons born in U.S. and others born elsewhere.

- parents, children, and grandparents,
- parents, children, and other extended family members such as cousins, uncles, etc.

CATEGORY 3 (Recruit 10 persons)

- single person living alone, [NO MORE THAN 3 OF THESE TOTAL]
- married couple;
- married couple with children,
- single parent with children,
- siblings or other relatives living together.

11. IF NECESSARY: Just let me confirm please, at least one person currently living with you was born outside the United States – is that correct?

YES :__:
NO :__:[TERMINATE]

***DETERMINE WHETHER RESPONDENT IS NEEDED IN THIS STUDY.....**

IF TERMINATED, TELL THEM WE MAY BE ABLE TO USE THEM IN ANOTHER STUDY WE ARE CONDUCTING FOR THE CENSUS BUEAU IN THE NEXT FEW WEEKS.

APPOINTMENT

Thank you for answering my questions. I'd like to schedule an appointment for an interview at a time that's convenient for you. Let me read you some times I have available and you can choose the time that is best for you:

[SCHEDULE DAY AND TIME]

May I please have your name, address and phone number? (We need your address so that we can send you directions on how to get [here/to the hotel and find the room.]

Name: _____

Address: _____ **ONLY ONE PER HOUSEHOLD**

Phone number: _____

I will send the directions out to you shortly. It will include instructions on where to park. If you have to cancel your interview, please call back so that we can schedule someone in your place, OK?

Appendix C
Cognitive Interview Guide
(mail version)

Date: _____ Time: _____ Interviewer: _____ Respondent ID: _____

**ACS Cognitive Interviews - Migration
Self-administered (Mail) Version**

Thank you for agreeing to help us out. Let me tell you what this is about.

I work for Westat (in Rockville, MD), which is a research company that conducts research on many different topics, under contract for many different organizations. This particular project is for the U.S. Census Bureau. One of the things the Census Bureau does is count the number of people in the country's population, as well determine how many people are in different age groups, how many were born inside and outside the U.S., and so on. They do this in a variety of ways - mainly by mailing a questionnaire to households, where someone is asked to fill it out and send it back to the Census Bureau.

Today we are helping the Census Bureau improve the questions that are in one of the surveys they use for this purpose, called the American Community Survey (It's something they have recently started doing on an annual basis to update the information collected in the census done at the start of the decade). For the information gathered in the survey to be useful, it's important that the questions are ones that people can easily understand, and provide meaningful answers for. By reading or hearing the questions and sharing your reactions with us, it will help the Census Bureau make sure their questions are good ones. That's what we've asked you to help us out with.

Before we get into it, I should make you aware of a few things. First, I want to assure you that everything we cover in the interview will be kept confidential - Only people actually working on the project will have access to the information you share with us. But if it's OK with you, we would like to audio-tape our conversation, just so that we can review it later - it helps us analyze our results. (And finally, some of our friends from the Census Bureau that we're working with are here today - behind the mirror - to observe and see for themselves how the questions work.)

CONSENT FORM: *Here is a form I must ask you to look over and sign - it basically covers the points I've just gone over with you and indicates you have agreed to take part. (ONE COPY FOR RESPONDENT, ONE COPY FOR US)*

So the way this will work is: I'm going to ask you to fill out parts of a test copy of the Census Bureau survey I was telling you about. You can sort of pretend you got this in the mail and you've sat down at home to fill it out. As much as possible, just to read and respond to it the way you normally would. Also, because we're testing the questions, I'd appreciate it if you could read aloud as you go along - that will help me keep track of where you are and what exactly you're reading.

Also, because we're testing the questions, I'd appreciate it if you could try and "think aloud" while you answer the questions - as much as possible, just verbalize whatever comes to mind while you think about the question and come up with your answer. I'm going to give you an example of what I mean by this, and then let you try it. Suppose one of the questions is: "How many windows are in your house or apartment?" In coming up with their answer, someone may think aloud by going:

“Well, there are 2 windows in the living room, one in the kitchen, one in the bathroom, and we have three bedrooms with one window in each. So that’s a total of 7 windows – Oh, but then our basement has two very tiny windows – am I supposed to count those? If I counted those, it would a total of 9 windows.”

