

SUMMARY RESULTS OF INDIVIDUAL EVALUATIONS AND EXPERIMENTS

Category A - Response Rates and Behavior Analysis

Summaries of the twelve evaluations in the Response Rates and Behavior Analysis category follow.

Evaluation A.1.a, Telephone Questionnaire Assistance (Chesnut 2003a)

The Telephone Questionnaire Assistance program (TQA) was implemented to assist the public in completing their census forms. Six language specific national toll-free numbers were printed on Census questionnaires and Language Assistance Guides. The English and Spanish toll-free numbers connected to an Interactive Voice Response (IVR) system where a caller obtained information by selecting from a series of menu options, and if needed, was transferred to an agent. The toll-free numbers for the Asian languages (Chinese, Korean, Vietnamese, and Tagalog) connected directly to bilingual agents. This evaluation profiled the Census 2000 TQA program and, where appropriate, assessed the performance of the TQA system.

The TQA program received 6 million calls, with approximately 51 percent serviced by an agent. There were three peaks for the daily total call volume. The first peak occurred after the initial mailout of census questionnaires, the second peak occurred after the mailout of the reminder postcard, and a third smaller peak occurred the week of Census Day. Each of the peaks occurred on Mondays. Excluding the peak days, a trend was observed in which Monday was the highest call volume day with a gradual decline in call volume throughout the week ending with a low call volume on Sunday.

The IVR system allowed callers to obtain or enter information by selecting from a series of menu options. The Census Bureau found that callers primarily used the IVR system to obtain information on completing a census form. Second, they used the IVR system to request a census form by mail and third, callers used the IVR system to obtain general information and other information pertaining to the census. Of the main services provided by an agent, the most frequently requested service was the request for a census form. Callers also requested answers to frequently asked questions about the census, requested answers concerning specific items on the census questionnaire, or registered a complaint about the census.

Based on the results of this evaluation, the following recommendations were made:

- Predict the call volumes on peak call volume days based on Census 2000 experience.
- Monitor the performance of the network provider based on their contractual requirements.

- Deliver the census mailing pieces on a day other than Monday to avoid an additive effect due to mailing strategy and day of the week.
- Conduct future research in assessing the expanded use of IVR technologies.
- Provide on-site technical support to all call centers
- Provide equal levels of service in both English and Spanish systems.

Evaluation A.1.b, Telephone Questionnaire Assistance Customer Satisfaction Survey
(Stevens 2002)

The purpose of this evaluation was to measure how satisfied callers were with the TQA operation for Census 2000. The Census Bureau measured customer satisfaction from two perspectives: the caller and the agent who handled the call.

Overall, callers were satisfied with the TQA operation. At least 72 percent of the respondents to the customer satisfaction survey replied favorably. The survey questions asked about ease of moving through the automated menu system, quickness of the agent in understanding their request, agent's level of interest in helping, overall satisfaction with the call, and other customer concerns.

Overall, agents widely supported most aspects of the TQA operation mentioned on the agent debriefing questionnaire. Only three out of nineteen questions were viewed negatively by the majority of the agents. The agents' satisfaction with the operation supports the callers' satisfaction. The agents agreed that: they understood the caller's requests, the visual design of the Operator Support System made it easy to read the prepared answers, training helped them understand Census concepts, and it was easy for them to use the Operator Support System to find the information that callers requested. However, the agents felt the callers seemed dissatisfied when they repeated the same verbatim information and they felt they could have used more practice with the Operator Support System before fielding calls.

Based on the results, recommendations are:

- Continue to research the caller's expectation at the first and subsequent menu selections in the IVR system, as part of or prior to development.
- Research the caller's expectations of waiting times and make adjustments accordingly to the maximum time agents should keep callers waiting.
- Design the Operator Support System script so that less information is repeated when the agents are responding to an incoming question. In addition, increase training on how the agents can read the Operator Support System script appropriately.

- Provide the agents with extra practice time and include more realistic examples of different types of calls.
- Allow agents to respond to requests for replacement forms as soon as the reminder postcards are delivered to mailout/mailback addresses.
- Provide the agents and/or the IVR system with tools for verifying whether the Census Bureau received a caller's census form, being sure to address confidentiality issues.

Evaluation A.2.b, Internet Data Collection (Whitworth 2002)

Census 2000 was the first U.S. census to include an Internet on-line reporting capability. Though it was met with many challenges, the Internet data collection (IDC) was an operational success. It proved to be secure and there were no hardware or software failures and no known security breaches. However, it did not fully reach its potential. This is probably because of a conscious decision by the Census Bureau not to advertise this response mode. The IDC system could have handled tens of millions of forms. Instead, the Census Bureau received 89,123 initial on-line requests for an Internet census form. About 16.7 percent of these initial requests were invalid (mostly requests for the long form). There were 63,053 households representing 169,257 persons that were counted using just the Internet census form. Fewer than 4 percent of Internet submissions had other returns from different modes.

The Internet form collected data for six persons, with a continuation roster similar to the short-form paper questionnaire. This restriction of a six-person form required followup for large households, but did not greatly affect Coverage Edit Followup (CEFU) operations. Fewer than 2 percent of the Internet households reported household size to be larger than six. Internet households had approximately the same reported household size as reported for all response modes. There were some differences between the demographic makeup for the Internet household members when compared to the overall population. This is not surprising when one considers the typical Internet user. Demographic highlights include:

- There were slightly more males (52.0 percent) than females (48.0 percent).
- There were higher percentages of White and non-Hispanic household members than for all response modes.
- There was a higher percentage of persons between the ages of 25-54 responding on the Internet, as compared to all response modes.

The Internet is here to stay. The exact form and function of Census Internet options, however, is largely undeveloped. Recommendations include:

- Focus future research not only on how to securely implement the form itself, but also on how to promulgate the Internet form as a major response option.

- Focus future research on how to use Internet response as a tool to increase data quality by implementing real-time data feedback and analysis.

Evaluation A.2.c, Census 2000 Internet Web Site and Questionnaire Customer Satisfaction Surveys (Stapleton 2002)

As part of a comprehensive plan to simplify public participation and to increase response rates to Census 2000, the Census Bureau designed a single web site to service Internet users. The site contained two major components: Internet Questionnaire Assistance (IQA) and IDC. Customer satisfaction surveys were conducted and the survey results were used to analyze the degree of respondent satisfaction with each system.

Most respondents were not satisfied with the IQA. Nearly 62 percent of the respondents indicated that, overall, they were not at all satisfied with the Internet help screens. While nearly 77 percent of the respondents found it easy or very easy to understand the help screen information, about 58 percent said it was not at all easy to find the help topics for which they were searching and 65 percent stated that the help screen information was not at all helpful. These findings suggest that while the information presented on the site was easy to interpret, it may not have been the appropriate information for the users. Those respondents who did find the information helpful were more satisfied overall. While the information on IQA was easy to understand, it was difficult to locate and generally unhelpful. In short, the IQA did not provide the information that respondents were seeking. However, the high correlation between helpfulness and overall satisfaction indicates that the Census Bureau might improve customer satisfaction by focusing on IQA elements that are helpful to users.

The Census Bureau measured satisfaction on seven aspects of the Census 2000 Internet Form: time required to load the form, moving through the form, availability of help screens, understanding the help screen information, ease of sending the form, security and confidentiality procedures, and overall satisfaction. Respondents were largely satisfied with most of the seven aspects. The percent of respondents indicating they were satisfied or very satisfied with a specific aspect was as high as 94 percent (for the item 'ease of sending form'). However, satisfaction lapsed slightly for the two items which dealt with help screens: availability of help screens and understanding the help information (74 percent and 73 percent, respectively). Overall, 91 percent of respondents were satisfied with the Census 2000 Internet Form. High levels of customer satisfaction indicate a strong potential for large-scale implementation in 2010.

Key recommendations include:

- Implement a content redesign of the Internet census help instrument.
- Conduct research on knowledge and perceptions of the decennial census as well as the needs of potential users of the Census 2000 web site.
- Look beyond restricting on-line assistance to questionnaire help.

Evaluation A.3, Be Counted Campaign for Census 2000 (Carter 2002a)

The four goals for the Be Counted Campaign in Census 2000 were to count persons who did not receive a census questionnaire, to count persons who believed they were not included on any other census form, to encourage participation of persons who are traditionally undercounted in the census, and to provide a means for persons with no usual residence to be counted.

Respondents returned 804,939 Be Counted Forms to the Census Bureau. The Census Bureau expected approximately 1 million Be Counted Forms to be returned. Of the 605,905 Be Counted Forms that were included in census processing, 239,128 Be Counted Forms added persons to the census not included on other Census forms. There were 236,482 households where the household contained some persons who were only enumerated from the Be Counted Form return. Of these households, 116,019 were enumerated only by Be Counted Forms and the remaining 120,463 were enumerated by Be Counted Forms as well as other census forms.

There were 560,880 persons added to the census through Be Counted Forms. This is more than double the number of persons added from the “Were You Counted?” program in the 1990 Census. There were higher percentages of groups traditionally undercounted than were observed in the census overall. These groups include renters, children, and minorities. Approximately 40.7 percent of all Be Counted Forms that were picked up by respondents from distribution sites were non-English forms, most of which were Spanish. There were also approximately 15,410 Be Counted Forms that were returned to the Census Bureau that were determined to be from persons with no usual residence. This means that Be Counted Forms increased coverage in groups that have been traditionally hard to count.

Overall the Be Counted Campaign was a success. It added 560,880 persons to the census. While this number is small, these are people that would have been missed without this program.

The Census Bureau should consider the following points if implementing an operation like the Be Counted Campaign in 2010:

- Conduct further analysis to investigate the number of Be Counted Forms matched and geocoded by the automated system and by clerical staff.
- Design the Field Verification to permit the enumerator to record the Census identification number (ID) of the Be Counted Form duplicates.
- Record the language of the returned Be Counted Forms and whether translation or transcription was needed.

Evaluation A.4, Use of Non-English Questionnaires and Guides in the Census 2000 Language Program (Smith and Jones 2003)

For Census 2000, households in mailback areas were mailed an advance letter. The advance letter provided households an opportunity to request an alternative language questionnaire in one of five languages. Short and long form questionnaires were available upon request in Spanish, Chinese, Tagalog, Vietnamese, and Korean. Respondents were asked to return their advance letter indicating which language questionnaire they preferred. This approach spearheaded the effort to encourage respondents in linguistically isolated households to complete a census questionnaire. The Census 2000 Language Program also made available language guides in forty-nine languages. The language guides were user-friendly visual aides that assisted respondents in completing the questionnaires for both long and short forms.

This evaluation analyzed information about non-English forms requested by households who were mailed advance letters. The results are compared to the number of requests for Language Assistance Guides.

There were over 2.2 million requests for non-English questionnaires. Most (83.7 percent) of the households that requested alternative language questionnaires requested them in Spanish. Although there were requests for non-English questionnaires from households in all states, most (about 70.6 percent) of the requests were from households in four states: California, Texas, New York, and Florida. About 39.4 percent of the households that requested alternative language questionnaires were in census tracts designated as hard-to-enumerate. Less than half (45.1 percent) of the households requesting an alternate language form returned these forms by mail.

The number of respondents requesting language assistance guides was substantially lower than the number of households requesting a non-English language questionnaire. At least 93,672 respondents requested language assistance guides from Questionnaire Assistance Centers and 77,191 respondents requested language assistance guides through the TQA program. About 34.3 percent of the language assistance guides requested at Questionnaire Assistance Centers and 18.9 percent of the language assistance guides requested through the TQA program were in languages other than Spanish, Chinese, Korean, Tagalog, or Vietnamese.

Based on the results of this evaluation, the recommendation is:

- Continue to take initiatives to help respondents overcome language barriers in completing census forms by providing an opportunity for households to make requests for non-English questionnaires, providing guides and questionnaires at assistance centers, and providing alternate language telephone assistance.

Evaluation A.5.a, Analysis of the Linguistically Isolated Population in Census 2000 (Lestina 2003)

The four objectives of this evaluation were to create a universe of households identified as linguistically isolated, to determine how they were enumerated in Census 2000, to examine the education attainment of the householder, and to examine geographic clustering at the tract and county levels.

A household is classified as linguistically isolated if all household members age 14 years or older speak a language other than English and have limited English proficiency. There are 3,141 counties in the nation with at least one linguistically isolated household. Each of these has up to 35.0 percent of its households that are linguistically isolated. Of the 3,141 counties, 91.53 percent have fewer than 5 percent of their households that are linguistically isolated. There are eight counties in Texas with at least 25 percent of their households that are linguistically isolated.

Of the 64,960 tracts in the nation with at least one linguistically isolated household, 77.5 percent have less than 5 percent of their households that are linguistically isolated. There are 11 tracts in the nation where at least 75 percent of their households are linguistically isolated. They are as follow: one in Maricopa County, Arizona, one in Pinal County, Arizona, three in Los Angeles County, California, one in San Francisco County, California, one in Lafourche Parish, Louisiana, two in Bronx County, New York, one in Dutchess County, New York, and one in Charleston County, South Carolina.

At the tract and county levels, the linguistic isolation variable may help with identifying areas for special enumeration procedures, including language programs, for the 2010 Census. Further analysis should be done by specific languages that are spoken at home to identify the level and whether they are clustered.

Evaluation A.6.a, The U.S. Postal Service Undeliverable Rates for Census 2000 Mailout Questionnaires (Kohn 2003)

In Census 2000, the questionnaire mailout/mailback enumeration method was the primary means of census-taking. The U.S. Postal Service (USPS) was the primary vehicle for delivering census questionnaires. Based on the Decennial Master Address File (DMAF), the Census Bureau mailed questionnaires on March 13-15, 2000 to about 96 million housing units in areas designated as being mailout/mailback. Questionnaires that were undeliverable were called Undeliverable as Addressed (UAA). Since this study of questionnaires concerns undeliverability by the USPS, only mailout/mailback housing units are included.

The Census 2000 Local Census Office (LCO) redelivery operation for UAA questionnaires took place in preselected ZIP codes and was conducted by specially trained enumerators. The UAA questionnaires were routed back through the USPS and returned for check-in at the LCO until March 18, 2000. By redelivering questionnaires identified as UAA in areas where they were clustered, the Census Bureau sought to efficiently boost response by getting questionnaires back

into the hands of the households early in the mail response period and to address geographic clustering of UAA questionnaires.

Addresses remaining UAA after the census redelivery were included in the nonresponse followup (NRFU) workload. During NRFU, many housing units for UAA questionnaires were enumerated as occupied households. Housing units which were delivered a questionnaire either by the USPS or the census redelivery could have returned their questionnaires by mail. Those which did not respond by mail were also included in NRFU and many of these were also enumerated as occupied housing units.

The major objectives of this evaluation were to examine the decrease in undeliverable housing units as a result of the redelivery operation and to study relationships between UAA status and demographic data.

Nationwide, the Census Bureau delivered to nearly 600,000 occupied housing units in the redelivery operation. Age of the householder, tenure, and the size of the household are the best predictors of USPS delivery. Minority and nonminority counts were obtained with the aid of a race edit/allocation variable which allocated each person to one and only one of six major race categories. Minority households were more likely to be in LCOs selected for the redelivery operation than nonminority households. For USPS UAA units for which redelivery was attempted, nonminority households were more likely to have a successful redelivery than minority households.

Evaluation A.6.b, Study of the U.S. Postal Service Reasons for Undeliverability of Census 2000 Mailout Questionnaires (Chesnut 2003b)

During the mailout of Census 2000 questionnaires, the USPS designated each questionnaire as UAA if it could not be delivered successfully to the labeled address. The mailing pieces were annotated with a reason for undeliverability and sent back to the Census Bureau. From the undeliverable questionnaires received, a stratified systematic sample was drawn for the purpose of conducting a study of the reasons for undeliverability. This evaluation examined the distribution of reasons for undeliverability. In addition, inferences are drawn from the sample to the universe of Census 2000 undeliverable questionnaires at a national level. This evaluation also investigated whether the reason for undeliverability is a valid indicator of the final census status of a housing unit.

From the results of the sampling procedure, a total of 9.7 million UAA questionnaires were received at the National Processing Center in Jeffersonville, Indiana. The most common reason questionnaires were not deliverable was due to the fact that the housing unit was identified as vacant by the USPS. Almost half of the undeliverable questionnaires received were stamped or annotated with a “vacant” reason for undeliverability. The USPS policy is that mail is not delivered to vacant units. Addresses identified as “no such address” and those identified as not having a mail receptacle composed the next largest portions of undeliverable questionnaires. The remaining undeliverability categories (duplicate, under construction, demolished, nonresidential, no such apartment, post office (P.O.) box, not deliverable and unable to forward,

outside delivery limits, refused, blank other, and illegible) each contributed 6 percent or less to the universe of undeliverable questionnaires.

Thirty-one percent of the total undeliverable questionnaires received a final census status of vacant and 47 percent received a final census status of delete. Approximately 22 percent were given a final census status of occupied. For the questionnaires marked “vacant,” approximately 50 percent actually received a final census status of vacant. Also, 22 percent of the questionnaires marked vacant received a final status of occupied. For the questionnaires marked “post office box” or “no mail receptacle,” an estimated 48 percent were occupied. This highlights the troublesome aspect of using mailout/mailback to enumerate respondents who do not receive mail at their place of residence. In a much more favorable result, an estimated 85 percent of questionnaires marked “demolished,” “new construction,” or “nonresidential” received a final status of delete and that 77 percent of questionnaires marked “no such address” or “no such apartment” received a final status of delete.

Based on the evaluation results, recommendations include:

- Use USPS products/services such as the Address Element Correction service prior to Census mailout to provide corrections to addresses or to identify potentially undeliverable addresses in the mailout/mailback address list.
- Capture the USPS’s reasons for undeliverability and use these in determining final census status.

Evaluation A.7.a, Census 2000 Mail Response Rates (Stackhouse and Brady 2003a)

The response rate is a measure that represents the percentage of addresses eligible for NRFU that returned questionnaires prior to the designation of the NRFU universe. Due to the expected higher level of data quality and the lower cost associated with self-enumerated responses relative to enumerator-collected responses, it is important for response rates to be as high as possible.

The mail response rate is defined as the number of mail returns received prior to the cut date for the NRFU universe divided by the total number of housing units in mailback areas that were eligible for NRFU. The final response rate is similar but includes all mail returns through the end of the year. Mail returns included in the response rates are paper questionnaires, interviews during the TQA program, Internet data captures, Be Counted Forms, and CEFU returns.

The mail response rate as of April 18, 2000 was 64.3 percent. The final response rate was 67.4 percent as of December 31, 2000. Reflecting the higher response burden of the long form questionnaire, the short form mail response rate of 66.4 percent was 12.5 percentage points higher than the long form mail response rate of 53.9 percent. Many residents with long forms held onto them and returned them after April 18. After that date a larger proportion of long forms were returned than short forms. The final response rate was 69.1 percent for short forms and 59.4 percent for long forms.

Mailout/Mailback areas had a mail response rate of 65.4 percent, which is higher than either the Update/Leave (U/L) areas mail response rate of 59.3 percent or the Urban U/L areas mail response rate of 50.5 percent. Final response rates by Type of Enumeration Area (TEA) were 68.5 percent for Mailout/Mailback, 62.6 percent for U/L, and 54.8 percent for Urban U/L.

Most questionnaires were returned in the period between March 15, when questionnaires in Mailout/Mailback areas were mailed, and March 28. There were slight surges in the number of mail returns corresponding to the delivery of reminder postcards beginning on March 20 and on Census Day (April 1). These two surges in response were more pronounced for long forms than short forms.

An additional 1,052,712 returns were received between April 18 and April 25, representing 28.4 percent of the mail returns checked in after April 18. These returns represent a potential decrease in the NRFU workload of 2.5 percent. Therefore, work needs to be done to determine what is the optimal date for determining the NRFU universe, by considering the cost benefits versus the operational challenges to other operations.

Evaluation A.7.b, Census 2000 Mail Return Rates (Stackhouse and Brady 2003b)

The mail return rate is a measure of respondent cooperation in Census 2000. It is defined as the number of mail returns received prior to the cut date for the NRFU universe divided by the total number of occupied housing units in mailback areas that were on the DMAF prior to NRFU. The final return rate is similar but includes all mail returns through the end of the year. Mail returns included in the return rates include actual paper questionnaires, interviews during the TQA program, Internet data captures, Be Counted Forms, and CEFU returns.

The mail return rate is different from the mail response rate. The denominator of the mail response rate includes all housing units in mailback TEAs that were eligible for NRFU and had addresses that were considered adequate to attempt delivery by either the USPS or census field staff. The response rate denominator is larger than the return rate denominator, largely because the response rate denominator includes vacant housing units, UAA addresses, some addresses deleted in U/L and Urban U/L delivery, NRFU, or Coverage Improvement Followup (CIFU).

The mail return rate as of April 18, 2000 was 74.1 percent. This rate represents 75,163,020 mail returns that were received by April 18, 2000 out of a return rate denominator of 101,398,131 households. Another 4,367,080 questionnaires were returned after April 18, resulting in a final return rate as of December 31, 2000 of 78.4 percent. Reflecting the higher response burden of the long form questionnaire, the short form mail return rate of 76.4 percent was 13.4 percentage points higher than the long form mail return rate of 63.0 percent. Many residents with long forms held onto them and returned them after April 18. After that date a larger proportion of long forms were returned than short forms. The final return rate was 80.1 percent for short forms and 70.5 percent for long forms.

Mailout/Mailback areas had a mail return rate of 75.1 percent, which is higher than either the mail return rate for U/L areas (69.6 percent) or the mail return rate for Urban U/L areas (63.7

percent). Final return rates by TEA were 78.6 percent for Mailout/Mailback, 77.9 percent for U/L, and 70.8 percent for Urban U/L.

Differential return rates were observed for different demographic groups. The likelihood of responding to the census increased with householder's age. Whites had a higher mail return rate (77.5 percent) than the total mail return rate, while all other race groups had lower return rates than the total mail return rate. Non-Hispanic householders had a mail return rate of 75.0 percent, 10.5 percentage points higher than the Hispanic mail return rate of 64.5 percent. Households consisting of two persons had the largest proportion of residents who responded to the census. Larger households of five persons or more had increasingly lower mail return rates as household size increased.

Evaluation A.8, Puerto Rico Focus Groups on Why Households Did Not Mail Back the Census 2000 Questionnaire (Berkowitz 2001b)

Census 2000 was the first time that residents of Puerto Rico were asked to complete and return their questionnaires by mail. Fifty-three percent of the households in Puerto Rico returned their questionnaires by mail, a low response rate compared with the national rate of 65 percent. This evaluation explored the reasons residents of Puerto Rico didn't mail back their questionnaires. Contractors conducted focus groups in nine sites. Sites were selected for geographic and socioeconomic diversity from among municipalities (municipios) with lower-than-average (under 50 percent) mailback response rates. Although interesting and suggestive, these results are based on a small, purposive sample and cannot be generalized to the population of Puerto Rico as a whole. However, they do provide a useful jumping off point for further thinking and research.

Participants' reasons for not returning their questionnaires by mail fall into four clusters:

- Motivational and process-related reasons. Several participants were unclear about or misunderstood the census' purpose. Considerable confusion existed over the process by which the questionnaires were distributed, the rules for returning them, and especially the role of enumerators.
- Practical and logistical reasons. Lack of time figured as a reason, as did difficulties of mailing and getting to the post office.
- Cultural and political attitudinal reasons. These included fears that the information wouldn't be kept confidential, as well as the belief that any funds that would be allocated would only end up enriching dishonest politicians. Participants also expressed a strong preference for a more personal approach to collecting the information as being more appropriate to the culture.
- Reasons related to questionnaire content and design. Some participants were affronted by the race and ethnicity questions, which were seen as divisive as well as inappropriate to the realities of Puerto Rico.

Almost everyone had heard something about Census 2000 from television and radio ads, newspapers, schools, or informal sources such as relatives and neighbors. But in many cases, advertising efforts fell flat. Nearly all of the 41 nonresponding household heads urged a return to the system of collecting data door-to-door as practiced in 1990.

Key recommendations for improvements and for further research and exploration include:

- Provide more comprehensive information, in different forms, on the purpose and uses of the census.
- Conduct further research in Puerto Rico on views and perceptions of the Census 2000 questions on race and ethnicity.
- Use the findings and recommendations from this study, as well as other supporting testing and research, to develop advertising appeals for census participation more attuned to Puerto Rico.
- Consider using the results of this study to create a close-ended survey to be administered to a probability sample of residents of Puerto Rico.

Category B - Content and Data Quality

Summaries of the eight evaluations in the Content and Data Quality category follow.

Evaluation B.1.a, Analysis of Imputation Rates for the 100 Percent Person and Housing Unit Data Items from Census 2000 (Zajac 2003)

This evaluation provides information on data quality, specifically data completeness, for the 100 percent person and housing unit items. For this evaluation, data completeness is measured by imputation. It should be noted that the definition of imputation can be interpreted in various ways which could lead to different methods of computing imputation rates. Therefore, when comparing imputation rates across reports, it is imperative to understand the way the rates are computed to ensure they are comparable.

Imputation is divided into three categories. They are defined as follows:

- An *assignment* is performed when a response for a data item is either missing or not consistent with other responses and an item value can be determined based on information provided for that same person.
- *Allocations* are performed when a response for a data item is either missing or is not consistent with other responses and an item value cannot be determined based on information provided for that same person. An allocation uses a response from another person within the household or from a person in a nearby household. When every 100

percent characteristic for a person requires allocation, the person is considered totally allocated when at least one person within the household has data that do not require allocation. However, when every item for every person in the household requires allocation, it is covered by substitution.

- A *substitution* occurs when all the 100 percent characteristics for every person in the household are either missing or are not consistent with other responses. A nearby housing unit with complete 100 percent data is selected to represent the missing or inconsistent data items. This nearby housing unit is selected using the nearest neighbor hot deck. This is also called a whole household substitution.

In addition to these three types of imputation rates, a data completeness statistic was produced to determine the number of 100 percent population items within each person record that were not imputed.

Almost 1.5 million households were substituted, representing 1.39 percent of the occupied housing units. Within these substituted households, there were over 3.4 million substituted persons, accounting for 1.26 percent of the persons in housing units.

Total item imputation rates for the 100 percent person data items range from 1.98 percent for the sex item to 5.08 percent for the age item. The tenure item imputation rate was 5.48 percent. In general, short form data were more complete than long form data, self-response data were more complete than enumerator return data, English forms had more complete data than forms designed for other languages. For all items, data for owners were more complete than for renters.

Overall, the data completeness statistic shows that about 97 percent of nonsubstituted person records have at least four of the five 100 percent population items with nonimputed data.

Evaluation B.1.b, Analysis of Item Nonresponse Rates for the 100 Percent Housing and Population Items from Census 2000 (Norris 2003)

This evaluation determined the extent of item nonresponse for the Census 2000 hundred percent items. Inconsistent responses are not considered nonresponse. Rates were reported for each of the 100 percent household population items and tenure classified by form type (long versus short) and response mode (self versus enumerator). In addition, results of the Internet returns are reported. Some of the breakdowns within return characteristics are subject to interpretation.

The definition of item nonresponse is sometimes interpreted in various ways depending on the scope of a particular analysis. This could lead to different methods of computing item nonresponse rates and may potentially lead to conflicting rates between reports. Therefore, when comparing item nonresponse rates across reports, it is imperative to understand the way the rates are computed to make sure that they are comparable.

Item nonresponse for the 100 percent items ranged from 1.13 percent for the sex item to 4.14

percent for the tenure item. Generally item nonresponse was higher for enumerator returns than for self-responses and higher for long forms than for short forms. Tenure had a relatively higher overall nonresponse rate compared to other items.

Recommendations include:

- Review the procedures and debriefings of field staff to see if they can provide useful information about problems that could have led to item nonresponse.
- Continue to test question wording and placement.
- Investigate ways to reduce item nonresponse.
- Review the results of Evaluation B.1.a in conjunction with the results of this evaluation to obtain a more comprehensive view of data quality.
- Investigate the use of content followup for relatively high item nonresponse items.
- Look at the age item in combination with the date of birth item during future analysis.

Evaluation B.3, Census Quality Survey to Evaluate Responses to the Census 2000 Question on Race: An Introduction to the Data (Bentley et al 2003)

Data on race from most federal surveys currently reflect a collection methodology that asks respondents to mark only one category. Census 2000 was the first decennial census to ask respondents to “mark one or more races.” Some data users may want to compare the race distribution from Census 2000 with those of other data sources where respondents were asked to mark only one race for each person in a household. The Office of Management and Budget refers to this comparison as “bridging.”

The Census Quality Survey enables users to make comparisons between race data obtained using “mark one race” and “mark one or more races” methods by collecting race data using both methods from the same people. The Census Quality Survey was designed with the primary objective of producing a data file that could be used to bridge between “single” and “one or more races” distributions. The Census Quality Survey had a nationally representative design with two data collection points. Respondents were asked at one point to “mark one race” and at another point to “mark one or more races.” The sample was split into two panels. Panel A received the “mark one or more races” instruction at the initial contact and Panel B received the “mark one race” instruction first. During the recontact, each panel received the alternate instruction. Data from these two contacts can be used to produce “bridging parameters” to compare race distributions collected under single race and one or more race methodologies.

Initially, about 27,500 housing unit addresses were designated to be in sample for each panel. Of the eligible addresses, 97 percent completed an interview in the initial contact. In the recontact, sample housing units were contacted only if an initial questionnaire was completed.

Of the eligible recontact addresses, 87 percent completed an interview in Panel A and 94 percent completed an interview in Panel B. The results from the question on race suggest that each panel appears to be representative of race distributions in Census 2000.

Forty percent of the non-Hispanic respondents in Panel A who reported Two or More Races in Census 2000 also reported Two or More Races in the initial contact. Similarly, 41 percent of those in Panel B who reported Two or More Races in the census also reported Two or More Races in the recontact. The effective sample size for computing bridging parameters is reduced because of the generally low level of consistency in the reporting of Two or More Races. In contrast, 97 to 98 percent of those who reported a single race of White, Black, or Asian in Census 2000 reported the same race in the Census Quality Survey.

The “mark one or more races” data collection contact was cross-tabulated with the “mark one race” contact to assess how individuals respond when asked to choose a single race for people for whom multiple races have been reported. Even with the “mark one race” instruction, a significant portion of respondents reported Two or More Races. This portion was greatly reduced when the followup race probe was used in the Panel A recontact. Users of the data file will need to determine how best to treat these reluctant cases when computing bridging parameters.

Evaluation B.5, Content Reinterview Survey: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview (Singer and Ennis 2003)

The Content Reinterview Survey was designed to evaluate the consistency of responses to the Census 2000 questionnaire. Previous content reinterview surveys attempted to evaluate both response variance (the variation in responses over repeated questioning) and bias. In 2000, only response variance was studied. To reduce cost and the burden to respondents, the 2000 Content Reinterview Survey asked population questions about only one sample person per household, who was randomly chosen from a roster for each unit that was collected at the beginning of the Content Reinterview Survey.

Prior to Census 2000 enumeration, 30,000 households that were initially selected to receive the census long-form questionnaire were randomly selected as potential participants in the Content Reinterview Survey. After a household returned the census questionnaire, it became eligible to participate in the reinterview survey. Experienced census field representatives called the selected households to re-ask most of the census long-form questions. Personal visit interviews were allowed if the households could not be reached by telephone.

For the Content Reinterview Survey, data were analyzed from about 20,000 of the preselected households. Around three-quarters of the cases analyzed had completed the mailback forms for Census 2000. About three-fifths of all preselected reinterview households completed Census 2000 mailback forms.

Since the Content Reinterview Survey was conducted by enumerators who used either telephone interviews or personal visits, collection mode for the reinterview survey was different from that of the census in the majority of analyzed cases.

Based on data collected in the census and the reinterview survey, analysts computed the index of inconsistency (a measure to detect response variance) and used it to evaluate the consistency of each item at the national level. A high index of inconsistency (50 or more) for a question indicated that the question was problematic because the data elicited by the question was not consistent. A low index (below 20) indicated that the data elicited by the question was probably consistent. A moderate index (20 up to 50) indicated that the question was somewhat problematic.

Of the 58 population characteristics evaluated by the Content Reinterview Survey, 16 showed low inconsistency, 26 showed moderate inconsistency, and 16 showed high inconsistency. The items that showed low inconsistency included questions about sex, age, Hispanic origin, marital status, school attendance, language spoken at home, place of birth, citizenship, year of entry to the U.S., and veteran status and period of military service. The items that showed high inconsistency included questions about language usage, disability, grandparents as caregivers, work experience in 1999, and income. For the first time ever, Census 2000 allowed the respondent to choose one or more races in response to the race question. The edited race data displayed moderate inconsistency.

Of the 36 housing characteristic items measured, five showed low inconsistency, 15 showed moderate inconsistency, and 16 showed high inconsistency. The items with low inconsistency included questions about the number of people in the household, whether the unit was owned or rented, heating fuel, whether there was a mortgage on the property, and if real estate taxes were included in the mortgage payment. The items with high inconsistency included questions about utility costs for gas and for electricity, second mortgages and home equity loans, loans on mobile homes, the value of the property and insurance costs for the property, the number of rooms in the house/apartment/mobile home, whether there was a business at the site and the total value of agricultural sales for the property, and plumbing facilities, kitchen facilities, and telephone service.

Sufficient data were gathered to compare indexes of inconsistency by collection type for 87 items. At the 90 percent confidence level, 51 items showed less inconsistency for mailback forms and two showed less inconsistency for enumerator forms. The two that were less inconsistent when collected by enumerators were “Do you speak a language other than English at home?” and “What is the annual cost for Gas?”

Key recommendations follow:

- Use cognitive experts to recommend improvements to problematic questions.
- Use results from content tests in developing questionnaires for the 2010 Census and the American Community Survey (ACS).

- Provide better instructions on the 2010 Census and the ACS for the Hispanic-origin question.
- Use separate “Yes/No” questions for each response category of “mark all that apply” questions.
- Create a database that links changes in identifiers for census cases to enable locating Content Reinterview Survey cases when there are changes to Master Address File (MAF) IDs.
- Plan the content reinterview surveys of the 2010 Census and the ACS as early as possible and conduct them within three or four weeks of completing the original data collection.
- To the extent possible, use the same data collection modes, data capture methods, data processing procedures, and enumerators for both the 2010 Census and its content reinterview and for both the ACS and its content reinterview.
- For time-sensitive questions, refer to the date of the original survey in the content reinterview for both the ACS and the 2010 Census.

Evaluation B.6, Master Trace Sample (Hill and Machowski 2003)

The Master Trace Sample database project merged Census 2000 data from multiple sources to provide information about cases in the various phases of data collection and processing. The objective of this effort was to support future methodological and operational analyses and decisions regarding the 2010 Census by creating a complex, relational database for research purposes. The prototype database merged Census 2000 address frame, collection, enumeration, capture, processing, response, and coverage files. This merge yields a sophisticated database which allows quantitative insight into the relationship of key census processes. In addition to being an innovative research tool, the Census 2000 Master Trace Sample database is intended to serve as a model upon which the Census Bureau will improve in future censuses.

The Master Trace Sample database contains a sample of Census 2000 housing unit records that allow Census Bureau researchers to trace response and operational data through stages of Census 2000 processing. These stages include address list development, data collection, data capture, and data processing. For the sample of housing unit records, the database contains all returns, which include 100 percent housing unit and person data.

The Master Trace Sample database also contains data not typically analyzed in census evaluations. For example, the number of times an enumerator visited a housing unit during nonresponse followup is contained in this database. In addition, the database links micro-level data such as enumerator production data with response data, which are not traditionally linked in census evaluations. The purpose of the Master Trace Sample database is to facilitate research on relationships among Census 2000 operations beyond the current Census 2000 Testing, Experimentation, and Evaluation Program.

The database contains a total of approximately 1.5 million MAF housing unit IDs from a systematic ID sample and a block cluster sample. The block cluster sample contains all housing unit IDs within selected block clusters.

The primary recommendations are:

Recommendations for Expanding the Census 2000 Master Trace Sample Database:

- Expand the Master Trace Sample to include data on Group Quarters.
- Expand the Master Trace Sample database to include coverage measurement data associated with persons.

Recommendation for the 2010 Census Master Trace Sample Database:

- Implement a formal evaluation to assess both the usefulness of the database for research and the benefits to the Census Bureau of resulting products.

Recommendation for Creating Master Trace Sample Databases for other Censuses and Surveys:

- Provided the proposed formal evaluation of the Census 2000 Master Trace Sample finds the database useful, the Census Bureau may wish to consider building such a ‘trace’ database specific to each of its major surveys, as well as the Economic Censuses.

Evaluation B.7, Accuracy of Data for Employment Status as Measured by the CPS - Census 2000 Match (Palumbo and Siegel 2004)

This evaluation presents the results of an exact match study that used the Current Population Survey (CPS) - Census 2000 Match to evaluate the labor force data in Census 2000 by making estimates of their content error (content error refers to the accuracy of the data, as opposed to coverage error, which refers to how completely people and housing units are counted). The evaluation contains a description of the methods used to create the match file and a description of how the file was used to measure levels of content error.

For people in Census 2000 who were also in the CPS sample in February through May 2000, the CPS - Census 2000 Match brought together each person’s census report with the same person’s CPS report. This linkage provided the opportunity to compare two independent observations of the same event (the person’s relationship to the work force at a particular time) and to use the outcome of the one observation (the person’s labor force classification in the CPS) to measure the accuracy of the outcome of the other (the same person’s labor force classification in Census 2000). The CPS was used because it is considered to be the standard of comparison for census labor force data. The analysis in this evaluation was restricted to the national level.

Cross-tabulation of the two observations presents estimates of the potential quantities of *response error* in published census figures. A response error is said to occur when a person’s

labor force classification in Census 2000 as either *employed*, *unemployed*, or *not in labor force* differs from that same person's classification in the CPS. To make these quantities meaningful, two relative measures of response errors (percentage distributions) and two summary measures of response errors were derived from them. The derived measures are the focus of the evaluation. The percentage distributions reveal the success rates of Census 2000 in classifying people to their correct (same as CPS) labor force categories and away from incorrect (different from CPS) categories.

Findings from the CPS - Census 2000 Match include:

- Census 2000 and the CPS are reasonably consistent in classifying people to the employed and not in labor force categories, but they exhibit considerable variability in classifying people to the unemployed category.
- Previous studies of census-CPS employment classifications, which were done for the 1960 and 1970 censuses but not the 1980 and 1990 censuses, revealed patterns similar to those described above. However, for Census 2000, consistency slipped somewhat from the 1970 levels, in spite of efforts, particularly after the 1990 Census, to make the census employment questions conform more closely with the CPS questions.
- As was true in the 1970 and 1960 studies, the index of inconsistency measures for the unemployed category were in the high range (above 50), which calls into question whether the unemployed concept is measurable in a census context.
- The results for the employed and not in labor force categories indicated that, although the census is able to measure these concepts reasonably well, improvements are needed. The study suggested, for example, that it may have been a mistake to use the CPS wording for the "work last week" question in Census 2000.
- The underestimate of employment and the overestimate of people not in the labor force in Census 2000 relative to the CPS is likely related to the failure of the census classification system to filter more employed people out of the not in labor force category and into the employed category.
- The difference between the reference periods for the labor force estimates of Census 2000 and the CPS is probably not a major contributor to the gaps between the estimates.
- A tendency for people classified as employed in the CPS to be classified as *not* employed in Census 2000 appeared to be associated with particular age categories, class of worker categories, and educational attainment categories. The finding suggests that some groups of workers may have had difficulty in understanding or correctly responding to the work-last-week question in the census.

Evaluation B.12, Puerto Rico Census 2000 Responses to the Race and Ethnicity Questions (Christenson 2003)

Prior to Census 2000, race data were last collected in Puerto Rico by enumerators in the 1950 Census of Puerto Rico. Hispanic origin had never before been asked in Puerto Rico. In Census 2000, the questions on race and Hispanic origin in Puerto Rico were identical to the questions asked in the 50 states and the District of Columbia (DC). Both the newness of the data and the uncertainties associated with the race question contributed to a high level of interest in the responses by residents of Puerto Rico to these questions.

The goal of this study was to examine the data resulting from the responses to the race and Hispanic origin questions by the residents of Puerto Rico and to compare them with those resulting from responses to the race and Hispanic origin questions by residents of the 50 states and DC.

The analysis shows that the residents of Puerto Rico identified themselves as overwhelmingly of Hispanic origin and of a single race. In terms of race, the great majority identified themselves as White, with a substantial minority reporting themselves as Black or African American. When compared with those of Hispanic origin in the 50 states and DC, substantially fewer reported themselves to be of Some Other Race, and a lower percentage identified themselves as of Two or More Races. The analysis also shows that two typical indicators of problematic questions, item nonresponse and differences between respondent and enumerator completed questionnaires, did not indicate major problems with either question.

From these findings come the following recommendations for improvements and for further research.

- Include Puerto Rico in the cognitive testing or efforts to field test different versions and formats of questions and questionnaires that are currently underway for the 2010 Census.
- Investigate further the use of the “Some Other Race” category to assure that the range of responses excludes those that better fit the Hispanic origin question.

Evaluation B.13, Puerto Rico Focus Groups on the Census 2000 Race and Ethnicity Questions (Berkowitz 2001a)

This evaluation explored the views and perceptions of residents of Puerto Rico on the Census 2000 Puerto Rico short-form mailback questionnaire items on race and Hispanic origin. Contractors conducted focus groups in 12 sites across the island selected for geographic and socioeconomic diversity, recruiting participants of diverse ages and educational levels, including some who had lived in the U.S. for an extended period. Although these results are based on a relatively small, purposive sample and cannot be generalized to the Puerto Rican population as a whole, they provide an interesting jumping off point for additional thinking and further research.

There was unanimous agreement among focus group participants that the question on race is inappropriate to the Puerto Rican context. Participants could not find themselves reflected in the available answer categories, which they viewed as foreign to the Puerto Rico's history and culture of mixing or blending across racial groups. Some also felt the question was inherently racist, discriminatory, and divisive, and suspected a hidden political agenda. Taking their lead from the national origin terms in the second half of the question, some participants chose to define race as nationality. However, they were bothered that "Puerto Rican" did not appear as a preprinted category. Participants were not satisfied with the option of checking off multiple racial categories for an individual, because they did not perceive themselves as biracial or multiracial but, rather, as mixed.

The presence of a preprinted "Yes, Puerto Rican" answer category rescued the Hispanic origin question from the same fate as the question on race. In general, participants' interpretations of this question were highly context-dependent. They disagreed as to whether the terms "Spanish," "Hispanic," and "Latino" all meant the same thing and in their assessments of its uses as a blanket category. "Origin" was variously understood as birthplace, ancestry, nationality, and self-identification. In the end, participants decided this question was not nearly as simple and straightforward as it had first appeared.

On the relationship of the two questions, the largest cluster of focus group participants viewed them as so redundant that it made no sense to ask both and favored eliminating the race question as the more offensive and less informative of the two.

Key recommendations include:

- Include residents of Puerto Rico (in addition to persons of Puerto Rican origin living in the U.S.) in any subsequent cognitive testing or efforts to field test different versions of questions and questionnaires for future censuses.
- Provide more extensive public education in Puerto Rico on the larger mission of the census, the rationale for asking questions about race and ethnicity, and the intended uses of the data.
- Consider using the results of this study, along with the results of Evaluation A.8 to create a survey to be administered to a probability sample of residents of Puerto Rico.

Category C - Data Products

Evaluation C.1, The Effects of the Disclosure Limitation Procedure on Census 2000 Tabular Data Products, Abridged (Steel and Zayatz 2003)

The full report for this evaluation is not available because it contains proprietary information. Most information in the full evaluation is Census Confidential. The full evaluation cannot be removed from Census Bureau facilities and is available to Census Bureau personnel on a need-

to-know basis. Abridged information is primarily descriptive and qualitative. Quantitative information can only be found in the unabridged evaluation.

Data swapping was used to protect the confidentiality of the Census 2000 tabulations. The procedure was performed on the underlying microdata and all tabulations from the 100 percent (short form) and from the sample (long form) data were created from the swapped files. It affected pairs of households (or partnered households) where one or both of those households had a high risk of disclosure. The set of census households that were deemed as having a disclosure risk was selected from the internal census data files. These households were unique in their geographic area (block for 100 percent data and block group for sample data) based on certain characteristics. The data from these households were swapped with data from partnered households that had identical characteristics on a certain set of key variables but were from different geographic locations. The swapping procedure was performed independently for the 100 percent data and the sample data.

To maintain data quality, there was a maximum percent of records that were swapped for each state for the 100 percent data and another maximum percent for the sample data. Presumably, the higher the rate of swapping, the greater the confidentiality protection but the lower the data quality. However, the way the procedure is targeted to records with disclosure risk and the choice of variables that are controlled on and the choice of variables that are not swapped also affect the resulting levels of protection and quality. The main goal was to see if the Census Bureau were able to strike the right balance between protecting confidentiality and maintaining data quality.

To answer questions on data quality, evaluators compared tables from swapped versus unswapped data, examined the changes in cell values due to the swapping for cells of different sizes, and compared swapped and unswapped sample estimates of 100 percent data items. Evaluators also compared the effects of swapping among different geographic levels. To answer questions on data protection, evaluators looked at how often the Census Bureau was able to swap households with a high disclosure risk. Some calculations were performed on all 50 states. For the most detailed analysis, calculations were performed on three states (Oklahoma, Massachusetts, and Mississippi) for the 100 percent data and three states (West Virginia, New Jersey, and Vermont) for the sample data.

For the 100 percent data, all records were given a chance of being swapped. The swapping was applied consistently in each state. Records were assigned a level of disclosure risk from 1 to 4 with 4 having the most disclosure risk. The procedure for assigning the levels of disclosure risk is Census Confidential. All level 4 records were swapped. The performance on levels 3, 2, and 1 varied from state to state and was generally better for urban states with a diverse population.

For the sample data, all records were given some chance of being swapped. A small percent of households were swapped in each state. Again records were assigned a level of disclosure risk. Records were chosen for swapping based on their level of disclosure risk and the ability to pair records with high levels of disclosure risk. Most records deemed as having a disclosure risk were swapped.

Using variables that are common to both the 100 percent and sample data, the Census Bureau found that the confidence interval about the swapped sample estimate covers the true 100 percent value nearly as often as the interval about the unswapped estimate. Results were better in urban states with a diverse population.

The data swapping procedure was checked for quality. It was conducted correctly and consistently. Minimum but necessary changes were made to the data in such a way that maximized data quality.

The disclosure limitation model used for Census 2000 is useful and the Census Bureau should continue future research on disclosure limitation techniques. The Census Bureau should include confidentiality protection as part of the process when planning a census.

Category D - Partnership and Marketing

Summaries of the three evaluations in the Partnership and Marketing category follow.

Evaluation D.1, Partnership and Marketing Program Evaluation (Wolter et al 2002)

The Census Bureau contracted with the National Opinion Research Center to evaluate whether the Partnership and Marketing Program increased the public's awareness of the census and mailback response rates, especially among historically undercounted populations. The National Opinion Research Center implemented a before, during, and after research design with three waves of interviewing. Wave 1 occurred in Fall 1999 before the launch of the education phase of the advertising program and before most partnership activities had commenced; Wave 2 took place in Winter 2000 before the mailout of census forms; and Wave 3 began in Spring 2000 following Census Day and continued during the census NRFU operations. Across the three waves of data collection, the National Opinion Research Center completed just under 10,000 interviews of American households. The surveys sought to interview the person in the household who opens the mail or the one most likely to open and answer the census form.

The research design incorporated representative samples of several race/ethnicity populations, including Hispanics, non-Hispanic Blacks, non-Hispanic Whites, Asians, American Indians, and Native Hawaiians. It enabled separate analysis and conclusions about the effectiveness of the Partnership and Marketing Program for these race/ethnicity populations, as well as for the total population. The design also incorporated an exact match of the survey responses to the actual census returns for the households interviewed in Waves 2 and 3.

Overall awareness of communications about Census 2000 increased significantly over time. It was greater after the Census 2000 Partnership and Marketing Program than before the onset of the program. Awareness of communications about Census 2000 increased for all six of the race/ethnicity populations, including historically hard-to-enumerate populations such as Hispanics, non-Hispanic Blacks, and American Indians. It appears that the program was effective for all targeted populations in stimulating awareness.

The effects of census marketing and partnership activities are confounded with one another. As such, it was impossible for the evaluation to measure their effects separately. The study did examine, however, the public's recall of eighteen sources of census communications, each of which exhibited a combination of advertising and partnership influences. To strengthen the analysis, evaluators combined the eighteen sources into two composite measures: mass-media and community-based communications. Mass-media communications included television, magazine, radio, newspaper, and billboard ads. Community-based communications included religious groups, community or government organizations, informal conversations, schools you attended, schools your children attended, census job announcements, conference exhibit booths, signs inside buildings, speeches, articles, the Internet, paycheck or utility bill, and participation on a complete count committee. Results showed significant evidence that awareness of both types of communications was greater after the Census 2000 Partnership and Marketing Program than before the onset of the program.

Four race/ethnicity populations (non-Hispanic Blacks, non-Hispanic Whites, Asians, and Native Hawaiians) indicated that they were more likely to return the census form after the Census 2000 Partnership and Marketing Program than before its onset. Higher awareness of communications about Census 2000 correlates with a greater likelihood or intention of returning the census form for five of the targeted populations (Hispanics, non-Hispanic Blacks, non-Hispanic Whites, Asians, and Native Hawaiians). Hispanics show this effect even though their mean intended participation did not increase from before to after the Census 2000 Partnership and Marketing Program, suggesting that the program had less impact on them.

According to the data, attitudes towards census confidentiality declined at the close of the 20th Century. Favorable attitudes started at a low level prior to Census 2000 and never recovered to the levels reported in 1990. On the other hand, respondents' views of the importance of participating in the census remained quite stable: both censuses exhibited similarly favorable attitudes and neither displayed a trend from wave to wave within the census period. Finally, the Census 2000 Partnership and Marketing Program seems to have achieved greater success than comparable efforts in 1990 to create a favorable attitude that the census cannot be used against you.

In light of these findings, the recommendations are:

- Repeat a program of mass-media and community-based communications in general form, content, and intensity for the 2010 Census.
- Evaluate current communications channels, with an eye towards optimizing the allocation of Partnership and Marketing Program resources among the various channels.
- Reevaluate what promotional messages resonate best with the American population, overall, and with targeted race/ethnicity populations.
- Build on the success of the Census 2000 Partnership and Marketing Program for the Black population.

- Reevaluate the communications approach for the Hispanic, Native Hawaiian, non-English speaking Asians, and American Indian populations.
- Conduct a formal cost-benefit analysis, attempting to demonstrate the trade-offs between increased expenditures on Partnership and Marketing Program activities and reduced followup costs.
- Use an experimental design to measure the effectiveness and benefit of a partnership and marketing program.

Evaluation D.2, Evaluation of the Census in Schools Program: Materials and Distribution
(Macro International 2002)

The Census in Schools Program aimed to raise awareness of Census 2000. The program offered teaching materials that provided information on the purposes and methods of the census and that sought to engender an interest in the census. The program was particularly targeted at schools in hard-to-enumerate areas. This evaluation was intended to measure the effectiveness of the approach for disseminating Census in Schools Program materials and use of and satisfaction with the materials among teachers. It was not intended to measure the impact of the Census in Schools Program materials on children or their parents or on ensuring a high rate of participation in Census 2000.

All elementary school teachers and all secondary math or social studies teachers in hard-to-enumerate areas were sent an invitational packet. This invitational packet consisted of an informational letter and an order form. This packet provided teachers with the opportunity to order Census in Schools Program materials, which included a Teaching Guide, lesson plans, and a Giant U.S. map. Principals, other than those in hard-to-enumerate areas, administrators, and curriculum coordinators also received an invitational packet. Additionally, all elementary school teachers and middle school social studies teachers were sent Take-Home materials for students to learn about the census and share with their parents at home, thereby having the potential to reach each kindergarten to eighth grade student in the country.

ORC/Macro International conducted the evaluation of the Census in Schools Program based on a survey they fielded in Spring 2000. They mailed to a stratified random sample of 4,000 teachers selected from all primary and secondary teachers in the 50 states, DC, and Puerto Rico and from teachers ordering the Census in Schools materials. Of the teachers selected for the survey, 1,101 responded.

Approximately 56 percent of all teachers heard of the Census in Schools Program. The single most important conduit for information about the program was the invitational packet. Fifty-four percent of teachers heard about the program through invitational packets. Approximately 23 percent heard about it from their principal. Overall, 63 percent of all teachers who heard about the Census in Schools Program received at least one component of the Census in Schools Program materials. Thirty-nine percent of teachers who received the materials acquired them from their principals. About 23 percent of the teachers who received the materials ordered them

in response to the invitational packet. Teachers who heard about the Census in Schools Program from their principal rather than hearing of it through other sources, were more likely to actually receive the materials.

The Giant U.S. Map was popular. Of the 85 percent of teachers who received the map, 92 percent used it in classroom activities. Of the 33 percent of teachers who received the Take-Home materials, about 79 percent sent them home with their students.

About 39 percent of all teachers in the 50 states, DC, and Puerto Rico taught in hard-to-enumerate areas. Of those in hard-to-enumerate areas who heard of the Census in Schools Program, almost 61 percent received at least one component of the Census in Schools Program materials. Sixty-four percent of teachers in hard-to-enumerate areas who heard about the Census in Schools Program through the invitational packet received at least one component of the Census in Schools materials. Teachers in hard-to-enumerate areas were more likely to have ordered the materials themselves (34 percent) than teachers in other areas (16 percent).

Fifty-three teachers provided answers to an open-ended question about why they did not send Take-Home materials home. Responses included that teachers need more lead time to examine the Census in Schools Program materials and incorporate them into their curricula. Some teachers found the Census in Schools Program materials too difficult for their students and others thought they were too elementary. Some teachers received the Census in Schools Program materials in the wrong language or targeting an inappropriate age group.