Why don’t you try it now: How many windows are in your house or apartment?

And if you’re not sure what a question is asking, by all means let us know that. Occasionally, after you’ve answered a question, I might ask you to tell me a little bit about why you answered the way you did, or to tell me what something means to you. And once we’ve gone through the questions we’ll review some of them more in-depth. Okay?

Probing Strategy:

- A ‘think aloud’ approach will be used including a practice think aloud exercise after the respondent signs the consent form
- Have respondent fill out a copy of the mail survey for up to five household members. Have them read the question and their answer as they do so. Fill in their answers and note any observations on your own copy of the mail survey.
- Majority of probing will be retrospective - done AFTER all ACS items have been asked for household members. Repeat the question (“One of the questions I asked you was.....”) along with the respondent’s answer. Then probe. As necessary, tailor probes to the respondent, and the person or situation for which the respondent is reporting. Also, you may need to probe the same question/issue for more than one person in the household (e.g., people came different countries, or emigrated at different times, etc.)
- While completing the form/interview, probe IMMEDIATELY on obvious ‘issues’ – a noticeable problem such as respondent confusion, changing of answers, response is ‘don’t know’ or the cognitive interviewer needs some clarification on something the respondent said as part of the think aloud response. Example generic probes are:
 - *I noticed you hesitated before answering – can you tell me what you were thinking about?*
 - *You seem a bit confused – can you tell me about that?*
- Follow-up with more specific probes (e.g., those regarding specific phrases), as appropriate.
- Before dismissing respondent, ask if they have any additional comments on the questions asked in the interview.

7. Where was this person born?

- Inside the United States—*Print name of state.*

- Outside the United States—*Print name of foreign country, or Puerto Rico, Guam, etc.*

Print name of specific state/province/region of foreign country.

PROBES

What does the phrase “state/province/region” mean to you?

8. Is this person a citizen of the United States?

- Yes, born in the United States—*Skip to 10a*
- Yes, born in Puerto Rico, Guam, the U.S. Virgin Islands, or Northern Marianas
- Yes, born abroad of a U.S. citizen parent or parents
- Yes, U.S. citizen by naturalization—*Print year of naturalization*

- No, not a citizen of the United States

PROBES:

What do you think is meant by “Born abroad of a U.S. citizen parent”?

What do you think is meant by “naturalization” (or “U.S. citizen by naturalization”)?

IF PROVIDED A YEAR OF NATURALIZATION:

How did you figure out XXXX as the year (you/name) naturalized? Is that date connected with any event in the naturalization process? (When did the process of naturalization end?)

9. When did this person come to live, work, or study in the United States?

a. Print year of first arrival.→

b. If this person has entered the U.S. more than once, print year of most recent arrival.→

PROBES:

In your own words, what is this question asking?

What does the year XXXX represent in your answer for “year of first arrival”?

*What do you think is meant by part b – “If this person has entered the U.S. more than once...”?
[IF R HAS MISINTERPRETS THIS QUESTION, SHOW THEM THE ALTERNATIVE ON THE NEXT PAGE – ASK IF IT CHANGES WHAT THEY THINK THE QUESTION MEANS AND HOW.]*

What do you think “year of most recent arrival” means in this question?

IF ENTERED A YEAR FOR ITEM B:

What does the year XXXX represent in your answer for “year of most recent arrival”?

*How long were you outside of the United States before you entered this last/most recent time?
How many times have you left the US and then returned again? [PROBE FOR DETAILS OF WHEN/HOW LONG*

[ALTERNATE WORDINGS OF QUESTION 9B]

9. When did this person come to live, work, or study in the United States?

a. Print year of first arrival.→

b. If this person has entered the U.S. more than once to live, work, or study, print year of most recent arrival.→

9. When did this person come to live, work, or study in the United States?

a. Print year of first arrival.→

b. If this person has entered the U.S. more than once, (*not including holidays, vacations, or other brief visits*) print year of most recent arrival.→

SCENARIOS (Mail Version) [End of Interview]

Let me ask you to imagine some other situations, and you can tell me how you think you would answer certain questions in the survey for these people.....