In examining the results of this evaluation, some themes appeared:

- Principals were an important conduit for transferring information about Census in Schools Program materials as well as for ordering the materials.
- It seems that the invitational packet did not draw the attention of many teachers to whom it was sent. Teachers receive many items in their mailboxes and the invitational packets did not appear to stand out from other materials sent to teachers.
- For those using the Census in Schools Program materials, satisfaction was high.

In view of these findings, the recommendations are:

- Conduct research/testing to understand how and when teachers react to various types of mailings and to better understand how teachers can be reached.
- Design the materials to better meet the needs of the teachers.
- Use principals to transmit the Census in Schools Program materials to teachers.

Evaluation D.3, Report of Survey of Partners (Westat 2001)

The mission of the Partnership Program was to develop an aggressive and comprehensive program that incorporated the efforts and resources of government units, community-based organizations, religious groups, and businesses to assist the Census Bureau in conducting an efficient, accurate census. The primary goals of the program were to increase mail response rates, reduce the differential undercounts, and communicate a consistent Census 2000 message. To achieve these goals, the Census Bureau formed partnerships with state, local, and tribal governments, non-governmental organizations, community groups, the media, and private sector businesses. The Census 2000 Partnership Program also included 690 partnership staff at headquarters and across all 12 regions.

A survey-based study was conducted to evaluate the program's effectiveness from the partners' viewpoint. A model of organizational relationships was used as an organizing framework for the evaluation. The components of the model addressed by the survey were:

- Benefits partners expected to achieve from their partnerships with the Census Bureau.
- Census Bureau contributions to the partnership (the wide variety of materials Census provided to participating organizations).
- Partner contributions to the partnership (activities partners conducted to publicize and increase awareness of the census, to get their target populations counted, and to assist with Census Bureau operations and initiatives, financial contributions, and in-kind contributions partners made to support and promote Census 2000 efforts).
- Structures and processes that existed between the Census Bureau and partners to accomplish partnership goals.

Census Bureau staff drew a stratified random sample of 15,803 from a frame of partners that were entered in the Contact Profile Usage and Management System at the time the sample was drawn. Data were collected over a six month period by both mail and Computer Assisted Telephone Interviewing (CATI) methods. The survey achieved a 67.9 percent response rate.

Seventy to 81 percent of partners responded that they placed "Moderate emphasis" or "A lot of emphasis" on each of five Partnership Program goals. Partners' expected benefits of participation were aligned with the Census Bureau's goals for the program. From the partners' view, contributions the Census Bureau made were highly valued. The majority of partners (ranging from 71.3 percent to 88.0 percent) that used each of the 18 types of materials rated the material as "Moderately Helpful" or "Very Helpful." Non-English materials were used by more than 90 percent of all organizations that received them and these materials were also rated as "Moderately Helpful" or "Very Helpful" by more than 80 percent of partners that used them.

Seventy percent of respondent organizations reported that they conducted one or more activities. Mean ratings for all activities were above the 3.1 level of the four point scale (1 = "Not at all

Helpful” to 4 = “Very Helpful”), indicating that across all partners, every activity was considered to be at least “Moderately Helpful” in achieving Partnership Program goals. Relatively few of the partners responded that they made any type of financial contributions to the partnership. Results indicated that partners contributed more in terms of resources (e.g., staff time, space, materials, etc.) rather than spending organizational funds.

The majority of partners (70 percent) reported that Census Partnership Specialists helped them promote Census 2000. More than half of the partners reported that the direct Census support and Census participation in their activities was helpful. Overall, partners were satisfied with the process in place to furnish them with Partnership materials. A majority of partners indicated that the Partnership Program helped them to reach their goals for participating, more so for goals of reaching and educating the target population (67 percent and 72 percent, respectively) than for minimizing the target population's fear of providing information to the government (60 percent).

Of partners expressing an opinion (79 percent of all partners), 84 percent were positive about their intent to participate as partners again. This result suggests the overall success of the program.

Recommendations for practices that should stay the same:

- Continue to define common goals that partners perceive as benefits.
- Continue use of the variety of materials for education and awareness. Continue to make use of the specific materials that were rated most used and most helpful. Continue to develop and use language-appropriate materials.
- Encourage future partners to conduct the types of activities that partners considered successful during Census 2000.
- Continue to provide liaison support to partners through Partnership Specialists.
- Continue to provide direct Census support for partner activities and Census participation in those activities.

Recommendations for practices that should change:

- Make partner benefits more explicit.
- Re-evaluate the use of specific materials that were rated least used and least helpful.
- Improve the process for furnishing materials to partners.
- Incorporate qualitative feedback obtained from partners into any plans for improvement for the next census cycle.

Category E - Special Places and Group Quarters

Summaries of the three evaluations in the Special Places and Group Quarters category follow.

Evaluation E.1.b, Evaluation of the Facility Questionnaire (Computer Assisted Telephone Interviewing and Personal Visit) (Stevens 2003)

This evaluation compared the classification of special places, also known as special place type coding, collected at three different times during Census 2000. It specifically included comparisons of the special place type code recorded during the Special Place Facility Questionnaire operation, Reinterview, and the Special Place Advance Visit operation. Special place type coding was conducted to determine the path the remainder of the interview would follow in the Special Place Facility Questionnaire and Special Place Advanced Visit operations. In addition, subsequent operations of data collection were dependent on it. This evaluation was designed to answer two questions related to special place type coding which help describe the consistency of the special place type coding across the three operations listed above.

Note that, although the Special Place Facility Questionnaire and Special Place Advance Visit operations recorded special place type codes, the primary purpose of these operations was to collect group quarters type codes. This evaluation focused on special place type codes. The conclusions from this evaluation should not be applied to group quarters type codes. The 2010 Census plan is to emphasize group quarters type codes rather than special place type codes. Therefore, results from this evaluation may not be influential to the 2010 design.

About 25 percent of the special place codes changed from the Special Place Facility Questionnaire to the Reinterview and 38 percent of the special place type codes changed from the Reinterview to the Special Place Advance Visit. Omitting the Reinterview operation, 25 percent of the special place type codes changed from the Special Place Facility Questionnaire to the Special Place Advance Visit. However, of the 25 percent that changed from the Special Place Facility Questionnaire to the Special Place Advance Visit, most are different because the special place was deleted from the census and not because of a change in the coding of the special place.

Discrepancies in the special place type code differ somewhat by type of special place. Specifically, large/complex special places were more likely to have a discrepancy in the special place type code than other special places. However, this difference may not be of practical significance.

Recommendations include that future evaluations base the comparison on the group quarters type code instead of, or in addition to, the special place type code. A group quarters type code comparison is more appropriate since it is the classification by which data are tabulated in census products.

Evaluation E.5, Revision 1, Group Quarters Enumeration (Jonas 2003b)

The Group Quarters population consists of all persons who do not live in housing units such as single-family houses, apartments, and mobile homes, but rather in group situations such as college dormitories, nursing homes, military barracks, prisons, juvenile institutions, migrant worker dormitories, convents, and group homes. A Special Place is an administrative entity containing one or more Group Quarters. The Group Quarters are where people sleep. For instance, a university is a Special Place and each dormitory is a Group Quarters. This evaluation was designed to document the counts of Special Places, Group Quarters, and Group Quarters' population and other operational aspects of the Group Quarters enumeration.

Group Quarters enumeration succeeded in its underlying mission of gaining a fundamentally accurate count of the Group Quarters population. The Group Quarters universe was home to 7.8 million people in 2000. These people were enumerated in 192,286 Group Quarters in 100,358 Special Places. Universities, military bases, and correctional institutions were the largest special places, as measured by both size of population and number of Group Quarters per Special Place. Special Places with 1,000 or more people were made up almost entirely of persons in colleges, prisons, and military bases (99 percent of that population were in these types of places).

Overall, forty percent of Special Places had less than ten residents and 61 percent had less than 25 residents. These were mostly group homes and Service-Based Enumeration (SBE) sites or "other" Special Places. The 40 percent of Special Places with less than ten residents had only 2.3 percent of the Group Quarters population.

More Group Quarters questionnaires were filled out from administrative data than by any other method, with nursing homes, hospitals, group homes, and correctional institutions using administrative data most frequently.

Over 200,000 Group Quarters person records (2.6 percent of all Group Quarters person records) had all characteristics imputed. Over 55,000 of these persons were enumerated on questionnaires that could not be processed because the hand-transcribed Group Quarters ID was either not entered on the form by an enumerator or was an erroneous ID that could not be identified with a Group Quarters.

An estimated 4.4 percent of persons in Group Homes or religious Group Quarters were counted twice; once by the Group Quarters enumeration process and again on a household questionnaire returned by mail. Scheduling constraints did not allow Census addresses to be removed from the housing unit address mail list if they were also found to be included as a Group Quarters address. As such, some addresses were identified as duplicate addresses but enumerated twice. This occurred primarily among Group Homes and religious Group Quarters.

Nearly 150,000 housing units were identified by Group Quarters enumerators at Group Quarters and at transient locations such as recreational vehicle parks. These housing units contributed over 260,000 persons to the census.

Recommendations include a number of changes to Group Quarters enumeration for 2010:

- Use available technology to track individual questionnaires from enumeration to data capture.
- Improve the address list creation process by gathering data on Special Places from Web-based sources, tailor address list creation and enumeration strategies to each major category of Group Quarters, and give large Special Places the option of providing Group Quarters data by electronic or printed records rather than by telephone and in-person interviews.
- Reduce duplication between the address files for Group Quarters and housing units.
- Be prepared for use of administrative data in enumeration.

Evaluation E.6, Service-Based Enumeration (McNally 2002)

The goal of SBE was to provide people experiencing homelessness an opportunity to be included in the census. The Census Bureau developed a specialized operation to enumerate selected service locations that serve people experiencing homelessness. It is important to note that the Census 2000 count of the SBE population does not represent a complete count of people experiencing homelessness.

Between March 27 and March 29, 2000, the Census Bureau enumerated people at emergency shelters, soup kitchens, regularly scheduled mobile food vans, and targeted nonsheltered outdoor locations. People on Be Counted Forms who marked the "No Address on April 1, 2000" box or indicated they were homeless in the address section also were included in the SBE universe.

There were 14,817 SBE sites in Census 2000. More than half (51 percent) of the locations were shelters. There were a total of 258,728 person records data captured from shelters, soup kitchens, regularly scheduled mobile food vans, and targeted nonsheltered outdoor locations. Most of the data captured person records (90 percent) were from shelters, soup kitchens and regularly scheduled mobile food vans.

The SBE operation appears to be a successful method of including in the census people who are experiencing homelessness. A total of 283,898 people were tabulated in Census 2000 as a result of the SBE operation. This total includes 9,963 persons imputed into the Census 2000 population.

There was a total of 258,728 questionnaires captured from the SBE program. Almost all (99 percent) of the data captured person records had at least two or more data characteristics (name, sex, age and/or date of birth, Hispanic origin, and race).

Approximately 87 percent of the data captured persons contained enough information for the questionnaire to be included in the unduplication process. That is, the questionnaire had a first and last name with combined fields containing at least three alphabetic characters and at least two person characteristics, one of which was date of birth or age. The Census Bureau was able to match and unduplicate 16,787 person records during data processing. However, it was discovered that 2,410 of these records were erroneously identified as a duplicate record.

A total of 38,415 people completed a Be Counted Form and marked the "No Address on April 1, 2000" box on that form or indicated they were homeless in the address section. Of these, the Census Bureau was able to match and unduplicate 3 percent to people enumerated during the SBE operation. The Be Counted program added 35,121 persons without a usual address to Census 2000. Most of these (31,994) were tabulated with the population for SBE sites. The remainder were tabulated with the Group Quarters population.

Based on the results of this evaluation, the Census Bureau should continue supporting SBE for the 2010 Census.

Category F - Address List Development

Summaries of the ten evaluations in the Address List Development category follow.

Evaluation F.2, The Address Listing Operation and Its Impact on the MAF (Ruhnke 2002)

The evaluation of the Address Listing operation for Census 2000 examined the operation's impact on creating the MAF for certain areas of the country. The Census Bureau conducted the Address Listing operation from July 1998 to May 1999 and used the results to create the initial address list for areas that would be enumerated using U/L methodology during Census 2000. In the Address Listing operation, census enumerators canvassed door-to-door to identify the mailing address and physical location of addresses in areas where the Census Bureau believed that problems were likely with developing an accurate mailing list and delivering census questionnaires through the mail. The enumerators also located each housing unit with a map spot on a block map and collected an occupant name and telephone number, when possible. This evaluation looked at the number, geographic location, characteristics, and quality of addresses listed during the Address Listing operation.

Stateside, about 22 million housing units were listed in the Address Listing operation. An additional 1.4 million addresses were listed in Puerto Rico. All of Puerto Rico was canvassed during the Address Listing operation and was enumerated using U/L methodology.

Despite Address Listing occurring in mostly rural areas of the U.S., over 73 percent of the units had complete city-style (house number, street name) addresses. About 14 percent of the units had incomplete or no address information, but location descriptions of the units were recorded for over 95 percent of those units. Both city-style address information and location descriptions enable enumerators to locate the units on the ground when they deliver the census forms during

U/L and other census field operations. The presence of a map spot, a unique identifier for a housing unit on a census map within a block, is also crucial when trying to locate a unit in rural areas. Over 99 percent of the Address Listing adds had map spots.

Addresses eligible for the DMAF included those that represent potential residential housing units that are coded to census blocks and have map spots. Over 99 percent of the Address Listing adds were delivered to the DMAF and approximately 94 percent of all Address Listing adds were included in the final Census 2000 counts.

In areas where most mailing addresses are city-style (for example, 101 Main Street), the Census Bureau created the MAF by combining addresses from the 1990 Census Address Control File with addresses in the USPS Delivery Sequence File (DSF). Approximately 43 percent of addresses added in Address Listing matched to addresses that were identified as residential on or before the September 1998 USPS DSF. About 280,000 blocks in U/L areas had all of their addresses match to the DSF. This is about 14 percent of all blocks in which there was at least one unit listed during the Address Listing operation.

Listers were allowed two telephone callbacks to collect mailing address information during the Address Listing operation. There were three additional personal visit callbacks used to obtain address information in 36 of the approximately 3000 counties in which Address Listing was done. The 36 counties were the sites of the 1999 ACS. The additional callbacks were made to maximize mail response in that survey. It appears that the additional callbacks may have contributed to the success of obtaining additional address information, although not in any significant manner.

Recommendations resulting from this evaluation include:

- Reassess the methodology of delineating Mailout/Mailback versus U/L areas. It may be reasonable in some Census 2000 U/L enumeration areas to use the DSF as an address list building tool.
- Since the impact of the additional callbacks on obtaining mailing address information appears small and the necessary cost data were not available to do an effective cost comparison, additional callbacks for a future Address Listing operation are not recommended at this time.

Evaluation F.3, Evaluation of the Local Update of Census Addresses 98 (LUCA 98) (Owens 2003)

The Census Bureau conducted the Census 2000 LUCA 98 program in mailout/mailback areas from May 1998 to June 2000. The Census Bureau invited local and tribal governments to participate. Those who participated were sent lists of housing units in the census blocks in their area. The address list for the LUCA 98 program included addresses from various MAF sources, including the 1990 Address Control File, two USPS DSF deliveries, and the Block Canvassing operation. There were approximately 81.5 million addresses from these sources on the MAF

that were eligible for review in the LUCA 98 program. Governments updated the lists by adding, deleting, or correcting addresses. The Census Bureau then verified most of those updates. This evaluation documented the results of the LUCA 98 program.

There were 17,424 governmental units eligible to participate in the LUCA 98 program. A total of 9,263 governments participated. The housing units in these jurisdictions geographically covered approximately 92 percent of the housing units in areas eligible for LUCA 98. Although about half of all eligible governments participated, a little more than a third of eligible governments provided any updates in the form of adds, deletes, or corrections. The Census Bureau should investigate ways to increase government participation, especially focusing on ways to aid the governmental unit in providing updates once they have agreed to participate. In general, smaller governments (as determined by the number of housing units in the government's jurisdiction in 1990) had lower participation rates than larger ones. Governments may have not participated because they did not have enough resources to do the task or they knew that larger governments in their area were already updating addresses for the Census Bureau.

LUCA 98 participants reviewed address lists and added addresses for residential units in their jurisdiction that they believed did not exist on their review materials. They added 5,302,094 addresses to the MAF, which represents a 6.5 percent increase in housing units in mailout/mailback enumeration areas. Approximately 95 percent of LUCA 98 participant adds were included on the initial census address list. Many were added to the initial list as "provisional" adds, to be verified after the first census mailing. Approximately 58 percent of adds were confirmed to exist as residential addresses in the Block Canvassing operation or the LUCA Field Verification operation. About 58 percent of adds were in the final census housing unit inventory.

The LUCA 98 participants deleted (or declared nonresidential) any address on their address list that they believed did not exist in their jurisdiction as a residential unit. They deleted 490,613 addresses from the MAF. LUCA 98 participants corrected 2,762,050 addresses. The corrections included geographic as well as address information.

LUCA 98 participating governments appealed 313,853 addresses. A total of 303,410 of those addresses were added to the MAF after approval by the Census Address List Appeals Office that was set up by the Office of Management and Budget. There were 141,580 appeal addresses that were included on the final Census address list.

The participants of the LUCA 98 program contributed to the address list in many areas. Although the updates had a large impact on the update of the MAF for Census 2000, the timing of the program with other Census 2000 address updating operations introduced some complexity in determining the true impact of updates to the final census results. However, about 505,530 addresses in the final census were provided by LUCA 98 participants and may not have been provided by any other census operation.

To understand the true impact of LUCA in the future, the Census Bureau should allow sufficient time for the completion of government updates prior to Block Canvassing activities. This would reduce the complexity of processing, as well as eliminate the need for another operation to validate updates.

Evaluation F.5, Block Canvassing Operation (Burcham 2002)

The Block Canvassing operation was one of the largest operations the Census Bureau conducted to update the MAF in preparation for Census 2000. It occurred in the winter/spring of 1999. The operation required field listers to conduct a 100 percent canvass of residential addresses in areas containing predominantly city-style addresses. Results from Block Canvassing were used to assign each housing unit to one of six basic action code categories: Verify, Add, Delete, Address Corrected, Geographic Corrections, and Add and Verify.

This evaluation quantified the impact of the Census 2000 Block Canvassing Operation on the MAF by profiling the housing units that Block Canvassing added to, deleted from, and corrected on the MAF. This evaluation did not provide a thorough comparison of Block Canvassing results to final census results or to the MAF building process as a whole.

Block Canvassing listers added 6,389,271 addresses to their listing pages. About 95 percent of the added units had city-style addresses. Based on preliminary results, Block Canvassing appears to have had a high level of geocoding accuracy. Over 94 percent of the adds showed a Block Canvassing block code equal to the official block code on the MAF. Around 29 percent of addresses added by Block Canvassing actually were on the MAF before Block Canvassing occurred but were either ungeocoded until Block Canvassing geocoded them, moved to different blocks by Block Canvassing, or considered nonresidential until Block Canvassing determined that they were residential units.

Block Canvassing listers deleted 5,146,320 addresses from their listing pages. The original source of an address is the first source that added the address to the MAF. In general, Block Canvassing deleted a larger proportion of addresses that had a newer original source than addresses with an older original source. For example, the 1990 Address Control File showed a lower percentage of deletes than the November 97 DSF.

Around 78 percent of the added units were valid housing units in Census 2000 and almost 24 percent of the deleted addresses were later enumerated as housing units in the census. About 96 percent of addresses coded as existing by Block Canvassing ended up as valid housing units in

the census. Also, 96 percent of all addresses sent to Block Canvassing to be verified showed consistent results between Block Canvassing and the census.

Block Canvassing provided a large number of updates to the MAF. Block Canvassing not only improved the coverage of addresses on the MAF, but also improved the geocoding of addresses on the MAF. Block Canvassing played a significant part in correcting unit designations in multi-unit basic street addresses. If the Block Canvassing listers had not checked individual addresses

within multi-units, but only verified the number of units at the multi-units, the MAF would not have this added improvement.

A relatively large number of Block Canvassing adds and deletes turned out to be inconsistent with final census results. However, the consistency between Block Canvassing and the census, as a whole, appears to be relatively good.

Recommendations resulting from this evaluation include:

- Continue to explore the possibility of targeting areas with certain characteristics as priority areas for updating the file. The clustering results in this evaluation should be a first step to showing how to target areas for MAF updates.
- Make additional efforts to see if quality review programs can reduce inconsistencies for added and deleted addresses.

Evaluation F.6, Evaluation of the Local Update of Census Addresses 99 (LUCA 99) (Owens 2002)

The Census Bureau conducted the Census 2000 LUCA 99 program in U/L and Update/Enumerate (U/E) areas from January 1999 to June 2000. This evaluation documented the results of the LUCA 99 operations.

There were 30,375 functioning governmental units eligible to participate in the LUCA 99 program. A total of 10,925 governments participated and they covered approximately 68 percent of the housing units in eligible areas. Although about 36 percent of all eligible governments participated, only 17 percent of eligible governments provided any updates in the form of adds, deletes, or corrections.

There were approximately 23,227,788 addresses from Address Listing (in the U.S. and Puerto Rico) that were geocoded with a map spot and eligible for review in the LUCA 99 program. The Census Bureau sent 2,186,765 addresses out for review to participating governments in the stateside LUCA 99 Recanvass operation. Field representatives verified that about 76 percent of them existed as residential units. They deleted approximately 6 percent of the addresses and determined that less than two tenths of a percent were nonresidential. They made corrections to the remaining 18 percent of addresses on their lists.

The Census Bureau sent a total of 35,563 addresses out for review in Puerto Rico. Field representatives verified that about 93 percent of them existed as residential units. They deleted approximately 7 percent and determined that less than one tenth of a percent were nonresidential. There were no corrected addresses in Puerto Rico.

Field representatives for the LUCA 99 Recanvass operation updated the address list and added any unit that existed as a residential unit in the block that was not already on the list. They added 328,174 addresses, which represents a 15 percent increase in housing units in U/L

enumeration areas in the U.S. (excluding Puerto Rico) that were recanvassed. Field representatives added 9,874 addresses in Puerto Rico, which represents about a 28 percent increase in housing units in areas that were recanvassed. Approximately 99.5 percent of LUCA 99 Recanvass adds in the U.S. and Puerto Rico were included on the initial census address list. About 85.2 percent of those adds were in the final census housing unit inventory.

LUCA 99 Recanvass field representatives deleted (or declared nonresidential) 145,378 addresses from their listing pages in the U.S. and 2,543 addresses in Puerto Rico. LUCA Recanvass field representatives corrected 388,838 addresses in the U.S. and Puerto Rico.

After participating local governments received feedback from the Census Bureau, they could appeal specific addresses. Participants appealed 18,442 addresses. Approximately 54 percent (10,053) of the addresses appealed by local governments were included on the final census address list.

The Census Bureau should continue to pursue LUCA type programs in non-city-style address areas for future censuses and tests. Also, the Census Bureau should investigate ways to increase government participation in LUCA programs.

Evaluation F.10, Evaluation of the Update/Leave Operation (Pennington 2003)

In Census 2000, U/L was intended for use in areas with some addresses that were not city-style. Noncity-style addresses, such as Rural Route and Box or P.O. Box, are often not linked to the physical location of the housing unit. When there is only a location description for a unit but no address, mail delivery of the questionnaire is not a possibility. U/L areas were primarily rural, but not too remote or sparsely populated. Designation of U/L areas was made by block. In Puerto Rico, U/L was the sole enumeration method. Questionnaires with preprinted address labels were hand-delivered to every housing unit on the U/L address list. Existing housing units that were not listed on the address register also required questionnaires, but these questionnaires were hand-addressed and added to the address register. Since staff were in the field delivering the questionnaires, they could also make other updates to the address list and to the maps during the U/L operation. This evaluation quantified the U/L operation as one means of assessing its effectiveness and value to the census-taking process.

There were 23,525,257 addresses in the stateside U/L operations and 1,471,225 in Puerto Rico. These numbers represent the addresses that had either a labeled questionnaire or a hand-addressed questionnaire. Questionnaires were to be distributed to all housing units appearing in U/L areas. Some of the addresses on the U/L address list were deleted as nonexistent or nonresidential in the U/L operation and the labeled questionnaires were not delivered. Deleted addresses are included in the workload calculation because it takes time and effort to try to locate such addresses.

Stateside U/L operations added 1,644,174 addresses, while 111,787 addresses were added during U/L in Puerto Rico. The number of corrections in stateside areas was 9,045,814, with 751,156 in Puerto Rico. Some number of these corrections were to the occupant name and telephone

number fields. These fields quickly became out of date. The number of deletes, either as nonexistent or as nonresidential, was 1,228,987 in stateside areas and 122,815 in Puerto Rico. Some units that were deleted in U/L were matched to U/L adds after processing, resulting in 24,265 moves, all of which were stateside. Units on the address list for U/L that did not receive any of these field actions were said to be verified. There were 11,582,017 verified addresses stateside and 485,467 verified addresses in Puerto Rico.

Not every address added in the U/L operation was included in the census. Some records were not included because they did not contain sufficient address information for adding to the address list or data sufficient to be assigned to a block. Other added records were found in subsequent operations to represent housing units that did not exist in the designated block, either because the unit was nonexistent or because the unit existed in another block. Of the 1,644,174 U/L adds in the U.S., 85.2 percent were in the final Census counts. In Puerto Rico, 83.7 percent of the 111,787 added addresses were included in the counts.

Recommendations resulting from this evaluation include:

- Research areas that could be converted to a mailout/mailback methodology. The analysis revealed large numbers of blocks in U/L areas that were wholly covered on the DSF.
- Assess the value of updates to the occupant name and telephone number fields.

Evaluation F.11, Urban Update/Leave (Rosenthal 2002a)

The Census Bureau conducted the Urban U/L operation from March 3 to March 31, 2000. The objective of the Urban U/L operation was to improve coverage by improving the deliverability of the questionnaires and updating address information and census maps. The Urban U/L operation targeted areas deemed unsuitable for Mailout/Mailback. Primarily, these areas included multi-unit buildings where the USPS delivers the mail to a drop point instead of individual unit designations and urban communities that had city-style addresses but many residents picked up their mail at a P.O. box. The Urban U/L operation relied on the local regions to identify areas based on their knowledge of whether the USPS could adequately deliver the census questionnaires. In Urban U/L areas, enumerators delivered the census questionnaires and updated address registers and census maps, concurrently. Residents were asked to complete and mail back their census questionnaires. Eight regions participated in Urban U/L: Atlanta, Boston, Chicago, Dallas, Denver, Detroit, Philadelphia, and Seattle. Twelve states and DC had Urban U/L areas.

Nationwide, 12,843 blocks were covered by Urban U/L. Almost 60 percent of these blocks contained housing units. The MAF had 314,059 residential addresses in Urban U/L blocks. After removing known duplicates, there were 310,114 addresses. Of the 310,114 addresses, 280,086 addresses, or 90.3 percent, were delivered to the DMAF. Ultimately, 238,216 addresses, or 85.1 percent of the DMAF addresses, were enumerated in the census as either occupied or vacant housing units.

Urban U/L contributed to the success of Census 2000 by improving the address list and successfully targeting hard-to-enumerate areas. Of the 267,005 addresses in the address registers, 18.1 percent were updated. There were 13,131 additions during questionnaire delivery, a 4.9 percent increase to the addresses printed in the address registers.

There were 2,114 blocks where 75 percent or fewer of the housing units in the block matched the DSF. These blocks contained 15.3 percent of the housing units in Urban U/L areas. Such blocks would presumably present mail delivery challenges for the USPS.

Less than 1 percent of Urban U/L housing units were drop delivery; that is, mail is delivered to a central location instead of to individual units of a multi-unit structure. While these addresses were included in Urban U/L, they did not make up a large part of the Urban U/L housing units in the census. Furthermore, the variable used to identify drop delivery status is not robust.

Close to one-quarter of the housing units in the census with hard-to-count scores were in the hardest hard-to-count class. The Planning Database provided a 1990 Census tract-level hard-to-count score, a composite measure of characteristics correlated with success in counting people. Evaluators classified each hard-to-count score into one of ten hard-to-count classes. Matching the Census 2000 census tracts to the Planning Database, 189,045 addresses, or 79.4 percent of the Urban U/L housing units in the census, were in census tracts that could be matched between the 1990 Census and Census 2000.

Persons under 18 years old, African Americans, and renters were over-represented in Urban U/L areas as compared to the nation. These traditionally undercounted persons were enumerated by mail at lower percentages than the average household or persons in Urban U/L areas.

Recommendations resulting from this evaluation include:

- Designate areas for Urban U/L based on headquarters' objectives supplemented by RO input instead of the current practice of the regions designating areas as Urban U/L.
- Conduct more field work or receive better USPS input to identify drop delivery status.
- Expand use of the Planning Database to target hard-to-count areas deemed suitable for Urban U/L.
- Consider enumerating traditionally undercounted areas using U/E instead of Urban U/L.

Evaluation F.12, Update/Enumerate (Rosenthal 2002b)

The U/E method targeted communities with special enumeration needs and areas where most housing units may not have had house number and street name mailing addresses. These areas included resort areas with high concentrations of seasonally vacant housing units, selected American Indian reservations, and colonias. In U/E areas, enumerators updated their address registers and census maps and enumerated the housing unit at the time of their visit. LCOs,

using general guidelines, designated areas for U/E. Every Regional Census Center (RCC) except Detroit had areas enumerated using the U/E methodology. Thirty-five states had U/E areas.

Nationwide, 183,889 blocks were covered by U/E and 75,827 of these blocks (41.2 percent) contained housing units. The MAF had 1,169,090 residential addresses in U/E blocks, after removing known duplicates. Of the 1,169,090 addresses, 90.4 percent were delivered to the DMAF. Ultimately, 956,214 U/E addresses (90.5 percent of the DMAF addresses) were enumerated in the census as either occupied or vacant housing units.

This evaluation looked at the extent of address updating, descriptive statistics of the addresses, demographic characteristics of the households and people living in U/E areas, and timing and cost of the operation.

The U/E operation contributed to the success of Census 2000. U/E helped the Census Bureau improve the address list and demonstrated that areas suited to field enumeration were identified.

Of the 926,861 addresses in the address registers, 37.2 percent were updated. The most frequent updates, corrections (change in the address), were made to 284,127 addresses. The remainder of the updates were nearly all deletions. There were 129,692 U/E additions during field enumeration, a 14.0 percent increase to the addresses printed in the address registers.

For 71.9 percent of blocks, no more than 25 percent of the housing units in the block matched the DSF. These blocks contained 60.6 percent of the U/E housing units. Such blocks would presumably present mail delivery challenges.

Of the addresses in the census, 15.2 percent had no address information; that is, the housing unit was missing the house number, street name, rural route, and P.O. box information.

The Planning Database provided a 1990 Census tract-level hard-to-count score, which is a composite measure of characteristics correlated with success in counting people. Evaluators classified each hard-to-count score into one of ten hard-to-count classes. Matching the Census 2000 tracts to the Planning Database, 59.2 percent of the U/E addresses in the census were in tracts that could be matched. While about one quarter of the addresses were in the top three hard-to-count classes and few addresses (0.6 percent) were in the bottom two hard-to-count classes, U/E was not limited to the most difficult hard-to-count classes. These results show that the Census Bureau followed the 1995 Census Test recommendation to not target U/E based on hard-to-enumerate criteria.

The higher-than-national enumeration rates of American Indians/Alaska Natives, Hispanics, and vacant housing units indicate successful targeting of areas with special enumeration needs.

The average household size in U/E areas was 2.9 persons, compared to 2.6 persons nationally. The U/E vacancy rate of 38.7 percent was higher than the national vacancy rate of 9.0 percent. Most vacants were seasonal vacants. Of occupied housing units, 76.1 percent were owned, compared to 66.2 percent nationally. Of persons, 49.6 percent were male, compared to 49.1

percent nationally. Of persons, 31.9 percent were under 18 years old, compared to 25.7 percent nationally. Of persons, 23.6 percent were Hispanic, compared to 12.5 percent nationally. Of persons, 1.5 percent were African American, compared to 12.3 percent nationally. Of persons, 27.7 percent were American Indian/Alaska Native, compared to 0.9 percent nationally.

U/E was conducted from March 13 to June 5, 2000, one week past the planned May 30, 2000 end date. The extra time was needed to enumerate an American Indian reservation at the bottom of the Grand Canyon, accessible only by mule.

Evaluation F.13, List/Enumerate (Zajac 2002)

List/Enumerate (L/E) was an operation used in sparsely populated areas of the country. Census enumerators were assigned areas to canvass and were given census maps for these areas. The enumerators were responsible for listing addresses within their area on blank address register pages, locating the addresses on census maps (map spotting), and conducting an interview to collect census information for each address. The operation, which included reinterview and field followup components, was carried out from mid-March 2000 to the beginning of July 2000. This evaluation examined the characteristics of addresses added to the MAF by L/E.

L/E was responsible for adding 392,368 addresses nationwide to the MAF. Of these addresses, 391,276 met the eligibility criteria to be in the census. This is about 99.7 percent of all added L/E addresses. Of the addresses eligible to be in the census, 389,749 addresses were actually included in the final census count. This represents 99.6 percent of the eligible L/E addresses and 99.3 percent of all added L/E addresses.

A total of 47,927 blocks had at least one L/E address. Of these blocks, 4.7 percent had all of their addresses recognized by the USPS. This indicates that these blocks could have possibly been converted to the Mailout/Mailback enumeration methodology. These blocks contained 1.4 percent of the addresses added during L/E.

L/E appears to be successful for the following reasons:

- Coverage: A total of 392,368 addresses were added from the operation.
- Future Locatability of Addresses: Over 50 percent (197,525 of the 392,368) were complete city-style type addresses. Of the 160,232 addresses that were not complete city-style or not complete rural route, 85.2 percent had location description information. In addition, 98.7 percent of all added L/E addresses had a valid map spot.
- Quality of Addresses: About 99.3 percent of the 392,368 addresses made it into the census.
- Targeting of Areas: Only 1.0 percent of the blocks in L/E had all of their addresses recognized by the USPS. These blocks represent just 1.4 percent of the addresses added during the operation.

Evaluation F.15, An Assessment of Addresses on the MAF "Missing" in the Census or Geocoded to the Wrong Collection Block (Ruhnke 2003)

One of the results of the Accuracy and Coverage Evaluation (A.C.E.) included a representative sample of addresses that were coded as "missing" from the census. This evaluation conducted additional research to better understand these "missing" addresses and to examine the reasons for their status of "missing" after the A.C.E. Final Housing Unit work was completed. Evaluators matched the addresses coded as "missing" to all non-duplicate housing units on the MAF in a larger geographic search area than the one used by the A.C.E. and searched for matches in the tract which included each address and in all surrounding tracts. The main focus in understanding these "missing" addresses was to determine if they were actually included in the census as housing units, but were incorrectly geocoded to a collection block outside of the A.C.E. geographic search area. Since matching was not limited to census addresses, but included all non-duplicate housing units on the MAF, evaluators were able to examine addresses that were on the MAF or the DMAF but were excluded from the census.

About 8,900 of the sample units coded as "missing" by the A.C.E. were matched to units on the MAF during this evaluation. About 4,800 of them were matched to addresses that were included in Census 2000. Of those census matches, about 3,100 were geocoded in error in the census to a collection block that was different than the block provided by the A.C.E. The other 1,700 units were matched to census addresses that were geocoded to the same block as the A.C.E. "missing" addresses. There are two primary reasons that these census units were not included in the census address list used for the A.C.E. address matching.

The first reason is that some of these units were identified as potential duplicates during the Census 2000 Housing Unit Unduplication operations and were therefore kept out of the A.C.E. Final Housing Unit matching operation. About 78 percent of the matches to in-census units in the same block were potential duplicates that ultimately were reinstated in the census. The remaining 22 percent of the in-census matches to A.C.E. "missing" units in the same block were not reinstated duplicates. A reason that these units were excluded from the address list used for the A.C.E. address matching is that they were not geocoded to an A.C.E. sample block at the time of the Final Housing Unit matching, but were moved into an A.C.E. sample block in time for evaluation work.

Of the approximately 8,900 sample addresses coded as "missing" by the A.C.E. that matched to the MAF in this evaluation, about 4,000 were not included in Census 2000. That is, these units were listed and confirmed as good, residential addresses during the A.C.E., but the Census Bureau's rules for creating the DMAF and the Hundred Percent Census Unedited File excluded them from the census. Those units represent a weighted estimate of 1.3 million units coded as erroneously excluded from the census as measured by the A.C.E. and this evaluation.

About 28 percent of the cases coded as erroneously excluded units were never delivered to the DMAF. There are a number of reasons units on the MAF would have not been sent to the DMAF as a result of the Census Bureau's rules for developing the Census 2000 address frame. One of the reasons a unit would not be included on the DMAF is if it was coded by the USPS as

nonresidential on the DSF. The Census Bureau excluded those addresses from the original census address list because it would not be prudent to mail questionnaires to all nonresidential addresses. The Census Bureau relied on field listing operations to add those units if they were actually residential units by Census Bureau definitions.

About 49 percent of the cases coded as erroneously excluded units were on the DMAF, but were deleted during the Kill Process. The goal of the Kill Process was to identify units that were most likely bad addresses (for example: a unit for which no census form was received and the unit was deleted in both the NRFU and CIFU operations) and remove them from the census.

About 22 percent of the cases coded as erroneously deleted units were on the DMAF but were determined to be potential duplicates during the Housing Unit Unduplication operations through address and person matching algorithms. The Census Bureau ultimately decided to exclude those units from Census 2000. The amount of erroneous deletions from the Unduplication operation as measured in this evaluation is potentially overstated. This comes from the fact that the A.C.E. may have coded something as missing from the census, when it was actually included in the census with a different form of the address. The Unduplication operation may have recognized the duplication but removed the version of the address that the A.C.E. listed.

The estimated percentage of census addresses that were geocoded to the incorrect Census 2000 collection block is 4.8 percent (standard error is 0.3 percent). The estimated percentage of geocoding error in the census was significantly higher in Mailout/Mailback enumeration areas (5.5 percent) than in U/L (1.7 percent) or L/E areas (1.2 percent).

Geocoding error was more prevalent among housing units in multi-unit structures. Housing units in both small and large multi-unit structures had a significantly higher geocoding error estimate than single units or housing units in two-unit structures. Additionally, large multi-units (housing units in structures with ten or more units) had a significantly higher geocoding error estimate than small multi-units (housing units in structures with three to nine units). The geocoding error estimate for both single housing units and two-unit structures was about 3 percent, for small multi-unit structures was about 5 percent, and for large multi-unit structures was about 11 percent. Geocoding errors are expected to be higher for units in multi-unit structures because geocoding error is a structure-based problem. Geocoding the structure to the wrong block causes every unit in that structure to be geocoded to the wrong block. The larger the structure, the larger the number of geocoding error cases there will be if the structure is geocoded to the incorrect block.

Geocoding error of census addresses was less frequent in certain regions of the country. The geocoding error estimate for the Midwest (3.8 percent) was significantly lower than the geocoding error estimate for the South (5.7 percent). There were no other significant differences. Geocoding error estimates also differed for some of the ROs. The Boston and Kansas City ROs both had a significantly lower geocoding error estimate than the national estimate of 4.8 percent.

One of the reasons addresses were coded as "missing" from the census during the A.C.E. was because they were incorrectly geocoded in the census to a collection block outside of the scope of the A.C.E.'s geographic search area.

Recommendations include:

- Research the possibility of collecting Global Positioning System coordinates for addresses in the census, to help enumerators find their assignments and to ensure geocoding units to the correct block. Consider getting better geocoding for the areas with the highest geocoding error rates, which are Mailout/Mailback areas. This research is currently underway.
- Conduct research to refine procedures for identifying and deleting units believed to be duplicates. The unduplication process appears to have deleted many units which should have been included in the census. Work has already begun on building an unduplication process into the 2010 Census.

Evaluation F.16, Evaluation of the Block Splitting Operation for Tabulation Purposes (Green and Rothhaas 2004)

This evaluation measured the percent of the country affected by collection blocks split for tabulation purposes and the accuracy of that block splitting. Collection blocks are geographic areas that are usually defined by visible features, and used by the Census Bureau to conduct field operations. Often, collection blocks cross governmental unit boundaries such as city and town or other required data tabulation boundaries. At the end of Census 2000, the Census Bureau redefined the census collection blocks by recognizing the boundaries of governmental units and other geographic entities required for tabulation of census data. One of the steps needed to achieve this involved using an automated system to split collection blocks in certain situations. This block splitting process was based on address ranges and map spot information in the TIGER database. To evaluate the block splitting process, evaluators selected a sample of 1,000 collection blocks that had at least one tabulation boundary that split the block for field verification. Field representatives determined whether the housing units in these blocks were allocated to the correct side or the wrong side of each tabulation boundary.

About 916,000 blocks out of the 5.1 million blocks in the country were split for tabulation purposes. A total of 282,457 blocks formed the sampling universe used to evaluate the block splitting process. The 633,337 split blocks excluded from the sampling universe were either located in remote Alaska, located in Puerto Rico, were split by the boundaries of special purpose governmental or administrative entities such as school districts, were split by the boundaries of statistical entities, or contained no housing units or group quarters. Remote Alaska and Puerto Rico were excluded from the evaluation to minimize cost. Boundaries of special purpose governmental, administrative, and statistical boundaries were excluded because this evaluation relied on the knowledge of residents of the block and they would not necessarily know where these types of boundaries existed in their blocks. Split blocks that contained no housing units or group quarters were excluded because the purpose of the evaluation was to measure the error

associated with placing housing units and group quarters in the wrong tabulation block. The estimated number of blocks that fell into each of these categories was not available, but the sum total was 633,337 blocks. A little more than 10 percent of the 115.5 million housing units in the country were located in the split collection blocks in the sampling universe.

Results showed that over 26 percent of these split collection blocks in the sampling universe contained at least one housing unit allocated to the wrong side of the tabulation boundary. Although this percentage is high, split collection blocks with at least one housing unit allocated to the wrong side of a tabulation boundary represent less than 2 percent of the collection blocks in the country. For housing units, about 3.65 percent of the 12 million housing units in the split collection blocks in the sampling universe were allocated to the wrong side of a tabulation boundary. These errors represent 0.37 percent of the housing units in the country. For the group quarters in the sample, none were allocated to the wrong side of a tabulation boundary. Although the estimate of the number of group quarters allocated to the wrong side of a tabulation boundary was zero, the Census Bureau cannot conclude that there were no group quarters in error throughout the country. The preliminary August 2002 results from an administrative program, in which the Census Bureau receives input from local governmental entities, showed that 1,867 group quarters in the country were in fact allocated to the wrong side of a tabulation boundary. This is less than 1 percent of all group quarters.

For this evaluation, collection blocks were categorized in the mailout/mailback, military, Urban U/L, and Urban U/E enumeration areas as “inside the blue line.” The term “inside the blue line” refers to areas where almost all mail delivery was to city-style addresses. Except for Remote Alaska, all other types of enumeration areas were categorized as “outside the blue line.” This term refers to areas where mail delivery was to noncity-style addresses. A mixture of city-style and noncity-style addresses occur in some types of enumeration areas, especially those “outside the blue line.” For the housing units affected by block splitting, the percent in error for enumeration areas “inside the blue line” was comparable to the percent in error for enumeration areas “outside the blue line”

For housing units affected by block splitting, fewer than 4 percent of the 11.1 million housing units with city-style addresses and fewer than 3.5 percent of the 773,000 housing units with noncity-style addresses were allocated to the wrong side of a tabulation boundary. The percent in error for both address types in the country was the same (0.37 percent).

Category G - Field Recruiting and Management

Summaries of the two parts of Evaluation G.1 in the Field Recruiting and Management category follow.

Evaluation G.1, Part 1, Census 2000 Staffing Programs, Recruiting Component (Jacobson and Petta 2002)

This evaluation described the factors that affected recruiting performance in 519 of the 520 LCOs during Census 2000.¹ The effect of several different types of factors were examined: census pay relative to locally prevailing pay; recruiting goals set by headquarters, area characteristics such as population density, private firm employment and per capita income, and manager's start date and turnover. This work was modeled on Westat's similar analysis of enumeration performance during the 1990 decennial census.

The analysis of factors affecting Census 2000 recruiting produced several important conclusions that were strongly supported by regression analysis, in keeping with reasonable expectations, and consistent with the observations of knowledgeable observers. The most basic findings were that there was considerable variation in recruiting performance across LCOs. Factors that accounted for about half of that variation fall into three categories: pay, area characteristics (such as the expected nonresponse workload and number of workers available to be recruited), and management factors (such as turnover of LCO managers).

Enumerator pay, relative to locally prevailing pay, was a key determinant of recruiting performance. The overall high levels of pay compared to 1990 greatly facilitated recruiting, but it was not possible to eliminate all cross-LCO variation in relative pay. Relative pay was about 77 percent of prevailing pay in LCOs with much below average recruiting performance and about 84 percent of prevailing pay in LCOs with much above average performance. That variation was associated with a 10 percent increase in the number of applicants in the LCOs with high versus low relative pay. This is an important finding since it suggests that pay increases could have been relied upon to further enhance recruiting performance had that been necessary.

In terms of area characteristics, the NRFU workload strongly influenced recruiting. On average, the NRFU workload was about 80,000 cases per LCO, but many LCOs had fewer than 56,000 cases and many had more than 104,000 cases. An increase in caseload of 24,000, about one standard deviation, was associated with a 13 percent increase in qualified applicants. This result suggests that LCOs appropriately attempted to secure enough applicants to fill the required number of enumerator slots, but other factors very strongly affected recruiting success.

The number of workers employed by all firms within a LCO area had a positive effect on recruiting, but the effect was small given that employment levels in LCOs showed huge

¹The Window Rock, Arizona LCO was omitted from the study because some critical data for this one LCO was missing.

variation. This suggests that there were plenty of individuals in virtually every LCO's area who could be attracted to apply for census jobs and it was not necessary to take variation in the number of people who can be drawn upon in allocating recruiting resources.

An unanticipated result was that a one standard deviation increase in test scores, an increase from 85.6 to 88.4, was associated with a decrease in the number of applicants of almost 11 percent. This result may have been due to recruiting being more difficult in areas where many people were likely to do well on the test (even holding relative wages and per capita income constant). The effect was sufficiently large that variation in test scores should be taken into account in setting wages and managers should anticipate that areas with high test scores will need additional assistance in meeting recruiting goals.

In terms of management factors, resignations, terminations for cause, or leaving for any other reason by LCO management during the recruiting period were associated with a reduction in the number of recruits by about 12 percent. This evidence strongly suggests that LCO management performance is a key determinant of recruiting success and that LCO management needs to be in place for a substantial period to be highly effective.

By far, the largest source of variation in recruiting performance was associated with a LCO being in one of three Census 2000 administrative regions. Even after taking other key factors into account, LCOs in the Seattle and Denver regions recruited about one-third more applicants and LCOs in the New York region recruited about 17 percent more applicants than LCOs in other regions.

Evaluation G.1, Part 2, Census 2000 Staffing Programs, Pay Component (Jacobson et al 2002)

This evaluation described analysis of the effect of pay, frontloading, and other factors on how quickly the Census 2000 NRFU was completed. From late April to late June of 2000, the houses of people who failed to return their census forms were visited by about 510,000 enumerators. The evaluation examined differences in how quickly the NRFU was carried out in the LCOs; differences in how quickly the NRFU was completed relative to the 1990 Census NRFU; and the underlying reasons for those differences. Overall, the Census Bureau's plan to raise wages to at least 75 percent of local levels and to put to work during the first week twice the number of enumerators that would be needed if there were no attrition, directly led to dramatic improvement in speed relative to the 1990 NRFU.

Hourly pay was increased by 37.8 percent on average relative to 1990 (adjusted for inflation) and the associated increase in enumerator retention was 22.6 percent. This increase in retention, coupled with introducing frontloading (increasing the number of enumerators at work at the outset relative to cases to complete), permitted the average Census 2000 LCO to complete the NRFU in 7.19 weeks compared to 9.72 weeks in 1990. Moreover, the slowest performing LCOs completed their work about 1.5 weeks faster than the fastest performing LCO in 1990.

Analysis of the variation in completion time across the 510 LCOs with adequate data (out of a total of 520 LCOs) shows that differences in the degree of frontloading was the primary source of variation in completion time, differences in the number of cases completed by individual enumerators played only a small role, and differences in retention of individual enumerators were too small to have much of an effect.

Analysis of the influence of factors within and outside of the control of the Census Bureau, using administrative databases covering the 510 LCOs plus a survey database covering close to 2,800 enumerators in 376 crews in 27 LCOs, showed that the NRFU was completed most rapidly in low-wage areas and areas where applicants' test scores were low on average, in the Denver and Los Angeles Census Regions where managers ensured that high levels of frontloading were achieved, and in LCOs that had fewer cases to complete (relative to larger scale offices) and in which managers did not turn over.

The first overall conclusion from this analysis was that differences in factors outside of the control of census managers, such as the labor force and area characteristics, had small effects on completion time and productivity. In contrast, factors largely within census management control, such as the total number of LCOs, the number of cases to complete within a given LCO, census pay levels, and Regional Office (RO) planning and oversight, had large effects on performance. Where the basic pay, recruiting, and frontloading plans were followed, LCOs succeeded in securing and retaining more than enough applicants to staff the NRFU with highly competent enumerators who also were strongly motivated to work as long as needed.

The second, but single most important, conclusion was that the degree to which LCOs exceeded schedules was largely a function of the amount of frontloading they achieved. About 80 percent of the LCOs met or exceeded frontloading goals. However, the roughly 20 percent of the LCOs that did not meet their frontloading goals took about two additional weeks to reach their week 1 goals. Understanding why frontloading goals were not met, therefore, is the key to understanding the source of variation in speed.

One or more of the following three hiring explanations probably led to recruiting shortfalls: hiring was inherently more difficult due to factors outside of census management control, hiring was not effectively managed, and managers did not feel it was essential to meet frontloading goals. Unfortunately, the data needed to definitively sort out the relative importance of the three explanations were lacking

The third, but somewhat speculative, conclusion was that improvements in the hiring process were needed to meet frontloading goals. Possible improvements include starting the hiring process earlier and ensuring enough hiring clerks and phone lines are available to offset unexpected hiring difficulties. Thus, much as frontloading of enumerators was the key to dramatically increasing the speed in conducting the enumeration, increasing hiring capacity appears to be the key to meeting frontloading goals.

A fourth conclusion was that setting pay competitively was essential to recruiting sufficient numbers of well-qualified applicants and to retaining enumerators as long as they were needed.

These results suggest that census pay exceeded the threshold above which people who agreed to accept enumerator positions were sufficiently competent to execute the work and would not lightly break their commitments to work while their services were required. About 90 percent of the Census 2000 enumerators showed themselves to be highly productive, as measured by the number of cases they were able to complete per hour. In contrast, during the 1990 Census NRFU, 50 percent or more of the enumerators had difficulty completing assignments and/or quit before completing even their initial assignment.

The sharp contrast between pay and performance in 2000 versus 1990 has several important implications. Perhaps the most important is that the Census Bureau should reassess how test scores and availability to work many hours are used as hiring screens. The analysis suggests that the capacity to quickly complete the NRFU would have been enhanced had test scores of about 82 percent been used as a threshold (unless applicants had some special skill such as fluency in a foreign language) and the contact order been based on hours of availability (reported in applications).

The equations produced in this evaluation could be extended to set the schedule and the degree of frontloading for the 2010 Census NRFU in a way that would substantially reduce cost without reducing the probability the schedule is met.

Category H - Field Operations

Summaries of the seven evaluations in the Field Operations category follow.

Evaluation H.2, Assessment of Field Verification (Tenebaum 2001)

This operational assessment focused on the Be Counted/TQA Field Verification operation. Enumerators visited the location of units without a confirmed census address to verify their existence before Census 2000 included the addresses. The operation also included addresses deleted in two or more previous operations but for which the Census Bureau received a mail return (double deletes with a mail return). If the enumerator located the address, he/she entered a checkmark on the assignment listing for each unit verified as a residential address. If the address was not a living quarters or was a duplicate of another address on the assignment listing, the enumerator coded it accordingly.

The non-ID questionnaire process for Census 2000 was a very complex operation consisting of many components. This operational assessment only discussed one aspect of that overall process, that is, the verification of addresses which could not be matched to the MAF but could be geocoded to a census block. No conclusions can be made regarding any other component of the non-ID questionnaire process.

Enumerators provided information about the 884,896 assigned addresses. They coded 51 percent of the assigned addresses as valid living quarters, 35 percent as nonexistent, and 14 percent as duplicates.

Overall, 49.18 percent of the addresses without a confirmed census address (the non-ID cases) were coded as valid census addresses. Overall, 52.86 percent of the addresses deleted in two or more previous operations but for which the Census Bureau received a mail return (the double deletes) were coded as valid addresses.

The Census 2000 procedures provided a good model for planning a field verification operation for Census 2010, with the following recommendations:

- Redesign the Field Verification procedures to capture enough information for duplicates to provide a link between the two addresses.
- Clarify the procedures concerning how far to extend the search for assigned addresses so enumerators do not erroneously delete addresses located in adjacent blocks.
- Conduct further research into the sources of the double deletes with mail returns since about half of them were coded as valid addresses.
- Consider a way to independently validate the accuracy of the results to determine if Field Verification improves the census files.
- Determine the effect that additional response options in 2010 might have on Field Verification.

Evaluation H.4, Questionnaire Assistance Centers for Census 2000 (Jones and Barrett 2003)

Questionnaire Assistance Centers were targeted locations designed to assist individuals who had questions about completing their Census questionnaires, who needed language assistance on their questionnaires, who had a general question about the census, or who never received a census questionnaire. These centers were open between March 8, 2000 and April 14, 2000.