SHOWCARD A

- How would you answer question 8 for Sofia?: (What year would you write in the box?)

Showcard A

Suppose Sofia moved to the U.S. to live in 1986. She submitted papers for becoming a U.S. citizen in 1992. The U.S. government told her in 1994 that she is now a U.S. citizen.

SHOWCARD B

- How would you answer question 9 for Robert? (*When did this person come to live, work, or study in the United States?*)

Showcard B

Suppose Robert came to the U.S. to attend college for about 3 months (one semester) back in 1992, and then returned to his home country. Then in 1999, Robert moved to the U.S. to work at a job, and has been living here since that time.

SHOWCARD C

- How would you answer question 9 for Maria? (*When did this person come to live, work, or study in the United States?*)

Showcard C

Suppose Maria moved to the U.S. to live in 1997. Since that time, she has visited family in her home country twice, in 1999 and 2003. Each visit was for two weeks.

Appendix D

Westat Recommendations for the ACS Migration Questions

--Mail Mode--

7. Where was this person born?

- Inside the United States—*Print name of state.*
- _____
Outside the United States—*Print name of foreign country, or Puerto Rico, Guam, etc.*

Specify city or town of foreign country.

8. Is this person a citizen of the United States?

- Yes, born in the United States—*Skip to 10a*
- Yes, born in Puerto Rico, Guam, the U.S. Virgin Islands, or Northern Marianas
- Yes, born abroad of at least one U.S. citizen parent
- Yes, U.S. citizen by naturalization—*Print year of naturalization*
- _____
No, not a citizen of the United States

9. When did this person come to live, work, or study in the United States?

a. Print year→

b. If this person has come to the U.S. more than once (not including holidays, vacations, or other brief visits), print year of most recent arrival.→

--Telephone Mode--

9a. (Was <Name>/Were you) born in the United States?

- Yes, born in the United States
- No (If No, skip to 9c)

If DK or Ref, skip to Question 10a

9b. In what state (was <Name> / were you) born?

Enter name of state, then skip to 12

9c. In what country (was <Name>/were you) born?

Enter country.

9d. In what city or town of <Country Name> (was <Name>/were you) born?

Enter state/province/region.

10a. (Is <Name>/Are you) a citizen of the United States?

Yes - if R REPORTED A COUNTRY WHERE BORN,, go to 10b, otherwise go to 10c
No - go to 11a

DK and Refuse go to 13

10b. (Was <Name>/Were you) born abroad of at least one U.S. citizen parent, or did (he/she/you) become a U.S. citizen by naturalization?

Born abroad of a U.S. citizen parent(s) - go to 11a
U.S. citizen by naturalization - go to 10d
Neither - go to 11a

--Telephone Mode—
(continued)

Ask if DK or ref place of birth, and 10a=yes (citizen)

10c. You have indicated that (<Name> is/you are) a citizen, and that you don't know (his/her/your) exact place of birth, but perhaps you could give us other general information about (him/her/yourself)...

(Was <Name>/Were you) born in the U.S., born in Puerto Rico, Guam, U.S. Virgin Islands or Northern Marianas, born abroad of at least one U.S. citizen parent, or did (he/she/you) become a naturalized citizen?

1. Born in the U.S.
2. Born in Puerto Rico, Guam, U.S. Virgin Islands, or Northern Marianas - go to 11a
3. Born abroad of a U.S. citizen parent(s) - go to 11a
4. Citizen by naturalization - go to 10d
5. None of the above - go to 12

10d. In what year did (<Name>/you) become a naturalized U.S. citizen?

Enter year of naturalization

Go to 11a

11a. What year did (<NAME>/you) FIRST come to live, work, or study in the United States?

Enter year of first entry

11b. (Have you / Has NAME) come to the United States, more than once, not including holidays, vacations, or other brief visits?

Yes, continue to 11c

No, DK or REF, skip to 12

11c. In what year did (you / NAME) MOST RECENTLY come to the United States?

Enter year of most recent entry