Census Bureau Partnership Specialists, in consultation with local officials played an important role in selecting the census tracts where Questionnaire Assistance Centers were placed. Most of the tracts chosen to have Questionnaire Assistance Centers were in areas known to be either difficult to enumerate, heavily populated by certain racial and ethnic groups, or in linguistically isolated areas known to be heavily populated by speakers of certain foreign languages. Publicly accessible locations such as community centers and social service centers were set up to house Questionnaire Assistance Centers. The centers were staffed by paid clerks and volunteers. Some of the paid clerks had foreign language skills, so they could provide expert assistance to potential census respondents experiencing language difficulties. Volunteers were chosen from local community groups or other organizations that were in partnership with the Census Bureau. Both paid and unpaid staff provided literacy assistance to those respondents in need of it. Staff were instructed to complete a Record of Contact for each potential census respondent that visited the center. The Record of Contact documented the type and extent of assistance needed.

There were 23,556 Questionnaire Assistance Centers established during Census 2000. However, data were collected and processed from only 14,222. There were no data processed from the remaining centers. Data were keyed for 559,027 potential census respondents that utilized the Questionnaire Assistance Centers during Census 2000. About 39.4 percent of these respondents were provided with a Be Counted Form. Some respondents (26.4 percent) needed assistance on a specific type of questionnaire. Of those who did need assistance, most asked for help in completing the short form. Census forms were printed in six different languages: English, Spanish, Chinese, Korean, Vietnamese, and Tagalog. Of the people who needed assistance on a specific questionnaire, most (64.6 percent) required assistance on the English short form.

Respondents generally did not ask for assistance on specific questions or specific census forms. Only a few respondents (2.3 percent) asked for assistance with a population or housing question on their census form. Only 5.8 percent of respondents indicated needing assistance with language, but another 4.6 percent marked that they could not read or understand the form. Respondents reported other questionnaire related problems, such as respondent did not receive a census form, respondent lost their mailed census form, or respondent had a problem with their questionnaire that was not listed on the Record of Contact .

There were numerous ways that people learned about the Questionnaire Assistance Centers. About 32.6 percent of the respondents learned of the centers in ways unlisted on the Record of Contact. Of the alternatives listed on the form, 15.5 percent reported having seen a poster announcement.

Language Assistance Guides were available in 37 different languages. Most respondents (83.1 percent) did not request a Language Assistance Guide. Of the 94,639 people needing Language Assistance Guides, more than half (53.0 percent) needed them in Spanish. Be Counted Forms were available in six different languages, including English. Most respondents (60.6 percent) did not request a Be Counted Form. Of those who did (220,489 people), most requested them in English (69.8 percent) or Spanish (24.4 percent).

The Census Bureau should continue to establish Questionnaire Assistance Centers during a census to help respondents. However, the following changes are recommended:

- Increase the number of languages in which the Census Bureau provides the Be Counted Form. More than 1,000 respondents requested Language Assistance Guides in languages for which a Be Counted Form was not available.
- Collect demographic information about the census respondents that utilize Questionnaire Assistance Centers. This would make it easier to tell if the centers served the targeted population.

Evaluation H.5, Nonresponse Followup for Census 2000 (Moul 2002)

The objective of NRFU was to obtain completed questionnaires from households in mailback areas that did not respond by mail, through the Internet, or to a TQA operator. If a questionnaire was not checked in before the NRFU universe selection process began, the housing unit was targeted for NRFU. There were 119,090,016 housing units in mailback areas (including Puerto Rico) that were potentially eligible for followup. The initial workload of 44,928,883 housing units was identified on a flow basis. A Late Mail Return operation subsequently identified 2,555,918 housing units that were checked in after the initial universe was identified. The resulting workload, which included Puerto Rico, was 42,372,965 (35.6 percent of the eligible universe). The NRFU operation was scheduled to occur from April 27 through July 7, 2000. Actual start and finish dates were April 27, 2000 and June 26, 2000, respectively.

Based on the following, NRFU was a success.

- NRFU officially ended early on June 26, 2000, ten days ahead of schedule.
- Less than 0.1 percent of the workload had an undetermined status at the end of NRFU.
- Compared to the 5.0 percent target, there was a low final attempt rate (approximately 2.7 percent).

However, the NRFU operation was not perfect. For example:

- For 5.4 percent of the returns, enumerators failed to indicate whether the interview was with a household member or a proxy.
- Of the 26.4 million occupied housing units, 0.4 percent had no population count.
- Approximately 4.2 million housing units were enumerated multiple times. Approximately 3.5 million of these were enumerated in NRFU and by a paper mail return questionnaire.
- Some housing units had an unrealistically large number of continuation forms attached, as many as 99.

Of the 42.4 million housing units, 62.3 percent were occupied, 23.3 percent were vacant, and 14.3 percent were deleted. Approximately 79.6 percent of the occupied returns were completed by a household member and 16.5 percent of the occupied returns were completed by a proxy.

There were 1,255,579 continuation forms used in NRFU. Approximately 93.6 percent of the returns had one continuation form attached, approximately 2.9 percent of the returns had two continuation forms attached, and fewer than 1 percent of the returns had three or more continuation forms attached. The number of continuation forms attached ranged from one form to as many as 99 forms.

NRFU enumerated a higher percentage of multi-units and rented units than were self-enumerated. NRFU also enumerated a higher percentage of males, young people, Hispanics, and people of all races except Whites.

There were 688,944 addresses added and 6,023,232 addresses deleted during NRFU. The majority of the added and deleted addresses were single units in the mailout/mailback areas. The adds and deletes were mostly complete city-style addresses.

Recommendations include:

- Monitor the followup workload in real-time to reduce the number of NRFU cases with unknown population counts and the number of lost NRFU enumerator returns.
- Periodically identify and remove late mail returns from the NRFU workload to reduce the NRFU workload and the number of housing units with multiple data captures.
- Implement a sufficient quality assurance (QA) program to ensure the accuracy of the NRFU production files and the proper use of enumeration techniques.
- Develop standards/benchmarks with which to measure/judge the results.

Evaluation H.7, Nonresponse Followup (NRFU) Enumerator Training (Burt and Mangaroo 2003)

The success of the NRFU operation was highly dependent on the Census Bureau's ability to quickly develop skilled employees who were able to effectively perform the tasks of NRFU enumeration. During Census 2000, the Census Bureau hired more than one-half million temporary workers to conduct the NRFU operation. This evaluation examined the effectiveness of the Census 2000 NRFU enumerator training program. The evaluation used the Kirkpatrick training assessment model to evaluate the trainees' satisfaction with the training program, their knowledge following training, and their on-the-job performance.² The methodology included a content review of the training materials, observation reports on training delivery and NRFU enumeration, and surveys and debriefings of enumerators and crewleaders.

About half of the recommendations from the Census 2000 Dress Rehearsal evaluation were incorporated into the 2000 training package, either completely or partially. The recommendations that were incorporated contributed to an improved training program. The recommendations that were not incorporated did not seem to significantly impact the effectiveness of the training in preparing the enumerators to collect Census information.

²Kirkpatrick, Donald L. *Evaluating Training Programs: The Four Levels*, San Francisco, CA: Berrett-Koehler, 1998.

Most enumerators were knowledgeable of census concepts. With the exception of reading questions as worded, enumerators consistently demonstrated effective interviewing skills.

Evaluation results indicated that the training prepared the NRFU enumerators to effectively perform their job and to carry out the tasks they were trained to do. Almost all of the enumerators displayed their Census ID at each household, properly identified themselves and the purpose of their visit, and determined Census Day residency status. Most consistently confirmed that they were at the correct address and provided a Privacy Act Notice. The majority of NRFU enumerators recorded answers accurately and legibly. A sizable proportion of NRFU enumerators, however, did not always read the questions exactly as worded and frequently did not use the flashcards provided. The two major areas in which enumerators seemed less likely to follow procedures were asking about Hispanic origin and race.

Overall, enumerators were satisfied with the amount and content of the training they received.

The evaluation results indicated that the Census 2000 NRFU training program was well received by trainees and produced enumerators who could effectively collect needed Census data. The results also suggested some areas the Census Bureau could focus on in developing NRFU enumerator training for 2010.

- Increase the training time allotted to areas in which enumerators' feedback indicated they felt less well prepared, with particular emphasis on interacting with reluctant respondents and refusals.
- Continue to place emphasis on reading all of the questions exactly as worded, adding additional explanations on why reading questions verbatim is so important to data quality.
- Conduct debriefings of enumerator and crewleaders in the 2004 Census Test to get insight on how to improve the use of flashcards and other job aids and on how to make improvements on preparing enumerators to ask the ethnicity and race questions exactly as worded.
- Continue to provide an opportunity for the field work component of NRFU training and enforce inclusion of field work in all training sessions.
- Restructure the NRFU enumerator manual to be more consistent with the training guide in terms of organization and content.
- Add to the training materials a "Frequently Asked Questions" job aid, outlining potentially difficult questions that respondents might ask and appropriate responses.
- Use additional media such as audio tapes, videos, flip charts, posters, and slides in training and assure they are used.

- Increase the use of role playing, varying the situations to include reluctant respondents and refusals.
- Use a trainer and training assistant to model interviewing skills, techniques, and styles.

Evaluation H.8, Operational Analysis of Enumeration of Puerto Rico (McNally 2003)

Census 2000 was the first time that an U/L mailback methodology was used to conduct the enumeration in Puerto Rico. Stateside U/L occurred in areas that were determined to contain some proportion of Rural Route, P.O. Box, or other non-city-style addresses. These areas were primarily rural, but not too remote or sparsely populated. In Puerto Rico, U/L was the only type of enumeration.

During the U/L operation, questionnaires with preprinted address labels were hand-delivered to every housing unit on the address list. Existing housing units that were not listed on the address register also required questionnaires, but these questionnaires were hand-addressed and the addresses were added to the address register. Staff delivering the questionnaires also made corrections, deletions, and additions to the address lists and maps.

Working with maps and map spots was reported as the most challenging situation due to using rural procedures in an urban location. Other problems encountered were training manuals arrived late, merging long and short form questionnaires was a very time consuming process, and the maps for NRFU and CIFU were not updated from previous operations.

There were almost 1.5 million addresses in Puerto Rico. This number represents how many addresses had either a labeled questionnaire that was to be delivered during U/L or a hand-addressed questionnaire for a unit that was added to the address list during the U/L operation. Some of the addresses on the address list were deleted as nonexistent or nonresidential during the U/L operation. This is included as part of the workload because the effort required to try to locate such housing units is a component of the operation.

Adds accounted for 7.6 percent of the workload in Puerto Rico. Of the 111,787 U/L adds in Puerto Rico, 83.7 percent were included in the final counts. Deletes accounted for almost 8.4 percent of the workload in Puerto Rico.

As of April 18, 2000 the response rate for Puerto Rico was 48.4 percent compared to 59.3 percent for stateside. Sixty-four percent of the households in Puerto Rico returned their Census 2000 questionnaire by mail, a low return rate when compared with the stateside rate of 77.9 percent. Many questionnaires that were mailed back were not recorded in the system and became part of the NRFU workload. This led to multiple, unnecessary visits to households.

Based on the Census 2000 experience, recommendations include:

- Build on experience from Census 2000 and the ACS in Puerto Rico to use the mailout/mailback data collection methodology for at least part of Puerto Rico.

- Conduct further research into ways to improve census maps to make them easier for enumerators to use and to improve the accuracy of map spotting.
- Ensure that training materials and training manuals arrive on time.

Evaluation H.9, Local Census Office Profile for Census 2000 (Imel 2003)

The LCO Profile for Census 2000 is a database of a wide variety of variables aggregated at the LCO level. Data were produced specifically for the LCO database, which also contains data from Census 2000 evaluations. The LCO profile contains data from the following types of files and systems: address frame files, data processing files, and field tracking systems. It contains more than 1400 statistics for each of the 520 LCOs. These data include total housing counts, occupied housing counts, vacant housing counts, total population counts, response rates, various operation summaries, as well as many demographic characteristics. The database is intended to provide internal Census Bureau researchers and planners a rich, comprehensive dataset which will add a historical perspective and provide benchmarks for planning the 2010 Census.

Evaluation H.10, Date of Reference for Age and Birth Date used by Respondents of Census 2000 (Carter 2002b)

This evaluation examined how well respondents answered the census as of Census Day, April 1, 2000, by looking at how respondents answered the age and date of birth question. The way respondents answer this question can be influenced by whether or not they are using Census Day as their date of reference.

A state's return rate seems to be correlated with the date of reference for that state. As the return rate increases, the date of reference for the state is closer to April 1, 2000. A higher return rate in a state means more respondents are returning their questionnaires through the mail. It is also very likely that these respondents will not be part of NRFU and they are enumerated closer to April 1, thus less likely to misreport their age. If the return rate is low, that would mean a higher percentage of people are being enumerated in NRFU. NRFU takes place at a later date, so the respondents enumerated in NRFU seem to have a greater propensity to use a date other than Census Day to report their age.

The analysis showed that 89.8 percent of persons had their reported age consistent with their calculated age. There were 1.8 percent that under reported their age by one year and 6.0 percent that over reported their age by one year. These people may have potentially misreported their age due to using some date other than April 1, 2000 as the date of reference when reporting their age. The remaining 2.4 percent misreported their age by more than one year, which means the misreporting can only be attributed to simple reporting error.

There were two situations where respondents may have had problems reporting age correctly; the date of check-in of the form with the person's data was before the person's birthday and both were before April 1, and the date of check-in of the form with the person's data was after the person's birthday and both were after April 1. In the first situation, 10.3 percent of the persons

in this category under reported their age. In the second situation, 40.1 percent of the persons in this category over reported their age. These percentages are higher than any percent observed in any of the other situations for that type of misreporting. This means that the time at which a person is responding to the census does affect how he or she reports age.

There were 80.4 percent of households that had every person in them with the age correctly reported. This compares to 89.8 percent of persons with age correctly reported.

The census moment or 'average' date of reference moved from May 5 in 1990 to April 20 in 2000. This improvement may be due to the change in questionnaire design and in the enumeration time frame. The 2010 Census questionnaire should stress that the respondents are to provide their age as of Census Day, April 1, as in Census 2000. This will help respondents not misreport age. Also, a compressed Census enumeration time frame may aid respondents to correctly report age.

Respondents enumerated by personal visit tended to over report age. Enumerators should have this problem explained to them and training should stress the importance of Census Day as the reference date. Enumerators should know that respondents need to hear April 1, 2010, so they can correctly provide their information.

The problems that are observed in age reporting revealed problems with respondents referencing April 1 when providing age date. This is somewhat trivial because age can be calculated from date of birth. There are other issues that are sensitive to the April 1 reference day, such as Residency Rules, that cannot be corrected.

Category I - Coverage Improvement

Summaries of the five evaluations in the Coverage Improvement category follow.

Evaluation I.1, Coverage Edit Followup (Sheppard 2003)

The CEFU operation for Census 2000 was used to increase within household coverage and improve data quality by collecting person data for all persons beyond the first six in large households and by resolving count discrepancies between the reported household population count and the actual number of data defined persons recorded on the census form. The purpose of this evaluation was to document the CEFU operation.

While the coverage edit failure workload was projected to be over 3.1 million, the actual workload included 2,544,072 CEFU cases from Census 2000 mailback and Internet forms. Large household cases made up almost 55 percent of the CEFU cases (1,395,623). The edit failure rate for large household cases was 1.7 percent. Count discrepancy cases made up the rest (1,148,449). The edit failure rate for count discrepancy cases was 1.4 percent. Just over 60 percent (699,379 cases) of the count discrepancy cases were selected because the number of data defined persons on the form exceeded the respondent-reported household size. The rest of the

count discrepancy cases were selected because the number of data defined persons on the form was less than the respondent-reported household size (449,070 cases).

There were 1,251,971 cases completed during the CEFU operation. This was 53.5 percent of all the eligible and attempted cases. The Census Bureau were more successful completing large household cases (57.4 percent) than count discrepancy cases (48.5 percent). The largest reason for incomplete cases was inability to contact the respondent by telephone. The Census Bureau made two attempts to obtain telephone numbers for cases where one was not correct or present on the mailback form. However, only 21.1 percent of the cases with changed telephone numbers were completed. Since there was no field followup, the Census Bureau was unable to complete any of the cases where valid telephone numbers could not be obtained. These represented 24.0 percent of all eligible CEFU cases.

The CEFU instrument was effective in its two main objectives: correcting incorrect rosters and collecting person data. The CEFU operation successfully improved coverage and decreased the differential undercount. In 232,777 cases, or 18.6 percent of all completed coverage edit cases, one or more persons were added, deleted, or removed as a duplicate. A total of 410,565 persons were added, deleted, or marked as duplicates to correct the roster of a household. The 152,683 persons who were added to the household roster were more likely to be members of traditionally undercounted populations than persons in the overall population enumerated in Census 2000. These persons were much more likely to be under 24, be of a race other than White, and to be Hispanic. They were slightly more likely to be 65 years old or older, be male, and have the householder be an owner. There were 257,882 persons who were deleted or removed as duplicates from the household roster during the operation. These persons were much more likely to be between 15 and 24 or over 65 years old and to be Black than persons in the overall population enumerated in Census 2000. They were slightly more likely to be of Hispanic origin, be female, and have the householder be an owner.

The CEFU operation actually resulted in a net loss of 105,199 persons compared to the originally completed Census 2000 self-response forms. However, while the net improvement to the census from CEFU operation was a decrease in the population, it did improve the accuracy of Census 2000. Through the probing interview, the CEFU increased the likelihood that the 410,565 people who were added, deleted, or marked as duplicates were counted in the correct household.

Given the results and limitations of the data, the following are some recommendations for the CEFU operation in Census 2010:

- Continue to conduct a CEFU operation in future censuses. Include count discrepancy cases and large household cases, as well as other cases identified as having a significant possibility of coverage problems.
- Develop ways to increase the completion rate for CEFU operations.

- Improve case file creation, management, software testing, and transmittal procedures of input and output files to avoid loss of data and to ensure information is available to conduct interviews as planned.
- Improve the design of the CEFU instrument to improve effectiveness and reduce respondent burden.
- Collect evaluation data in future tests of coverage followup operations to help improve the methodology used to conduct followup interviews.
- Assign the final household size for count discrepancy cases not completed during CEFU by assuming that the cases that are not completed behave similarly to those that are completed.

Evaluation I.2, Evaluation of Nonresponse Followup - Whole Household Usual Home Elsewhere Probe (Viator and Alberti 2003)

The Whole Household Usual Home Elsewhere (WHUHE) probe was a questionnaire coverage improvement used to determine if all members of a household on the day of the interview had another residence where they lived most of the time (their Census Day address). This probe was accomplished by implementing a set of screening questions on the Simplified Enumerator Questionnaire. In cases where a household indicated that all household members had another residence where they lived most of the time, enumerators completed a blank, unlabeled Simplified Enumerator Questionnaire for their Census Day address or “usual residence.” This questionnaire was used to ensure a complete and accurate enumeration at the address of the usual residence.

A total of 151,775 questionnaires were completed for WHUHE households for their usual place of residence. Of these returns, 58,027 matched to an existing address on the DMAF. Another 55,286 returns were geocoded but not matched to an existing address. Of these 55,286 returns, only 606 were geocoded in time to be sent to the Field Verification operation. Most of the remaining 54,680 returns were added to the census but not included in the Field Verification operation. For the 606 returns that were geocoded and sent to Field Verification, 273 were verified, 271 were deleted, 59 were duplicates, and results were not reported for three returns. Finally, 38,462 returns could not be geocoded or matched to an existing address.

There were 55,987 WHUHE addresses that were enumerated by another procedure. More than 14 percent of these were reported to be vacant or nonexistent (delete) on Census Day. There were 54,915 addresses that were not enumerated by another procedure. The respondent for nearly 62 percent of these addresses was a neighbor or other proxy compared with only about 32 percent for addresses enumerated by another procedure. More than 71 percent of the addresses not enumerated by another procedure were reported as vacant. For more than 76 percent of these vacant addresses the respondent was a neighbor or other proxy.

There were 113,991 data defined persons enumerated on occupied WHUHE forms. Of these persons, 75,254 were found on other census returns at the address and 38,737 were not found on other census returns at the address. Of the 38,737 persons who were not found on other returns, 29,302 were selected by the Primary Selection Algorithm for inclusion in the census. These are persons who were not enumerated by other operations and were added to the census by the WHUHE program.

There appeared to be considerable confusion among enumerators concerning how to enumerate the WHUHE households. Many of the usual residences for these households were reported as vacant or deleted housing units. This is contrary to the concept of a usual home elsewhere for the Census Day household. Many of the questionnaires completed for the usual home of the Census Day household were completed by a respondent who is not a member of the household. Anecdotal evidence shows many instances of multiple questionnaires filled by the same proxy respondent.

Frequently, there were missing and inconsistent data for responses to the introductory questions used to identify WHUHE households. Because there was no way to link a census return for the address on the day of the interview with the return for the WHUHE address, the missing and inconsistent responses prevented the accurate identification of responses that should have generated a WHUHE return.

Recommendations include:

- Take into consideration the small number of persons added to the census by the WHUHE probe in judging the potential effectiveness of this program for the 2010 Census.
- Research whether or not it is a sound practice to add respondent provided addresses, such as the WHUHE addresses, without verifying them through a Field Verification operation.
- Cover the purpose of and the procedures for the WHUHE program thoroughly in enumerator training to ensure a better understanding of the program and higher quality data.
- Address the treatment of proxy responses about usual home elsewhere information in future censuses.
- Incorporate edits into future interactive electronic enumeration devices to detect and correct inconsistent data problems as the data are being collected.
- Create a mechanism to link each WHUHE return with the census return that generated it to design new QA processes.

Evaluation I.3, Evaluation of Nonresponse Followup - Mover Probe (Keathley 2004)

All enumerator questionnaires contained a mover probe. The mover probe allowed enumerators to identify households that moved into NRFU or CIFU housing units *after* April 1, 2000 and did not return a census questionnaire for their census-day address (nonresponse in-mover households). Enumerators in both operations would then attempt to complete a separate enumerator questionnaire for every nonresponse in-mover household for the address at which it lived on census day. These enumerator questionnaires represented the nonresponse in-mover households' completed census questionnaires. The purpose of this evaluation was to determine the usefulness of the mover probe.

Of the 105,480,101 occupied housing units in the U.S. (the 50 states and DC) enumerated in Census 2000, 22,850 would not have been enumerated without the mover probe. This represents 0.02 percent of the total U.S. occupied housing unit count.

There were a total of 45,507,823 enumerator questionnaires from the two followup operations. The 22,850 enumerator questionnaires representing occupied housing units enumerated only by the mover probe represents 0.05 percent of all NRFU and CIFU questionnaires.

There were 125,585 enumerator questionnaires that had information indicating that they represented nonresponse in-mover households; only 18.19 percent (22,850) of these represented households that Census 2000 would not have otherwise enumerated without the mover probe.

Recommendations include:

- Reevaluate the mover probe in future census and test census operations.
- Using the mover probe improved the accuracy of the housing unit count (and other related statistics).
- The cost of keeping the mover probe in future census operations might be negligible.
- Computerized versions of future enumerator questionnaires could improve the effectiveness of the mover-probe.
- Develop a way to link separate enumerator questionnaires for the same followup operation housing unit (one enumerator questionnaire representing the followup operation housing unit, the other representing the housing unit at which a nonresponse in-mover household resided on census day).

Evaluation I.4, Coverage Improvement Followup (Moul 2003)

Coverage Improvement Followup (CIFU) was designed to improve coverage of housing units in the mailout/mailback, U/L, and Urban U/L areas. There were 121,894,831 housing units in these mailback areas that were potentially eligible for CIFU. The workload (including Puerto Rico) consisted of 8,854,304 housing units. Most of this workload consisted of units classified as vacant or delete in NRFU. The workload also included units that were identified as vacant or delete in two census operations, units identified as seasonal vacants, units identified as UAA, adds from the new construction operation, adds from the U/L and Urban U/L operations, blank mail returns, lost mail returns, nonrespondents in Panels 7, 8, and 9 of the Response Mode and Incentive Experiment, February 2000 and April 2000 DSF adds, adds from the LUCA 1998 and 1999 Appeals process, Hialeah, Florida NRFU units, POP99s (housing units identified as occupied during NRFU that did not have a population count), and Residual NRFU units.

The CIFU operation was conducted in three separate waves as groups of LCOs completed NRFU. Wave 1 included 342 LCOs and began on June 26 and finished on July 26. Wave 2 included 175 LCOs and began on July 10 and finished on August 10. Wave 3 included three LCOs and began on July 30 and finished on August 23.

This evaluation developed a profile of the CIFU units to provide Census Managers with critical information needed for planning the 2010 Census.

The CIFU operation followed up 3.9 million vacant units and 2.6 million delete units. Approximately 21.9 percent of the vacant units were converted to occupied and 24.6 percent of the deletes were converted to occupied. These converted units resulted in a net gain of approximately 3.1 million people. Approximately 18.1 percent of the deletes were converted to vacant.

In the 1990 field followup operation Vacant/Delete/Movers Check, approximately 8.7 percent of the vacants were converted to occupied and 6.4 percent of the deletes were converted to occupied; approximately 5.3 percent of the deletes were converted to vacant. Compared to Census 2000, the 1990 vacant and delete workloads were larger and the conversion rates were lower. These differences were the result of changes in the universe rules for inclusion (i.e., there were different rules for including/excluding vacant and deleted units).

At the end of CIFU, approximately 26.8 percent of the units were occupied, 43.4 percent were vacant, and 29.7 percent were deletes and only 542 of the 8.9 million housing units had an undetermined status at the end of the operation. More than 88 percent of the lost mail returns and 81.2 percent of the blank mail returns yielded valid housing units. Approximately 52.9 percent of the new construction adds and 58.5 percent of the DSF adds were deleted and approximately 63.6 percent of the LUCA Appeals adds were ultimately deleted, which confirms the findings in earlier operations that these addresses were not valid addresses.

There were 5.3 million people enumerated in CIFU. Like the NRFU operation, CIFU was successful in enumerating a higher percentage of the groups that are typically undercounted: males, young people (34 years old and younger), Hispanics, Blacks, and Some Other Race.

Recommendations include:

- Continue to improve coverage by following up vacant and deleted units from NRFU.
- Continue to follow up any lost or blank mail returns. Consider adding a “vacant” option to the mailback questionnaire so that respondents could indicate the unit was vacant on Census Day so that valuable resources are not wasted following up legitimate blank returns.
- Investigate ways to improve/screen the data received from local governments to avoid spending time and money following up invalid/bad data.

Evaluation I.5, Coverage Gains From Coverage Questions C1 and C2 on Enumerator-Completed Questionnaires for Census 2000 (Zelenak and Nguyen 2003)

The focus of this evaluation was the use and effectiveness of coverage questions on enumerator-completed questionnaires for Census 2000. The intent of these questions was to identify people who otherwise would have been missed or included in error. The census operations which used these questions were L/E, U/E, NRFU, and CIFU. Enumerators began by asking how many people were living or staying in the housing unit on Census Day. After collecting the appropriate person and housing unit information, the enumerator asked two coverage questions, which were designed to get an accurate enumeration of all people and housing units.

For the first question, C1, the enumerator referred to Census Day, April 1, 2000, and asked:

I need to make sure I have counted everyone who lived or stayed here on April 1, 2000. Did I miss -

- *any children, including foster children?*
- *anyone away on business or vacation?*
- *any roomers or housemates?*
- *anyone else who had no other home?*

For the second question, C2, the enumerator referred to Census Day, April 1, 2000, and asked:

The Census Bureau has already counted certain people so I don't want to count them again here. On April 1, 2000, were any of the people you told me about -

- *away at college?*
- *away in the Armed Forces?*
- *in a nursing home?*
- *in a correctional facility?*

Enumerators were supposed to mark either “Yes” or “No” for each coverage question and then take appropriate action whenever the response was “Yes.” However, about one-third of the time, enumerators left these questions blank. Approximately 1.1 percent of the responses were “Yes” for C1, meaning that someone had been missed, and 0.7 percent were “Yes” for C2, meaning that someone should be counted elsewhere. Among the operations, the CIFU operation saw the lowest percentage of returns with these questions marked, 49.4 percent and 49.2 percent for C1 and C2, respectively. The response rate for these questions in the NRFU operation was approximately 68 percent for both.

Inconsistencies appeared for the people associated with these returns. Only 21.8 percent of the returns that had the “Yes” box marked for C1, had at least one person added. Only 43.4 percent of the returns with C2 marked as “Yes” had at least one person deleted.

Lack of information made it difficult to get an accurate account of the people who were missed or included in error. Therefore, a net result of people added or deleted may not be inferred from the data in this report. For the same reason, the demographic data included in this document may not accurately reflect the distributions of the people who were truly added to or deleted from the household rosters.

Based on the information about the number of “Add” and “Cancel” boxes marked, the Census Bureau added 77,050 people and deleted 83,160 people. Among the people recorded as adds, 46.6 percent were a race other than White, 57.9 percent were young people (ages 0 to 24), 56.5 percent were males, and 51.2 percent were renters. These groups are traditionally undercounted, however, the Census Bureau cannot infer that these two coverage questions are good for improving the differential undercount because of the inadequacy of the data collected.

Recommendations include:

- Provide additional space immediately following the coverage questions (for entering the names of the people to be added or deleted) to improve the percentage of returns with the “Add” and “Cancel” boxes marked when the coverage questions are answered as “Yes.”
- Use automated instruments and have more training for enumerators (on the purpose of asking these questions) to improve the data collection process.

Category J - Ethnographic Studies

Summaries of the three evaluations in the Ethnographic Studies category follow.

Evaluation J.2, Ethnographic Social Network Tracing of Highly Mobile People (Brownrigg 2003)

Residential mobility has long been identified as a behavior that challenges accurate enumeration and coverage. To learn more about how residential mobility impacts census coverage, the

Ethnographic Social Network Tracing Project researched social networks which include highly mobile people. Highly mobile people were defined as people who make residential moves more often than most people in the U.S. or who habitually migrate among domiciles. Social networks were formally defined and modeled by observing people interact over a six-month period. Researchers traced participants interacting in the social networks to the addresses and locations of their domiciles and reported the identities and characteristics of participants, sets of co-residents, and the domiciles they occupied.

Participants in the six social networks researched were involved in diverse patterns and degrees of mobility: a social network of campers who cooked communally, including survival campers along with recreational campers; seasonal workers who habitually circulated among an average of three term assignments at different distant work sites; a folkloric dance group made up of Mexican former farm workers settling in the rural Midwest; older Haitians who worked together in agricultural fields in the South; commercial fishermen, their friends and family, and industry employees who socialized around a particular Atlantic coast fishing dock; and participants in a local chapter of an American Indian men's society.

Various associations were found between the character of individuals' mobility (whether sedentary, residentially mobile, or habitually mobile), their positions in the interacting social networks and matrices of co-residence, and "census outcomes." In the social networks traced, fewer of the residentially and habitually mobile individuals were found enumerated in Census 2000 than those who remained sedentary. If census operations did not list or enumerate the unit that was the census residence of one or more individuals, or did not place the unit in accurate census geography, then it was unlikely that any census records could be found for anyone living in that unit. Similarly, if census operations listed and correctly placed a census residence in geography, but then did not enumerate it or enumerated it as vacant or with entirely different people, it was less likely that records of any of its co-occupants could be found. If a unit had been listed and enumerated more than once, then all or most co-residents might be duplicated.

The omission or erroneous inclusion of certain individuals in their correct census residence where at least one of their reported co-residents was enumerated and served as the census respondent arise from respondent behaviors. These "within-unit" results reflect dynamics among co-residents that influence who is reported.

Relationships were found in these small scale social network studies between individuals' mobility behaviors and both "unit-based" and "within-unit" errors. Habitually mobile people often stayed in types of domiciles Census 2000 did not list. In listed housing, how household respondents perceived and interpreted an individual's current and historic mobility influenced whether or not the person was reported. These relationships combined to produce the net effect that more individuals traced as habitually mobile or residentially mobile were omitted than were found enumerated.

If an individual's census residence was not listed, then that person had no unit of enumeration where he or she could be correctly enumerated in Census 2000. Unlisted types included unconventional domiciles (e.g., camp grounds, docked fishing boats, cheap motels, farm workers' rental labor camps), but also conventional housing units such as single family homes, mobile homes, townhouses, condos, and apartments.

Most of the habitually and residentially mobile social network participants who were found enumerated shared certain traits: they had census residences in conventional housing and maintained ties with and repeatedly and routinely returned to the same set of residentially sedentary co-residents in one locality. As long as their census residence was listed and enumerated, records for habitually and residentially mobile people with all these traits were found, no matter how often or how far they went away. In this research study, far more habitually and residentially mobile people lacked at least one of the traits cited above and were omitted.

Evaluation J.3, Comparative Ethnographic Research on Mobile Populations (de la Puente et al 2003)

This evaluation provided results from four commissioned ethnographic studies conducted by social scientists with extensive knowledge of specific subpopulations that are typically residentially mobile. Each researcher had previously conducted research within his/her population of interest and was known, in most cases, as a trusted individual by the community he/she studied. These four ethnographic studies took place before, during, and after Census 2000 to evaluate the lifestyles of the groups and to observe residential mobility activities during these time periods. All researchers used a combination of observation and unstructured interviews in their field works. The four transient populations examined in the ethnographic studies are: urban gang members, Irish travelers, seasonal residents or "snowbirds" in Arizona, and American Indians residing in the San Francisco Bay Area. Across these four distinct populations, many common barriers to enumeration were found. Many of these barriers have been studied and documented in previous ethnographic studies of hard-to-enumerate populations.

The barriers to enumeration identified in the current study include:

- Residential mobility. Residents may be hard to contact (i.e., not reached by traditional enumeration methods – in person or by mail) or they may have difficulty providing a specific place of usual residence. Most individuals in the study who were aware of the census residence rules, as presented on the census form, did not find them helpful in reporting a usual residence.

- Distrust and/or fear. There are two related reasons why there is reluctance to provide the Census Bureau with personal information. The first is applicable to persons who engage in illegal or unconventional activities. This can range from violation of a civil or criminal law to involvement in living arrangements that violate either public or private housing rules. Underlying this phenomenon is the fear that information provided to the Census Bureau is not kept confidential by the agency and that divulging such information may result in some penalty or prosecution if it fell into the wrong hands. The second and related reason for the reluctance to provide personal information in the census is a broader sense of distrust in government coupled with the unwillingness to provide personal information to an entity whose intentions are questioned.
- Irregular and complex household arrangements. In some cases, violation of housing rules and distrust in government may prevent honest responses. In others, it is unclear to respondents whom to classify as a household member when some of those living in the house are transients.
- Disinterest. In some cases, the Census Bureau's outreach effort did not resonate for some mobile groups. Either they were not exposed to the campaign or they chose not to listen to it or believe the claims made in it. This segment of the population may also be unresponsive to mass marketing strategies. Consequently, they do not understand why the census is necessary nor do they understand the process.

Some of the key recommendations made by the ethnographers who conducted the field work are:

- Enlist support from community organizations to promote census awareness and encourage census participation.
- Direct outreach programs to specific transient groups.
- Clarify residence rules for transients.
- Enumerate in nontraditional sites.
- Make sure that all undeveloped and public land campsites are designated for enumeration.

Evaluation J.4, The Enumeration of Colonias in Census 2000: Perspectives of Ethnographers and Census Enumerators (de la Puente and Stemper 2003)

Colonias are generally unincorporated and low income residential subdivisions, lacking basic infrastructure and services, along the border between the U.S. and Mexico. These settlements have been in existence for decades, but the exodus of the poor to colonias began in full force during the 1980s and 1990s. The low cost of land in colonias provided opportunities for home ownership and relief from higher housing costs in border cities.

The aim of this report was to provide qualitative information on how Census 2000 was conducted in selected colonias. This information comes from two sources. The first source is four studies conducted by ethnographers with field work experience in colonias and with knowledge of these settlements working under contract for the Census Bureau. The second data source includes the results of focus groups with census enumerators and crew leaders who worked in the selected colonias studied by the ethnographers.

Ethnographers from four colonias identified and documented the presence of four major barriers to census enumeration. These are: irregular housing, little or no knowledge of English and limited formal education, concerns regarding confidentiality, and complex and fluid households. However, the extent to which these barriers posed problems for Census 2000 enumeration and the Census Bureau's success in dealing with these obstacles varied across the four colonias.

- Irregular housing appeared to be an obstacle in all four colonias. However, ethnographic observations revealed that, for the most part, census enumerators were able to successfully negotiate the obstacles presented by irregular housing.
- Limited reading skills and little or no knowledge of English was cited as an obstacle to enumeration in all four colonias. Regardless of site, the need for a Spanish language census form that can be easily administered by enumerators and readily understood by respondents was documented by all ethnographers.
- All ethnographers reported that colonia residents expressed concerns regarding the confidentiality of census data. Lack of trust in government and leeringness of non-colonia residents prevailed across all four sites. However, it appears that, for the most part, these concerns were counterbalanced by Census Bureau efforts to promote Census 2000 via paid advertisement in the Spanish language media.
- Complex households and households with mobile and ambiguous members were prevalent in all four colonias. However, this situation was particularly pronounced in the colonia situated in Riverside County, California because of the sizeable number of migrant workers residing in this county.

Although colonias on the U.S./Mexico border are, for the most part, ethnically homogeneous, there is consensus among ethnographers that it is inappropriate to assume the same degree of homogeneity on other key dimensions such as language, the extent of social cohesion (i.e., community) among colonia residents, and the level of infrastructure development.

Recommendations based on this research include:

- Consider revising the training method and training materials for enumerators and crew leaders working in colonias. Emphasize classroom training less and emphasize on-the-job training.

- Use cultural facilitators and *promotoras* developed by the Dallas RO and implemented in the El Paso County, Texas site as a starting point and initiate research that will inform the Census Bureau on how to best use these initiatives in colonias in all four border states.
- Continue and expand the practice of hiring Spanish speaking enumerators who are familiar with colonias.
- Continue to use targeted paid advertising in both English and Spanish.
- Initiate research that will help the Census Bureau determine if mailing out Spanish language forms in 2010 is a viable strategy.

Category K - Data Capture

Summaries of the two evaluations in the Data Capture category follow.

Evaluation K.1.a, DCS 2000 Data Capture Audit Resolution Process (Rosenthal 2003b)

This evaluation looked at results from the Data Capture Audit Resolution (DCAR) process. DCAR identified errors in interpreting scanned questionnaires.

The DCAR consisted of three phases: an automated review of data used to set person panel³ and roster entry⁴ statuses, an edit to compare respondent or enumerator responses on household size to a household population count derived from a tally of person panels and roster entries, and a clerical review of images and an update of data for questionnaires whose response records had conflicting household size information.

There were two types of clerical review of computer images of questionnaires that failed edit: the Audit Count Check (Count Check) and the Audit Status Review (Status Review). The Count Check required that clerks review and correct the Optical Mark Recognition (OMR) interpretation of respondent or enumerator responses on household size only. They did not make corrections to the OMR fields based on a review of person panels or roster entries. The Status Review required that clerks review and correct the OMR fields. In addition, they reviewed and corrected the status of person panels and roster entries.

The Census 2000 Data Capture System (DCS 2000) successfully captured the response data that were input to the determination of household size. It successfully captured numeric responses and accurately identified the presence of responses in check boxes.

³ The number of person panels equals the number of person records associated with a return.

⁴ The number of roster entries equals the number of names listed in the questionnaire roster.

Of the 126,866,759 returns that were sent to DCAR, 124,194,637 returns (97.89 percent) passed the edit. Of the 2,672,122 failed edits, the Count Check process included 882,555 returns (33.03 percent) and the Status Review process included 1,789,567 returns (66.97 percent).

The rate of edit failures varied only slightly across Data Capture Center within form type. The rate at which mail returns passed the DCAR edit varied greatly by household size. About 98 percent of mail returns with a household size between 1 and 9 passed the DCAR edit and about 61 percent of the mail returns with a household size of 10 or more passed the edit. The lower rate for large households may be due in part to the limit of 12 names that could be reported on a mail return. The rate at which enumerator returns passed the DCAR edit varied slightly by household size. The rate decreased slightly as household size increased. The rate for households with ten or more persons is much larger for enumerator returns compared to mail returns, 96 percent versus 61 percent.

As the check-in date of the return became further removed from Census Day, the percent sent to Count Check and Status Review increased for mail returns faster than for enumerator returns, indicating more consistent quality for enumerator returns over time.

The status of pre-audit duplicates among person panels and among roster entries on mail returns were compared to their post-Status Review status. There were 52,406 pre-audit duplicate person panels and 41,562 pre-audit duplicate roster entries. Only 507 (0.97 percent) of the person panels were determined not to be a duplicate and only 1,233 (2.97 percent) of the roster entries were determined not to be a duplicate by the Status Review process. The lower rate of change for person panels may indicate that without associated demographic characteristics it is more difficult to accurately identify duplicates.

The Status Review changed only a small percentage of pre-audit statuses. The Status Review process changed about 12 percent of the statuses for person panels with a pre-audit status of valid. The Status Review process changed about 13 percent of the statuses for person panels with a pre-audit status of invalid. The Status Review process changed about 29 percent of the statuses for short form mail return roster entries with a pre-audit status of valid. The Status Review process changed less than 0.5 percent of the statuses for short form mail return roster entries with a pre-audit status of invalid. The Status Review process changed about 10 percent of the statuses for long form mail return roster entries with a pre-audit status of valid. The Status Review process changed about 4 percent of the statuses for long form mail return roster entries with a pre-audit status of invalid.

The DCS 2000 successfully interpreted a very high percentage of the hand-written numeric entries for household size. Although the occurrence of interpretation errors was low for both the mail and enumerator returns, the entries made by enumerators were misinterpreted about twice as often as those made by the respondents to mail returns. The DCAR process corrected the household size for about 64 percent of the enumerator returns that failed the Count Check but correction were made to only about 32 percent of the mail returns that failed the Count Check.

The DCAR corrected the data on a large number of cases that would have been included in the CEFU had the DCAR process not been performed.

The recommendation is to incorporate a similar process into the 2010 Census.

Evaluation K.1.b, Evaluation of the Quality of the Data Capture System and the Impact of the Data Capture Mode on the Data Quality (Conklin 2003)

The purpose of this evaluation was to see how well the reading of census forms could be delegated to automated data capture and imaging technology. The raw data for this evaluation consisted of a sample of 768,000 short forms and 768,000 long forms distributed among: mailout/mailback short and long forms, enumerator short and long forms, and U/L short and long forms. The enumerator and U/L forms included Puerto Rico and U.S. versions. The mailout/mailback forms included both English and Spanish versions.

Production technology involved three modes: OMR to determine in an automated process if a check box on the form did or did not have a mark in it; OMR to capture the value of each character in a write-in response; and Key From Image (KFI) mode to have a person interpret and key responses from a computer image of the form.

Both the production automated technology and the evaluation KFI data capture were subject to any one of the following errors: failure to read a field on the form, picking up content that was not really there (as in trying to interpret a stray mark), incorrectly capturing the content on the paper, or correctly capturing what the respondent wrote but not what the respondent intended.

The check-box contents must be identical to be considered a match. Picking up the wrong check-box is a hard match error. The write-in contents did not have to be identical to be considered a match. Missing characters or dropping or adding characters can lead to soft match errors. The divergence between the contents was scored using a soft match algorithm. A soft match error occurred when the divergence score exceeded a threshold.

Content can differ from intent. This can happen for reasons such as stray marks being read as characters or if the respondent writes poorly. The standard for Key From Paper (KFP) was to capture content with no more than a 2 percent error rate. The performance of the automated technology depended on whether the character recognition algorithm determined the content was clear enough to process. If the automated technology determined the content of a write-in field was clear, it processed it with a typical error rate of 1.0 to 1.1 percent. If the automated technology determined the content of a check-box field was clear, it processed it with a typical error rate of 1.2 to 1.5 percent. If the automated technology rejected content as unclear, the typical error rate after remedial keying by human operators was 4.8 to 5.3 percent.

The evaluation concluded with 90 percent confidence the modes are all significantly different from one another. The KFI mode tends to deal with content particularly hard for human or machine to interpret and therefore the error rate is not necessarily a poor reflection on the automated technology or the keyers. The error rates shown in this evaluation should be considered conservative upper limits for the true rates attributable solely to the hardware and software configuration of the automated technology.

Respondent-returned forms had statistically significantly higher nonblank hard or soft match error rates for ethnic, name, and race fields compared to enumerator-returned forms. Although enumerator-returned forms had lower soft match error rates for name related fields compared to respondent-returned forms, the rates for name related fields were higher compared to rates for other fields on forms returned by enumerators.

For fields filled out for only one person on a form, the hard or soft match error rate was significantly affected depending on the specific field being considered; form type or field category did not have a significant influence. For fields filled out for multiple persons on a form, the soft match error rate was significantly affected by form type and field category.

Census 2000 RCC had a significant influence on the hard or soft match error rate. The soft match error rate for name related fields in RCCs covering areas of traditional immigrant concentration in Florida, Los Angeles, and New York City was significantly higher compared to other RCCs.

For fields that were filled out for only one person on a form, the largest significant factor affecting the nonblank error rate was form. For fields that were filled out for multiple persons on a form, the largest significant factor affecting the nonblank error rate was field category.

The race response had a statistically significant effect on the nonblank error rate. Within the limited data set for race, evaluators were not able to find individual error rates that were outliers. The effect of race may be part of other significant factors. It would be helpful to include other factors with race in a future evaluation.

For fields filled out for only one person on a form, the error rate was not significantly affected by data capture mode. For fields filled out for multiple persons on a form, the specific field being considered and the data capture mode interacted to significantly affect the error rate.

The most frequent causes for failing to capture the intended response were an extra check-box, missing characters, or a wrong character. The most common reasons found for these problems were poor handwriting, no reason found, or rules not followed.

The future role of the automated technology reduces to two possibilities. The automated technology has a supporting role in decennial census processing or the automated technology has a dominant role in decennial census processing.

Several possible research questions exist for tests leading up to the 2010 Census.

- Should the Census Bureau expand efforts to make certain groups of fields easier for respondents to understand and fill out?
- Do the outlier error rates for name related fields suggest challenges to the automated technology that require increased attention?

- Is the especially high nonblank error rate for name related fields in Census 2000 RCCs of traditional immigrant concentration something that requires more investigation?
- Should certain fields sent automatically to KFI be allowed to go through the automated technology for processing?

Category L - Processing Systems

Summaries of the five evaluations in the Processing Systems category follow.

Evaluation L.2, Operational Analysis of the Decennial Response File Linking and Setting of Housing Unit Status and Expected Household Size (Rosenthal 2003a)

The Decennial Response File is the first in a series of files which ultimately produces final census population counts. A return represented a single household enumeration. A return consisted of one or more Decennial Response File household forms. For example, an enumerator continuation form was linked to its parent enumerator form to create one return. Similarly, a Be Counted Form for a partial household was linked to a mail return form if the Be Counted Form contained information on additional household members not included on the mail return questionnaire.

In mailback areas, the use of two forms to enumerate large households most commonly occurred in the NRFU and CIFU operations. Typically the two forms were an enumerator first form and an enumerator continuation form.

Large households on mail returns generally did not involve linking. The Census Bureau produced a composite record for these large households from the originating mail return and a CEFU telephone interview.

The Census Bureau assigned to each return a housing unit status and, if the status was determined to be occupied, an expected household size. Then, the Census Bureau applied the Primary Selection Algorithm, a computer program run on the Decennial Response File to select one return to represent each housing unit in the census, if multiple returns were present for a housing unit.

This evaluation presented the results from the linking of census forms and the setting of housing unit status and expected household size.

Of 129,389,529 returns, 1,387,085 returns (1.07 percent) were linked; that is, they were returns comprised of two or more forms. Of these, 39,108 returns (2.82 percent), had three or more forms.

The TEA with the highest rate of linked returns was in U/E. Large households probably caused this result. Most linked returns (1,384,233 returns or 99.79 percent) were comprised of an enumerator first form and an enumerator continuation form. Enumerators used this combination of forms to enumerate large households in the L/E, U/E, NRFU, and CIFU operations. Linkage rates comparable to U/E did not occur in L/E, probably because of a processing error. Enumerator continuation forms in L/E—along with U/L adds and U/E adds—were erroneously omitted from the Decennial Response File.

Of 129,389,529 returns, 1,318,350 returns (1.02 percent) had one or more unresolved categories: Occupied with Unresolved Population Count, Unresolved Occupied/Vacant, and Unresolved Occupied/Vacant/Delete.

The Census Bureau sent mail returns to CEFU if there was an inconsistency in household size, leaving few self-response returns unresolved.

A programming error affected the status resolution for some Vacant enumerator returns. Mistakenly, any Interview Summary Population of 0 was recoded to blank. As a result, the Census Bureau may have classified up to 133,438 Vacant returns as Deletes and up to 258,963 Vacant returns as Unresolved Occupied/Vacant. As many as 145,367 housing units of the 191,826 housing units in the census that had their occupancy status imputed (75.78 percent) may have been affected by this latter error.

There were 712,858 unresolved enumerator returns (1.51 percent), 329,895 returns were Occupied with Unresolved Population Count, 329,266 returns were Unresolved Occupied/Vacant, and 53,697 returns were Unresolved Occupied/Vacant/Delete.

For occupied self-response (restricted to paper mailback questionnaires) and enumerator returns, setting the expected household size was usually straightforward. For 74,725,437 self-response returns (93.71 percent) the number of valid person records and roster names corresponded to the respondent-reported household size. For enumerator returns, most household size measures also were consistent. For linked returns or unlinked returns that had the “continuation form(s) attached” box checked and the expected household size equal to the Interview Summary Population, 1,475,382 returns (99.11 percent) had the same Interview Summary Population and respondent-reported household size. For unlinked returns with the Interview Summary Population less than or equal to five, 26,897,133 returns (99.52 percent) had the same Interview Summary Population, number of valid person records, and respondent-reported household size. For unlinked returns with the Interview Summary Population greater than five and the respondent-reported household size greater than five, 68,599 returns (99.58 percent) had the same Interview Summary Population and respondent-reported household size.

Recommendations include:

- Attempt to link only enumerator first and enumerator continuation forms, if such forms exist in the future. Doing so would simplify the linking process, cause very little loss of data, and would have almost no effect on the population counts.

- Redesign the Interview Summary Section, if this section exists in the future, to improve the consistency of responses.
- Use hand held computers to help ensure data capture and consistency of responses.

Evaluation L.3.a, Analysis of the Primary Selection Algorithm (Baumgardner 2002)

The full report for this evaluation is not available because it contains proprietary information.

While various methods of collecting data were implemented with the desire of obtaining a more accurate census count, the various methods also presented the possibility of receiving multiple responses for a single Census ID. The Primary Selection Algorithm was the computer program designed to resolve the receipt of multiple responses from housing units. Major features of the Census 2000 Primary Selection Algorithm design included performing person matching between returns, constructing Primary Selection Algorithm households, selecting the primary Primary Selection Algorithm household, and selecting additional persons for the census household that were not in the primary Primary Selection Algorithm household.

Less than 10 percent of all Census IDs on the Decennial Response File were enumerated by more than one return. More than 95 percent of these were enumerated by only two returns. About 55 percent of the Census IDs enumerated by two returns were the result of two enumerator returns and about 82 percent of these were the result of returns from NRFU and CIFU. About a third of all Census IDs with two returns consisted of one enumerator and one mail return; about 96 percent of these were the result of a mailback return and a NRFU return.

Some returns were defined as ineligible for the Primary Selection Algorithm process. There were 2,656,951 ineligible returns at all Census IDs. More than 67 percent of these returns were ineligible due to being classified as a deleted housing unit record. Taking these ineligible returns out of the universe, 8,960,245 Census IDs (less than 8 percent of Census IDs on the Decennial Response File) had more than one eligible return.

A Primary Selection Algorithm household was a set of associated persons at one Census ID. The set may contain no persons (a vacant Primary Selection Algorithm household) or one or more persons. If two or more returns for the same Census ID had at least one person in common (determined by person matching), then these returns formed a single Primary Selection Algorithm household.

Over 73 percent of Census IDs with multiple eligible returns had just one Primary Selection Algorithm household. Census IDs with two or more Primary Selection Algorithm households accounted for just over 2 percent of all Decennial Response File Census IDs. The primary Primary Selection Algorithm household was the Primary Selection Algorithm household that was used in further processing. When more than one Primary Selection Algorithm household existed, the primary Primary Selection Algorithm household was selected by sequentially applying criteria to all of the Primary Selection Algorithm households until only one Primary Selection Algorithm household was selected.

Most Primary Selection Algorithm households at Census IDs with multiple returns consisted of one or two returns. Two-return Primary Selection Algorithm households were most often formed by two enumerator returns or one mail return combined with one enumerator return. When two enumerator returns formed a Primary Selection Algorithm household, over 91 percent were the result of one return from NRFU and one return from CIFU. This was expected due to the design of the CIFU operation.

Of the 8,716,359 Census IDs with two eligible returns, over 70 percent had a redundant return (a return containing only person records represented on the basic return of a Primary Selection Algorithm household) and almost 57 percent of these redundant returns were not vacant.

Almost 85 percent of all redundant returns were enumerator returns. More than 55 percent of redundant enumerator returns resulted from NRFU and nearly 88 percent of these were occupied, most likely due to the receipt of a late mail return. About 43 percent of redundant enumerator returns resulted from CIFU and 97 percent of these were vacant, as expected.

Of the 2,349,988 Census IDs with two Primary Selection Algorithm households, more than half had an enumerator return as the basic return of both Primary Selection Algorithm households. Nearly 80 percent of these cases resulted from one return from NRFU and one return from CIFU. This is most likely due to a vacant return from the NRFU operation and an occupied return from the CIFU operation.

Most cases of multiple enumerations in Census 2000 were expected as a result of the design of census operations. The Primary Selection Algorithm was designed to be robust and handle as many unusual cases as possible, but the results show that these unusual cases were very few. If this operation is implemented in the same manner in the future, a processing step done prior to the Primary Selection Algorithm should remove from further processing a CIFU return that just confirms the status of a NRFU return. The NRFU return in this case should be flagged to indicate that its status was confirmed.

Evaluation L.3.b, Resolution of Multiple Census Returns Using a Reinterview (Baumgardner 2003)

The full report for this evaluation is not available because it contains proprietary information.

The focus of this evaluation was to examine whether or not the Primary Selection Algorithm made the best decisions on the determination of the census household, given the information collected from returns for the Census ID. The analysis concentrated on Census IDs with two returns since 97 percent of all Census IDs with multiple returns had exactly two returns.

A sample of Census IDs affected by the Primary Selection Algorithm was selected. An interview was conducted at each Census ID with someone familiar with the household enumerated during Census 2000. The goal of the reinterview was to determine the residency status of each person on the census returns at the Census ID. The residency statuses obtained in

the reinterview were then used to determine if the Primary Selection Algorithm made the best decisions on the determination of the census household.

One set of limitations to this study includes operational problems with data collection. Unit nonresponse, allowing the use of proxy respondents, and recall bias may contribute to missing or inaccurate responses. Another set of limitations to this study includes limits to the statistical analysis. This study did not investigate vacant returns or evaluate the person matching process. There were also person records with an unresolved residency status and errors in the inclusion of some Census IDs in the sample. These factors make it difficult to fully evaluate the performance of the Primary Selection Algorithm.

When person matching did not match people across the two returns at a Census ID, these two returns formed two separate Primary Selection Algorithm households. At Census IDs with two Primary Selection Algorithm households, the reinterview determined that there were residents in both households about 38 percent of the time, residents in just one of the households about 58 percent of the time, and no residents in either household about 4 percent of the time.

At Census IDs with two Primary Selection Algorithm households where the reinterview determined that there were residents in both of the households, person matching was performed and missed a duplicate identified during the reinterview in about 16 percent of the cases. Also of interest is how often the Primary Selection Algorithm picked the “best” Primary Selection Algorithm household to represent the Census ID since both of the households contained residents. “Best” is defined here as the Primary Selection Algorithm household with the greater net number of residents. The net number of residents is the balance after subtracting the number of nonresidents from the number of residents. At Census IDs with residents in two Primary Selection Algorithm households, the “best” household or a household which was identical in terms of net residents to the other household at the Census ID was selected about 80 percent of the time.

Nearly 58 percent of Census IDs with two Primary Selection Algorithm households had at least one resident in just one of those households. At about 65 percent of these Census IDs, the Primary Selection Algorithm selected the household that contained at least one resident. Of the Census IDs with two Primary Selection Algorithm households and residents in just one of those households, the effectiveness of the Primary Selection Algorithm household selection criteria was examined. As expected, higher priority selection criteria were more effective at selecting the correct Primary Selection Algorithm household than the lower priority selection criteria. However, the selection criterion “CEFU Status” did not perform as well as expected. It selected the correct Primary Selection Algorithm household only about 69 percent of the time that it was used.

In the future, an evaluation such as this one should set out to determine if the Primary Selection Algorithm selected the right household regardless of whether or not the Primary Selection Algorithm household is occupied. Furthermore, the questionnaire used for the reinterview should be designed to manage both occupied and vacant returns. It should also not rely on census residence rules to determine residency status for people at the Census ID, since the

Primary Selection Algorithm itself cannot take those rules into account when making decisions regarding the census household.

Evaluation L.4, Census Unedited File Creation (Jonas 2003a)

The Census 2000 Hundred Percent Census Unedited File contains all the household and person records included in Census 2000. It has all the attributes of the final Census file, except the imputation of person characteristics where needed. The purpose of Census Unedited File creation is to determine which addresses are in the census and to determine the count of persons at each such address.

Nearly 128 million addresses were either on the DMAF as Census 2000 began or were added to it in the course of Census 2000 operations. Prior to unduplication procedures, approximately 117.3 million were ultimately resolved as housing unit addresses. Just over 9 million addresses were determined to not be valid addresses and roughly 1.4 million addresses were determined to be nonvalid duplicates of valid addresses on the DMAF.

Of the 117.3 million addresses resolved as housing unit addresses, 106.7 million were determined or imputed to be occupied and the remaining 10.6 million were determined or imputed to be vacant.

Roughly half a million addresses had their status resolved by imputation. There were 195,245 addresses determined to be valid Census addresses whose occupancy status could not be determined and had to be imputed. There were 296,617 addresses whose validity as Census addresses could not be determined. As a result, their validity and their occupancy status were both imputed. There was no enumeration data on the Decennial Response File or the DMAF for 251,477 (84.8 percent) of the addresses whose validity as Census addresses could not be determined.

Recommendations for some changes affecting Census Unedited File creation include:

- Use stronger software QA processes to ensure more complete adherence to specifications.
- Refine the timing of late census followup operations to ensure that addresses added by those operations are placed on the DMAF in time for the questionnaires from those addresses to be included in the census.

Evaluation L.5, Operational Requirements Study: The Beta Site Systems Testing and Management Facility (Titan Systems Corporation 2003)

This study assessed the extent to which the requirements for the Beta Site operation and its internal processes supported various automated systems used during Census 2000. The findings presented are qualitative in nature as they reflect the varied opinions and insights of the Beta Site operations personnel and customers who were interviewed by the Titan Systems Corporation.

The Beta Site was a software evaluation facility within the Census Bureau that was involved in the testing and deployment of Census 2000 systems and related components. Its primary objective was to assess a system's deployment readiness; however, it also conducted security testing, provided software release services, and performed network monitoring and troubleshooting support.

Security evaluation was a distinct phase of the Beta Site testing. The Beta Site personnel worked in a cooperative fashion with the Information Technology Security Office to assure appropriate security considerations were proactively addressed. Overall, the structure of the testing processes and associated functions were comprehensive and were aligned to support the objectives of the Beta Site.

Planning for the Beta Site support for Census 2000 began in mid-1996 and continued through the census to accommodate changing operational requirements, as needed. The physical site was constructed in 1996. In addition to testing Census 2000 systems, the Beta Site had to address such challenges as ramping up the testing infrastructure and performing Year 2000 compliance testing. According to a post-assessment study of the Beta Site, over 1,200 software tests were performed by the Beta Site from late Fiscal Year 1997 through Fiscal Year 2000. The Beta Site also maintained system configurations for over 8,000 personal computers and 570 servers.

Given the success of Census 2000 and the unprecedented reliance on automated systems, it is evident that the Beta Site played an important role in the decennial census and contributed significantly to its success.

Although the software validation role of the Beta Site operation was widely seen within the Census Bureau as being a necessary function, many of the Beta Site's customers expressed concerns over the efficiency, consistency, and timeliness of the testing processes that were employed. The requirements for the Beta Site could have focused more attention on the impact that its internal processes would have on customers' operations. Conversely, developers needed to factor in time for Beta Site testing in their development process. In this regard, the Beta Site personnel noted that, from their perspective, there were a large number of urgent requests that might have been avoided with proper planning/scheduling by the program offices.

The issue of who was responsible for developing test plans was not fully addressed. Although the Program Master Plan discussed the Beta Site Workflow and the "receipt of requirements," which included test plans, data, and cases from the developer, the precision of those requirements was never fully defined. The perception of the Beta Site staff was that

responsibility for test plans was defined through meetings with customers. In discussions with testers and customers alike, it was evident that this issue was not fully resolved during the requirements phase.

Early planning may have addressed requirements for two-way communications. Although a set of physical, logistical, and procedural requirements was outlined in April 1997, they did not adequately address the need for a structure to ensure effective communications between the Beta Site testers and developers. Interviews confirmed that the Beta Site process was often unclear to most customers and this led to a significant number of communication difficulties, especially when the need arose to escalate issue resolution to a higher level.

In 1997, the Beta Site management opted to use the services of the General Service Administration's Federal Systems Integration and Management Center as a means of acquiring a capable prime contractor for the Beta Site. The Federal Systems Integration and Management Center manages multiple-award contracts with qualified system integrators who can be competitively selected in a relatively short period of time. Resources permitting, it may have been beneficial for the Beta Site to use the Federal Systems Integration and Management Center in the requirements planning area.

Requirements did not give adequate consideration to the complexities of managing the Census 2000 systems in a networked environment and the division of responsibilities within the Census Bureau's technical management infrastructure was problematic. This was a crucial issue in view of the size of the network. During the decennial census, network administration and configuration issues arose between the Technologies Management Office and the Beta Site.

These and other findings led to the following recommendations:

- Fully consider requirements for communication processes.
- Improve testers' knowledge about the purpose, use, and capabilities of the software.
- Major advancements in technology will require early scoping of the level of effort required from the Beta Site in 2010.
- Improve life-cycle model for the 2010 Census.

Category M - Quality Assurance Evaluations

Summaries of the two evaluations in the QA Evaluations category follow.

Evaluation M.1, Evaluation of the Census 2000 QA Philosophy and Approach Used in the Address List Development and Enumeration Operations (Morganstein et al 2003)

Census 2000 involved numerous field operations and, at its peak, employed almost a half-million temporary workers. Managing the quality of the data produced by this large, decentralized, and transient workforce was a major challenge, which the Census Bureau attempted to meet by introducing an extensive QA program into its ongoing operations. The objectives of this QA were to minimize significant performance errors, to prevent the clustering of significant performance errors, and to promote continuous improvement.

Evaluation M.1 evaluated the effectiveness of the QA programs developed for and implemented in the major field operations in Census 2000 and, specifically, those operations used to update the Census Bureau's address list and directly enumerate the population. Further, it identified strengths and major deficiencies and provided a critique of the Census Bureau's QA philosophy.

The evaluation, which was conducted by Westat, was accomplished through a combination of approaches, including a review of materials prepared for use in training and in collecting data, examining observation reports, interviewing a range of Census Bureau staff who had been closely involved in the many phases of the QA program, and by examining such materials as were available concerning the operations, as well as the successes or failures of the QA programs. As a final step, Westat held discussions with staff in the statistical offices of the United Kingdom, Canada, and Australia, all of which had recently conducted a national census, to determine how these countries approached the application of QA in their census programs, and the possible relevance of their actions to future QA planning of the Census Bureau.

Census 2000 continued the tradition, initiated in the 1960 Census, of incorporating into field operations numerous activities described as QA. This commitment to quality and QA certainly is a significant "strength." The Census Bureau also met its objective for QA that it be completely transparent in Census 2000. To that end, materials used to train enumerators and first level supervisors contained specific references as to why QA was important and as to how it would be implemented and all enumerators were exposed to the concept of and need for "quality" performance and, generally, were measured against the established standards. Finally, most operations, had some form of QA process in place.

A review of observation reports and interviews with Census Bureau staff who were involved in the QA program revealed that the overall perception was that the Census 2000 QA field program was an important element in preventing significant errors and in preventing the clustering of significant errors. Although errors of both types did occur, for the most part, they were caught expeditiously and rectified. On this basis, the QA field programs can be viewed, generally, as successfully meeting the first two elements of the Census Bureau's QA mission to prevent significant errors and to prevent the clustering of significant errors.

As to the Census Bureau's goal of "making the best use of the available technology and statistical process tools with respect to its desire to promote timely and continuous improvement throughout the field operations," the authors concluded that the answer is mixed. Based on the Census Bureau's evaluations to date and comments from those involved, many of the early activities in preparing for Census 2000 were seen as having utilized a full QA approach that met the Census Bureau's stated goal of promoting timely and continuous improvement. Examples of activities considered as having successfully utilized a QA approach include preparing the geographic framework and printing questionnaires and related forms. However, in the context of what actually transpired during the data collection phase, the perception was less clear and decidedly mixed.

QA was not perceived as an equal partner, nor was QA staff given either the necessary authority or the required freedom to complete its task successfully. Finally, dedicated QA staff was not assigned to RCCs. A vital aspect of the QA program for promoting continuous improvement, real-time capture and dissemination of data during the data collection process, with which to monitor, evaluate, and react, was not implemented.

Some key suggestions are noted below:

- Ensure that the QA effort is seen as an integral and important element in the 2010 Census program. To that end, provide adequate resources, both in funding and staff, and initiate early planning, research, and testing for the QA phase.
- Involve the executive staff in supporting and monitoring QA efforts, especially throughout the data collection phase.
- Establish a senior management team to coordinate and approve the overall QA plan and, throughout the decennial period, to review progress and resolve issues.
- Establish the equality of QA relative to production. Simply put, QA must be seen, understood, and accepted as an essential element of the census and as an equal partner at all levels of planning, implementation, and review.
- Develop and implement a Management Information System component which provides management, in real-time, with relevant information on the quality of the data collection elements.
- Expedite the documentation of Census 2000 and establish ready access to the information. Some examples include the QA evaluation program and the documentation of the experiences, problems, and solutions, suggestions, and recommendations of staff, and the accumulation of memoranda detailing problems, issues, and solutions.

Evaluation M.2, Effectiveness of Variables Used in the Model to Detect Discrepant Results During Reinterview and the Identification of New Variables (Johanson 2003)

The Census 2000 NRFU Reinterview program included three components: a random reinterview, an administrative reinterview, and a supplemental reinterview. A portion of completed enumerator questionnaires were selected to be reinterviewed and, once the reinterview was conducted, the unit status and household roster were compared to the original enumeration. The purpose of the reinterview program was to identify faulty data collection, both intentional and unintentional.

The random reinterview component was designed to verify work from each enumerator. Virtually all enumerators who completed a minimum of ten enumerator questionnaires had one or more of their questionnaires selected for random reinterview. Random reinterviews represented 93.09 percent of the cases selected for the reinterview program. The remainder of the reinterview cases were administrative and supplemental reinterview cases (4.34 percent and 2.57 percent, respectively).

Outlier enumerators were identified for administrative reinterview by comparing questionnaire characteristics of each enumerator against the average for their area. A high vacancy rate, a high rate of partial interviews, a high delete rate, a high rate of questionnaires with a population count of one, and differences in average population per household were variables used in the comparison. The reports identifying these outlier enumerators were run once a week. Over the entire NRFU operation, 291,441 enumerators were flagged as outliers for one of the reasons above. This is 62.57 percent of enumerators with completed work. Not all of these enumerators had cases selected for administrative reinterview. At the discretion of supervisors, approximately 5 percent of enumerators flagged for administrative reinterview had administrative cases selected, or 3.5 percent of all NRFU enumerators.

Supplemental reinterview could be used any time there was reason to suspect cases might not be completed correctly. Supplemental cases with complete reinterview information showed a higher frequency of enumerator error between the original enumeration and the reinterview (11.30 percent) than random and administrative cases (9.42 percent and 9.67 percent, respectively). This higher incidence of error identification shows the effectiveness of the supplemental reinterview component.

At the individual case level, administrative and random reinterview found a similar proportion of cases with discrepancies. Of the enumerators in administrative reinterview, 52.09 percent had one or more cases in error. This is much higher than the percent of enumerators in random reinterview with error cases. The range of error rates for enumerators, depending on the number of random reinterview cases selected, was approximately 10 to 14 percent.

The evaluation included a discriminant regression model to evaluate the effectiveness of various variables in the administrative reinterview. Evaluators were looking for interactions that might help improve the model to identify outlier enumerators for the administrative reinterview. Of the characteristics reviewed for the administrative sample, the high delete variable had the biggest

impact for identifying enumerators with error . However, the regression models showed that very little of the variance associated with the dependent variable (presence of error) was explained by the independent variables of interest. This indicates that the dependent variable could be expected to behave similarly for randomly selected enumerators and enumerators identified as outliers.

Interpreting these results was difficult because of operational limitations. The analysis showed that administrative reinterview was definitely effective in identifying enumerators with high error rates, yet the contribution of the variables used to select the enumerators was not meaningful. This is partially explained by the fact that although enumerators were targeted based on work characteristics, the selection of cases for reinterview did not reflect the characteristic(s) that caused the enumerators' outlier status. Other studies and academic experts have shown that a targeted reinterview can be very beneficial. An example of targeted reinterview would be if an enumerator was an outlier for high deletes, then deleted cases were specifically reinterviewed. Using targeted case selection will likely make the administrative reinterview more effective.

Recommendations include developing procedures to reduce the number of enumerators identified as outliers. This can be accomplished by increasing the critical cut-off levels or accounting for multiple tests and the recurring time periods. Reducing the outliers identified by the administrative test will create a smaller workload to review. This, in turn, should increase the percentage of enumerators reinterviewed and the effectiveness of the administrative reinterview program in identifying enumerators with discrepant results.

Category Q - Organization, Budget, and Management Information System

Evaluation Q.1, Management Evaluation of Census 2000 (IBM Business Consulting Services 2003)

This evaluation documented the effectiveness of the management approach used in the Decennial Census in each of seven areas including the management model, organizational structures and processes, the decision-making process, management information tools, staffing, external influences, and the use of contracts.

The evaluation was conducted by a contractor, IBM Business Consulting Services. The contractor's understanding of the external and internal environment formed the foundation upon which the detailed approach for assessing organizational and management performance was based. The contractor evaluated the management approach by conducting interviews with key Census Bureau staff. A total of 52 interviews were conducted with Census Bureau personnel including nine Executive Staff, 13 Division Chiefs (including Regional Directors), 18 Assistant Division Chiefs, and 12 Branch Chiefs. The contractor interpreted the information collected in the interviews using management theory and experience from previous management reviews to evaluate the decennial census management practices. The interviewees had varying levels of exposure to the issue areas mentioned in the Interview Guide, and they generally commented on

areas with which they had direct knowledge or experience. There was no attempt to stratify the comments from the interviewees, since all the comments were treated equally, independent of organizational unit, experience, or other differences.

Performance Measures:

Key performance indicators revealed that, in certain respects, Census 2000 was the most successful decennial census ever conducted. According to senior Census Bureau officials, the most critical performance measure was Net Coverage. In Census 2000, the percent of net undercount estimate was minus 0.49. A net undercount estimate of minus 0.49 represents a small estimated overcount of the household population for Census 2000. Achievement of a small net coverage error that is close to zero is an important success factor.

The national response rate that determined the NRFU workload was 65 percent, which matched the 65 percent response rate from the 1990 Census and stemmed the decline that had been the trend over recent decades.

The NRFU effort was completed ahead of schedule.

Organization and Structure:

Post 1998, the Census Bureau operated within an organization that was well structured to support its performance objectives. The decennial organization was organized by business process drawing from functional capabilities residing within the participating divisions as required. Within the decennial census business process, a single process owner, the Associate Director for Decennial Census, controlled both line and funding authority for delivering the decennial census product. The overall Census Bureau organizational structure is logical, clearly defined, and, for the most part, tailored toward achieving desired results.

In many of the sub-structures and teams within the decennial organization, however, the leaders of the teams and decision-making bodies were not given or did not choose to exercise true decision-making authority. Leaders often served exclusively as facilitators and consensus builders. Although the intent behind the creation of these organizational bodies was to push decision-making to the lowest management levels technically possible, there was no decision-making authority in place at these lower levels to support that intent.

Management Approach:

Mid-decade, the Census Bureau attempted to institute matrix management to encourage teaming and to distribute decision-making to lower levels. This early attempt failed because there was no centralized integrating or coordinating process for census plans and operations, and decision-making mechanisms were not properly implemented. After 1998, a "centralized" management approach added strong coordination, integration, and decision-making roles in the Decennial Management Division to the teaming arrangements from the failed matrix approach. The centralized approach fulfilled the desired results of matrix management because it fit the

matrixed business model of the decennial organization. The centralized management approach is the approach that best represents true matrix management because it features a more robust Decennial Management Division on the y-axis counterbalancing and coordinating the interests of and the capabilities resident within the participating divisions on the x-axis.

Despite operational successes, the evaluation of the management approach revealed areas that warrant improvement. In assessing the management model, it is important to consider two external drivers that Census Bureau management must respond to (political influences and technological advances), but over which the Census Bureau does not have significant control. During the 1990s, the congressional debate over the use of a sampling strategy versus traditional enumeration strategy caused significant challenges to the Census Bureau.

As the debate over strategy continued, funding levels limited the number of staff members working on Census 2000. This constrained the Census Bureau's ability to generate detailed implementation plans for either a sampling or traditional enumeration approach. The funding shortfalls and an inability to maintain a core decennial staff throughout the decade also led to critical weaknesses in the decennial census knowledge base and had ramifications for many processes and procedures being executed during Census 2000.

Communications and Knowledge Management :

There is little record of an internal communications approach or the implementation of a formal communications structure for Census 2000.

A knowledge management capability to retain corporate knowledge, to support responses to external reporting requirements, and to communicate programmatic changes to decennial census participants in a timely manner would assist in improving communications and in stabilizing and maintaining the decennial census knowledge base throughout the decade.

Although the Census 2000 centralized management approach is an appropriate approach that matches the decennial census business model, improvements could be made that would make the implementation of the approach more effective. Recommendations for improvement include:

- Clearly define and communicate to all decennial census staff the roles and responsibilities of program managers.
- Develop program master plans describing operational plans earlier in the census cycle than was accomplished during Census 2000.
- Develop risk management plans as an integral component of census planning.

Category R - Automation of Census Processes

Summaries of the twelve evaluations in the Automation of Census Processes category follow.

Evaluation R.1.a, Telephone Questionnaire Assistance System Requirements Study (Titan Systems Corporation 2001a)

TQA was a large-scale program that provided telephone assistance to the public during Census 2000. A network of 22 call centers used a combination of automated technologies and agent responses to handle calls from households within the 50 states, DC, and Puerto Rico. This study presents information based on debriefings with personnel involved with the TQA program. A separate customer service survey evaluation (Evaluation A.1.b) provides the user perspective of this system.

TQA was the first time that the Census Bureau outsourced a call center operation. The program was considered the largest operation of its kind implemented in the call center environment. The system was designed to accommodate 11 million calls and received slightly over 6 million throughout its operation, from March 3, 2000 through June 30, 2000. An outbound service for CEFU started while TQA was still in operation. This program used some of the same technologies and a subset of the call centers.

A number of issues confronted the TQA development team such as the high call volume, call distribution across centers, the range of questions and topics that would require responses, and the completion of the automated short form questionnaire by call center agents. Despite the challenges and short time frame for development, the program was considered extremely successful. Both Electronic Data Systems and Census personnel provided exceptional support and dedication to ensure that TQA was successfully implemented.

Compromises on some requirements to collect certain data were made because of the limited time to develop the system. Although many compromises were made through negotiations with subject matter experts and program managers, the lack of some data from TQA impacted the completeness of post-Census evaluations. Some requirements were dropped because of timeline and resource constraints, but the majority of the missing management and evaluation data can be attributed to the GeoTel Intelligent Call Router.

The provision of required data through the Intelligent Call Router proved to be a significant deficiency. The requirements for performance and evaluation data apparently exceeded what the Intelligent Call Router typically provides. The Census Bureau requirements necessitated customized programming which apparently exceeded the system's technical capabilities. It was also not clear that the technical support involved in programming the application completely understood the implications of those requirements. The product was intended to meet the call routing and information tracking requirements that had been clearly specified by the Census Bureau, but the product fell short of its expectations and either did not provide the data or provided the data too late in the process.

These and other findings led to the following key recommendations:

- Begin development early.
- Establish realistic development timelines. For each development effort, timelines must consider the complexities associated with translating and implementing high-level user requirements into a functional system. In addition, time for testing and rework is required to ensure that each system is sufficiently stable for production.
- Establish agency-wide guidance.

Evaluation R.1.b, Coverage Edit Followup System Requirements (Titan Systems Corporation 2001b)

The CEFU program was a large-scale effort designed to provide outbound calling services from a network of 13 call centers as a means to resolve coverage edit failures. The system was developed within a very strict time frame using many of the same Census Bureau and contractor resources devoted to the development of the TQA program. This study presented information based on debriefings of personnel involved in the CEFU program.

Census 2000 was the first use of outbound calling combined with an automated, scripted instrument to collect the necessary census data. An estimated 3.1 million cases were anticipated for the system with 2.36 million cases actually identified for followup. The program commenced on May 8, 2000 and was scheduled for completion on June 15, 2000. Telephone followup was extended through August 13, 2000 to permit followup with language difficulty cases and to maximize the number of households for which the Census Bureau could obtain a completed interview. CEFU had a 70.8 percent completion rate for cases with a valid telephone number.

The call center infrastructure and automated features used for TQA provided the foundation for the implementation of CEFU. It was known that some form of outbound calling services would be used in Census 2000 when the TQA contract was awarded to Electronic Data Systems; however, the scope and specific requirements for the program were not defined until very late in the development process. Given the time frame for development, not all requirements were implemented and testing was limited. Regardless of the tremendous pressures and issues associated with its development, the contractor and government personnel provided exceptional support and dedication to ensure that CEFU was successfully implemented.

Compromises on evaluation data requirements for CEFU were made because of the limited development time. Although these compromises were made through negotiations between subject matter experts and program managers, the lack of some data from the operation impacted the completeness of Evaluation I-1, Coverage Edit Followup for Census 2000.

These and other findings led to the following key recommendations:

- Begin development early.
- Establish realistic development timelines. For each development effort, timelines must consider the complexities associated with translating and implementing high-level user requirements into a functional system. In addition, time for testing and rework is required to ensure that each system is sufficiently stable for production.
- Establish agency-wide guidance.

Evaluation R.1.c, Internet Questionnaire Assistance System Requirements Study (Titan Systems Corporation 2001c)

IQA and its companion system, IDC, were Internet-based systems that were developed by the same team. While there was a technical linkage from an operational standpoint, for purposes of this report they will be treated as separate systems. This study presented information based on debriefings with Internet Questionnaire Assistance system designers. A separate customer service survey evaluation (Evaluation A.2.c) provides the user perspective of both systems.

The overall objective of IQA was to provide information to the public to assist respondents with completing census questionnaires. This was accomplished through a user-friendly interface which provided hyperlinks covering a variety of topics. Language assistance guides were an important feature of IQA and could be easily downloaded by non-English speaking respondents.

IQA was assessed by Census Bureau personnel as being very successful. One of the goals of the system was to lessen the burden on the TQA operation. The IQA website received 23,864,598 hits between March 3, 2000 and April 19, 2000. The system was developed under a very aggressive schedule and deployed within 18 months. The short time frame was due to an interruption of the planning effort that stemmed from high level concerns over the security implications of IDC.

IQA was positively perceived by the personnel interviewed as "the right system for the job." It appeared to provide an effective means of disseminating Census 2000 information and foreign language assistance guides to the public through a user-friendly interface. The few system shortcomings were all minor. Future versions of the system would benefit from enhancements such as email response or other feedback mechanisms.

One major requirement and design issue was the need for the systems to provide strong levels of security against unauthorized intrusion; IQA succeeded in providing this protection.

The success of the system was largely due to the involvement of a few highly talented and dedicated Census Bureau personnel. Contractor support (with the exception of penetration testing) did not play a role in the planning or development of the system.

These and other findings led to the following key recommendations:

- Establish agency-wide guidance.
- Begin early planning for 2010. Widespread Internet usage will place significant demands on the next generation questionnaire assistance system. Therefore, it is recommended that planning commence early and reflect an expectation of greater usage requirements.
- Use flexible project control structure. The use of an oversight body to guide and monitor system development activities is considered a “best practice;” however, any project control structure should be flexible enough to encourage technical innovation.
- Assess staffing risks. Historically, the Census Bureau relied on in-house expertise to develop decennial systems. In the case of IQA, a single individual was the driving force behind the development effort. Given the nature, scope, and complexity of future systems, there are risks inherent in relying solely on in-house staff without supplementing these resources with external support.

Evaluation R.1.d, Internet Data Collection System Requirements Study (Titan Systems Corporation 2001d)

Census Bureau personnel assessed IDC as being very successful in spite of limited use by the public. The overall objective of the system was to provide census respondents with a highly secure Internet filing option. IDC succeeded in replicating the paper short form’s key features. Software limitations prevented precise replication of the paper form in every detail, however, this did not impact the form's usefulness.

One of the goals of the IDC system was to ease the data capture burden on the DCS 2000; however, it did not meet this expectation as only about 66,000 forms were submitted through IDC in March through April 2000. During this time, system utilization was characterized by two very brief periods of heavy access. IDC did not achieve consistent and widespread usage owing to a lack of publicity that stemmed from Internet related security concerns at senior management levels. The system was developed under a very aggressive schedule and deployed within 18 months

IDC was positively perceived by the personnel interviewed as “the right system for the job.” It provided an effective means of capturing Census 2000 information through the submission of short form questionnaires over the Internet. The few system shortcomings were all minor. However, there were some usability problems that were not resolved.

One major requirement and design issue was the need for the system to provide strong levels of security against unauthorized access; IDC succeeded in providing this protection. This requirement interfered with including members of the public in usability testing.

Similar to IQA, the success of the IDC system was largely due to the involvement of a few highly talented and dedicated Census personnel. Contractor support (with the exception of penetration testing) did not play a role in the planning or development of the system.

These and other findings led to the following key recommendations:

- Establish agency-wide guidance.
- Begin early planning for 2010. IDC was essentially over-engineered for capacity and redundancy because the anticipated traffic levels did not materialize due to a lack of publicity. Widespread Internet use will place significant demands on the next generation data collection system in 2010. Therefore, it is recommended that planning commence early and reflect an expectation of radically different capacity and redundancy requirements.
- Assess staffing risks. Historically, the agency has relied on in-house expertise to develop decennial systems. Given the nature, scope, and complexity of future systems, there are risks inherent in relying solely on in-house staff without supplementing these resources with external support.

Evaluation R.2.a, Operations Control System 2000 System Requirements Study (Titan Systems Corporation 2002a)

The Census 2000 Operations Control System (OCS 2000) was perceived as a successful system that was deployed when needed, in spite of the lack of a standardized requirements definition process and the substantial changes in requirements that occurred during the life of the system. The overall objective of the system was to automate the management of field operations prior to and during Census 2000. This study presents information based on debriefings with personnel involved with the OCS 2000 program.

The system assigned and controlled work to all census enumerators, tracked progress of assignments, produced cost and progress reports on field operations, printed a wide variety of enumeration related materials, and assisted with the management and tracking of shipping documents. The OCS 2000 was operational between October 1997 and August 2000 and had six key interfaces with other systems, one of which was external to the Census Bureau (the FedEx interface). It provided support for several operations including NRFU, the largest single field operation in Census 2000.

Its success was particularly noteworthy in view of the impact of a Supreme Court decision which changed the focus of the decennial census from a sampling-based approach to full enumeration. The decision was handed down very late in the system development cycle. Although the Decennial Management Division provided a project oversight role to ensure that the right resources were applied to the project, uncertainty over which method the Supreme Court would favor resulted in a dual system approach (i.e., sampling and full enumeration) to development going into the Dress Rehearsal in 1998.

Primary system deficiencies centered around two areas: the need for a FedEx interface and difficulties encountered with obtaining test data. The FedEx interface worked moderately well; however, once a package was shipped, the tracking data were not available in the OCS 2000. This required operators to go to the FedEx web site to monitor shipping status. Data to sufficiently test interfaces to other systems were not always available. This lack of data was caused because some systems were still under development and some operations in the field were not yet complete.

A perceived need for expanded system access and national level status reports was noted by some users during the interviews. Such capabilities could have been implemented but were problematic and inconsistent with fundamental system objectives. The inclusion of these capabilities would not have improved overall system functionality in a meaningful way.

These and other findings led to the following recommendations:

- Establish agency-wide guidance.
- Conduct contractor orientation. The agency utilized contractor support to develop the OCS 2000. Programming and database functions were performed by contractors working in-house under the direction of the Technologies Management Office. These contractors were well-integrated with the staff, were highly competent, and performed extremely well. However, they did not always understand business practices at the Census Bureau.
- Assign contracting officer to team. The assignment of a dedicated contracting officer to the OCS 2000 team proved to be an effective arrangement that facilitated the timely handling of contractual issues. Assigning a contractor officer as part of the overall team is a best practice that should be considered for large and/or critical system development projects.

Evaluation R.2.b, Laptop Computers for Accuracy and Coverage Evaluation System Requirements Study (Titan Systems Corporation 2002b)

The use of laptop computers to collect respondent data was perceived as a highly successful platform for the Person Interview phase of the A.C.E. program. A laptop-based automated questionnaire was used by interviewers to conduct personal and telephone interviews to assist with the determination of coverage error. In addition, laptops served as a platform for a case management system and provided remote mail services. This study presents information based on debriefings with personnel involved with the Laptops for A.C.E. program.

There were two automated instruments for Computer Assisted Person Interviewing (CAPI) for A.C.E., one for the Person Interview operation and another for the Person Interview QA operation. The latter was used to confirm that the individual conducting the Person Interview had actually contacted the original respondent. For cases when the respondent had not been contacted, the Person Interview QA instrument contained a complete version of the Person Interview thereby enabling the interviewer to collect the necessary information.

The Census Bureau built upon its experience using laptop computers for survey data collection that extended back as far as 1992. The laptop program for Census 2000 effectively utilized a customized version of an existing questionnaire authoring software package to develop the questionnaire for the A.C.E. program. The laptops provided a hardware platform for a logic-based instrument that guided the interviewer through the data collection process. Interviewers liked using the laptop. A side benefit of the laptops was that they lent a professional appearance to the interviewers that may have served to reduce concerns over the release of personal data by interviewees.

There were daunting logistical and support issues related to using laptops with automated survey instruments for data collection. Nevertheless, the Census Bureau opted to use them as the platform for the A.C.E. data collection operations. Many of those issues were formidable. For example, the Census Bureau had to acquire, configure and deploy over 9,000 laptops nationwide. Once deployed, there were major accountability, training, and maintenance issues. In spite of the costs and problems posed by these issues, the Census Bureau's decision to use the laptops proved to be a good one.

The laptops were a very effective tool and perceived as the right platform for the job in that they greatly facilitated case management, accelerated the data collection process, and improved data quality. One unique facet of the laptop program was its ability to exchange data with Headquarters and ROs via a remote dial-in telecommunications session. This method was very fast, secure, and reliable.

Off path data are generated when an interviewer needs to "back-up" to make a correction to the data entered. This action, in some cases, necessitates a shift into another logic "path" pertaining to the correct response. Off path data were identified as a requirement in the planning phase; the Computer Assisted Survey Execution Software was designed to either globally keep off path data or to ignore it entirely. Off path data were important because these data were not always incorrect; instead, the data may have resulted from a legitimate interview. Procedures were established to distinguish between the data that the Census Bureau wanted to keep and those data that were not important. However, these procedures were not correctly implemented. Since all off path data were captured, the Census Bureau was able to perform post-census processing edits to restore any missing values.

These and other findings led to the following key recommendations:

- Begin development early.
- Consider full and open competition. An open competition among vendors may require substantial time and effort in the short run; however, the competitive process usually serves to mitigate risks in the long run by assuring that the vendor has the necessary capability and experience to meet project requirements.
- Identify data exchange requirements early. The Census Bureau considered the requirements for transmitting data early in the laptops program. The early identification

of the requirements helped to ensure the timeliness and accuracy of the information being transmitted and served to maximize network and machine resources by transmitting during off-hours.

Evaluation R.2.c, Accuracy and Coverage Evaluation 2000 System Requirements Study (Titan Systems Corporation 2002c)

The 2000 A.C.E. control system (ACE2000) included both tracking and communication functions. The system was used to control A.C.E. field operations and some smaller operations such as relisting, Targeted Extended Search, and QA. A.C.E. production usage began in August

1999 and continued until May 2001. This study presented information based on debriefings with personnel involved with the ACE2000 system.

The original tracking and control system was character-based, operating in a Disk Operating System (DOS) environment. A Windows-based version of the software was developed as part of the Integrated Coverage Measurement program. The Integrated Coverage Measurement 2000 system was used to control and track all Integrated Coverage Measurement field operations in addition to the Computer Assisted Personal Interview operations during the Census 2000 Dress Rehearsal. Although the Dress Rehearsal was considered a success, issues such as data volume, changing survey requirements, and concerns regarding performance resulted in the need for further system enhancements. Integrated Coverage Measurement 2000 was renamed ACE2000 when the survey was renamed A.C.E.

The ACE2000 system was considered by those involved in the study to be the “right system for the job.” The system was successful, but its success results from the dedication of the Census Bureau and contractor staff, not from a well-planned development timeline or supporting development methodology.

The development of the system was subject to continuous changes. The changes forced the development team to focus on core functionality thus preventing some requirements from being implemented due to schedule constraints.

The number of changes limited the testing that could be accomplished. Although several layers of testing were in place, testing was implemented based on a careful review of resources, time available, and risk.

These and other findings led to the following recommendations:

- Define adequate resources. Development of the ACE2000 system was subject to wholesale changes in census methodology, technology, and business process. A census is not the time to try unproven approaches or to develop systems without sufficient time for development and a sufficiently large staff of both subject matter and development personnel. The Census Bureau must plan well in advance to ensure that the necessary personnel resources are available to support the project and that those resources can

devote a sufficient amount of time in requirements definition and testing without being diverted to other activities.

- Replicate census environment more closely. The Census 2000 Dress Rehearsal environment provided for only one LCO within each of the three ROs involved in the exercise. This environment did not provide the opportunity to test and evaluate certain aspects of the software. Recommendations include that the Census Bureau establish a more complete cross-section of the business process during the dress rehearsal so that the nuances of each operation can be better tested and evaluated.

Evaluation R.2.d, Matching and Review Coding System for Accuracy and Coverage Evaluation (Titan Systems Corporation 2002d)

The Matching and Review Coding System was software designed to facilitate clerical matching of records between census data and A.C.E. data. This study presented information based on debriefings with personnel involved with the Matching and Review Coding System.

A character-based system was developed by in-house resources and used for the Census 2000 Dress Rehearsal. This system lacked sufficient functionality to meet the needs of the matching process. A decision was made to outsource development of a new system that could provide the point and click interface of Windows-based software. The Matching and Review Coding software was developed, successfully deployed, and considered by those involved in the study to be the “right system for the job.” Although some requirements changed over time and new requirements were introduced, the dedication of the contractor and agency personnel resulted in a successful system development effort. This was accomplished despite significant time and resource constraints.

The software was developed as three distinct systems. This enabled Census Bureau personnel and developers to build upon their experiences and apply critical lessons learned during the development process. Although the same general process was used to identify, clarify, and implement requirements across the systems, the procedures and means of communication were fine-tuned over time.

The software provided point and click functionality to facilitate the clerical matching process. The software was more streamlined than earlier systems allowing the clerks to move faster through the matching process. Technological improvements such as split screen and filter capabilities also enhanced the system’s usability.

These and other findings led to the following key recommendations:

- Conduct team training. A team approach was utilized to identify requirements and evaluate the software. A team approach enables the organization to solicit different viewpoints and ensure that the widest range of needs are addressed.

- Clarify roles and responsibilities. Several testing processes were implemented between the contractor, alpha test groups, and the Census Bureau Beta Test Site. Care must be taken to clearly identify the roles and responsibilities of each group and procedures must be established to ensure an effective and efficient means of sharing test results. This will help ensure conformance with requirements specifications and comprehensive coverage of testing needs.

Evaluation R.3.a, Pre-Appointment Management System/Automated Decennial Administrative Management System System Requirements Study (Titan Systems Corporation 2002e)

The Pre-Appointment Management System/Automated Decennial Administrative Management System (PAMS/ADAMS) was the first fully integrated applicant, personnel, and payroll system developed for the Decennial Census. This enterprise wide system utilized state-of-the-art client server technology to manage and distribute data to 12 RCCs, the Puerto Rico Area Office, and 520 LCOs. The overall objective was to develop a comprehensive system for temporary employees that manages the complete employment life-cycle. The system was successful at integrating processes that were previously handled by noninteroperable (i.e., independent) systems. Time constraints imposed by late funding of the project placed limits on the implementation of some requirements, however this did not affect the overall performance of the system. This study presents information based on debriefings with personnel involved with the PAMS/ADAMS.

The PAMS/ADAMS was quite large in scope. In all, there were over 3,000,000 applications processed and weekly payroll reached a maximum of over 520,000 individuals. Previous systems were developed in-house. A commercial off-the-shelf product that could fully meet the unique needs of Census 2000 was not found, however the commercial product with the best fit was used as the basis for development.

A formalized method was successfully used to identify requirements. However, there were significant obstacles to implementing an effective system, especially prior to the dress rehearsal. The development of the system was impacted by requirements changes and performance issues. Although there were significant challenges that posed high risks throughout the development process, the production system performed to the satisfaction of the Census Bureau and its stakeholders.

Although the system produced many types of reports, other real-time reporting requests were not fulfilled because of late data warehouse implementation, inability of management reporting systems to fully use feeder system data, and heavy programmer workload.

PAMS/ADAMS incorporated a user-friendly, locally operated front-end interface for capturing applicant and payroll forms. It was developed to address data capture issues that arose after the original scanning requirement was dropped. This interface was designed for a wide variety of users, taking into account the level of education, physical limitations and the ability to minimize human error. This was a major factor that contributed to the system's efficiency and success.

The success of the system was largely due to the highly dedicated team of Census Bureau and contractor personnel. This experienced, cohesive team worked together to address technical and procedural issues capably, even under the pressures associated with time constraints and mandatory deadlines. All major performance issues were resolved before deployment.

These and other findings led to the following key recommendations:

- Design for flexibility. The system's processes were highly regulated and therefore subject to change in local and national laws which affected system requirements. The development team found it difficult, at times, to implement these policy changes in the system. The design of the system ought to be modularized and be adequately sized and flexible enough to accommodate these types of changes.
- Include all stakeholders. The system produced many automated and ad-hoc reports but there were many real-time reporting requests that were not fulfilled. Reporting requirements need to be considered from many different viewpoints and stakeholders from all key areas should be represented during this phase to ensure that their information needs are met. Identification of reporting needs during the requirements phase will maximize the benefit that can be derived from the system for all users.
- Implement formal process. Modifications were performed continuously and an efficient formal change control process was employed. Additionally, a well documented System Investigations Requirement log was used throughout the entire change management process. Changes should be systematically assessed in light of programmatic goals. The requirements for change control and supporting documentation should be included in the system development methodology. The Change Control Board also must have adequate resources to address programs with large and complex scopes.

Evaluation R.3.b, American FactFinder System Requirements Study (Titan Systems Corporation 2002f)

American FactFinder (AFF) is an Internet enabled information system which provides an efficient means of making a wide range of census information (demographic, economic, and geographic) available to Census Bureau personnel and external users. IDC was designed to meet the needs of all users, from novice to expert. Due to the diversity of system users, the system interface provided an interactive and user-friendly way to facilitate retrieval and use of information and data. This study presents information based on debriefings of personnel involved in the IDC program.

A contractor played a major role in designing, sizing, and operating the system. The decision to use contractor support stemmed from the realization that Internet technology was evolving rapidly and that outside expertise was needed to successfully implement a state-of-the-art system. An iterative development process was employed using a cyclical building technique (design, build, and test) that allowed for continuous feedback and evaluation. The contractor was also a partner with the Census Bureau in the requirements definition process. The process

included gathering requirements from key stakeholders, subject matter experts, and potential users via interviews and joint application development sessions. The results of the interviews were compiled and incorporated into Use Case documentation.

The principal contractor, IBM, had two years to develop, test, and deploy the system. A subcontractor with expertise in geographic information systems and mapping applications also was brought in to support the development of the system. The first production implementation of IDC was in March 1999 and provided access to economic data, the ACS, and 1990 Census data. A second implementation came in December 2000. It provided improved performance, addressed user comments and requests concerning the user interface, and scaled up the system to accommodate workloads associated with Census 2000.

As with other systems that were developed to support Census 2000, IDC did not benefit from an agency-wide standard process for requirements definition. The requirements methodology was provided by the contractor and the agency conducted analyses of security needs and user segmentation. Overall, IDC was “the right system for the job” in that it succeeded in providing an effective, though not always easy-to-use, tool through which many different types of users could--for the first time--access census data on demand.

The need to efficiently disseminate Census 2000 data was a main driver for developing IDC. However, the system disseminates other census data that are generated by various program areas of the Census Bureau (i.e., economic censuses and surveys, demographic surveys, and the ACS). The multi-faceted nature of the system and plans for further expansion require that this system remain active. Thus, unlike the other 11 automated systems that were evaluated specifically as supporting components of Census 2000, IDC is a corporate system--not a dedicated Census 2000 system.

IDC has been a major success for the Census Bureau from the standpoint of achieving a breakthrough in the delivery of voluminous data in an electronic format and in making these data available to external users. It also achieved a reduction (though not elimination) in the use of traditional media (printed hardcopy, magnetic tape, etc.). The system holds great promise for escaping the limitations of pre-defined census data by making customized queries possible.

Confidentiality was a major design factor from the outset. Given the need to prohibit unauthorized access to confidential microdata files and to minimize opportunities for ‘re-identification’ (i.e., combining multiple data sources in an effort to equate census data with particular people), the Census Bureau has undertaken precautionary and effective efforts to ensure security and prevent unauthorized access to data.

An evolutionary approach was used to develop the system that required constant fine tuning as development progressed. Census Bureau managers and program staff were aware, from the outset, that system development would be incremental because of the delivery cycles for Decennial Census, Economic Census, and ACS data products. Thus, there was an underlying assumption for IDC that the system would have to adapt to requirements growth and the contracting approach and development philosophy. Change control processes were governed by

this awareness. Changes to requirements were initiated throughout the development cycle in keeping with the iterative development approach. Adapting to the changing requirements had significant cost implications.

Communication between Census Bureau project management personnel and the contractor was frequent, well documented, and included an effective change control process. This process was especially important in view of the prototyping approach that was employed. The Census Bureau employed the Department of Commerce concept of operations for streamlined acquisition for its procurement methodology. This approach helped to explore system characteristics and development issues through pre-award, face-to-face meetings with vendors. Contractors were encouraged to utilize commercial-off-the-shelf software as a development tool due to software maintenance and other considerations.

These and other findings led to the following key recommendations:

- Define user base early. The Census Bureau made extraordinary efforts between 1995 and 1997 to define the system user base and address their needs by conducting focus groups with internal and external customers, meeting with private sector organizations, surveying participants involved in beta testing of the system, and interviewing data users. It is recommended that the Census Bureau continue the practice of conducting customer segmentation analyses as early as possible in the system development process.
- Establish agency-wide guidance.
- Customize user interface design by user type. The system posed a major challenge to designers/developers in the sense that it needed to serve a very diverse set of users. It is recommended that future refinements of IDC consider setting up user classes to make the system suitable for novices as well as power users.

Evaluation R.3.c, Management Information System 2000 System Requirements Study (Titan Systems Corporation 2002g)

The primary goal of the Management Information System 2000 system was to serve the information and decision support needs of the decennial census. Management Information System 2000 was the first executive information system used by the Census Bureau to aggregate scheduling and budget information and was the official source of management information for Census 2000. The Management Information System 2000 was an umbrella system with two components: Master Activity Schedule and Cost and Progress. This study presents information based on debriefings of personnel involved in the Management Information System 2000 program.

The Master Activity Schedule component contains information on the scheduling and duration of all census activities and was the official entry and update point for designated Census 2000 data. The Cost and Progress component of the system uses an efficient enterprise-wide database of cost and progress information designed to enable managers to assess and modify operational

plans as they relate to the Master Activity Schedule; manage operations and cost; and identify problems quickly.

Both components of the Management Information System 2000 contained a broad variety of functions to support the management of the decennial census. Among these were analytical tools used to assess progress and management functions to aid in the decision making process. Information was presented in various forms such as reports, graphs, summaries, Pert charts, and schedules. These tools were used to manage the operation of the census.

Although the Management Information System 2000 proved to be a useful tool for determining high level cause and effect relationships with regard to the operation of Census 2000, some users indicated that they continued to rely on their own control systems and scheduling tools. These users stated that the scope and purpose of the system were never clearly defined. The basic requirements as to what data were needed to monitor programs, and at what level, were not adequately addressed.

The Cost and Progress component was not used by some program management offices and divisions. This was due in part to the complex nature of using the software and the users' unfamiliarity with current technology such as graphical user interfaces. Regular usage along with training was needed to become proficient.

In the Cost and Progress component, it appeared that a lack of standardization between systems created problems interpreting information in the reports. Some terminology and data element definitions were different between systems and reports containing aggregated data often reflected different snapshots in time.

Any system design must consider the entire operational environment including associated business processes such as those designed to maintain the currency and integrity of the data. The Master Activity Schedule component was a useful tool because it provided scheduling information at various levels of detail that enabled managers to identify issues and activities with the potential to impact the project schedule. Current information related to task management was not always available as needed because the process of updating the schedule was not effectively implemented.

Testing serves to validate that system requirements have been met. Testing for the Cost and Progress component was considered comprehensive. In addition to the unit testing conducted by developers, an independent internal tester, who had participated in the Joint Application Development sessions, was assigned to conduct alpha testing. Knowledge of system requirements gained from being involved in the Joint Application Development sessions allowed the tester to determine if required functionality had been implemented.

These and other findings led to the following key recommendations:

- Increase senior management commitment. The requirements for the Management Information System 2000, in particular the Cost and Progress component, were developed very late in the decennial cycle. The system needed consistent senior level management focus and support throughout its development and operation.
- Increase coordination. Many of the systems feeding data to the Cost and Progress component were produced independently and used varying definitions for data elements. This impacted the ability of the Cost and Progress component to receive data from feeder systems without reprogramming. For reliable results, terminology and data element definitions should be standardized between feeder systems.
- Provide sufficient resources. Both components of the Management Information System 2000 required a high level of expertise to maintain. Due to a lack of technical support resources, trainers for the Master Activity Schedule became involved in other activities such as configuring personal computers and installing upgrades. The Census Bureau should address the need for on-going technical support during the requirements process and ensure that sufficient resources are available to support development teams. This would help ensure that developers stay focused on the development and deployment of the system instead of diverting their efforts to address more general support issues.

Evaluation R.3.d, Census 2000 Data Capture System Requirements Study (Titan Systems Corporation 2002h)

Census 2000 Data Capture provided state-of-the-art hardware and software to capture census data. A scanning process created a digital image of census forms; these images passed through OMR and Optical Character Recognition (OCR) devices to capture information electronically. Manual keying was used to enter data not captured electronically. The Census Bureau outsourced the two major components of the Census 2000 Data Capture program. The two components were the DCS 2000 awarded to Lockheed Martin and the Data Capture Services Contract awarded to TRW. This study presents information based on debriefings with personnel involved in both components of the Census 2000 Data Capture program.

Census 2000 Data Capture was a significant challenge involving leading edge technologies; outsourcing of software development, hardware/software integration, and operations; extremely complex requirements; and myriad changes. Despite the challenges, the project team was successful in implementing a system that efficiently and effectively processed 150 million forms, using innovative technologies and contracting techniques to accomplish this effort.

One of the reasons for the success of the Census 2000 Data Capture program was that the project team established a cooperative relationship with both the prime contractors. In addition, the program manager made it clear to each organization that their success was dependent on each other. This fact encouraged both contractors to establish close working relationships and cooperate in the identification and resolution of problems.

The project team established a stringent change control process at the working level that served to track, evaluate, and control changes to the DCS 2000. This process pleased program managers, as it mitigated risks to the data capture program. The requirements methodology and change control process implemented were praised by oversight organizations, such as the General Accounting Office and the Inspector General. In addition, TRW used a document management system and email to track, control, and issue changes to procedures and training materials as part of the Data Capture Services Contract. Although the stringent change control process was successful, some of those individuals requiring data capture information for evaluation purposes thought the process too rigid in meeting their data requirements.

A series of operational tests and dry runs were conducted at each of the Data Capture Centers. These tests provided an opportunity to assess the integration of software and procedures and identified changes and improvements to both components. An integrated “Four Site Test” was designed to measure the system’s ability to process a large volume of information.

There was a philosophical difference between the Census Bureau QA specialists and program managers regarding the application of quality standards on the DCS 2000 component. The specialists were assured that quality measurement and corrective actions were available, but it was never clear to the specialists through the available documentation or repeated explanation, how QA would be applied and measured during the data capture process. It was unclear to the program managers why their documents and presentations did not answer the QA specialists’ questions.

The DCS 2000 component was a technological success, but was too sophisticated for regular survey efforts. In retrospect, the reuse of the DCS 2000 for non-census work may not be realistic or ideal. Reuse cannot be achieved, unless the Decennial Census itself becomes less complex or has fewer specialized requirements that necessitate system customization.

These and other findings led to the following recommendations:

- Define requirements early. Starting the planning and development earlier would provide a greater chance that all identified requirements would be implemented and that sufficient time would exist for testing and refinement.
- Take advantage of institutional knowledge. Outsourcing should provide a means to augment and extend the capabilities of in-house personnel. The corporate knowledge and understanding of census processes needs to be maintained. Experienced system developers, operational customers, and census content experts should be actively involved throughout the requirements identification and development processes.
- Develop standard project management tool kit. The project team developed a number of processes and tools to support their contract management and development activities. Ideally, a standard set of tools would be available for each team project prior to development to avoid diverting project resources from the actual contract management and system development tasks.

Synthesis of Results from the Census 2000 Alternative Questionnaire Experiment (AQE2000) (Martin et al 2004)

This report summarizes the findings of three experiments included in the AQE2000. All three experiments were limited to the mailout-mailback universe.

The skip instruction experiment examined respondent performance in following skip instructions on the census long form. It compared different ways of aiding respondents' navigation through the questionnaire. One design incorporated instructions and visual features to help respondents prevent errors before they occurred and another was designed to help respondents detect errors after they occurred. In addition to these prevention and detection designs, other potential design improvements, such as using reverse print to attract respondents' attention to instructions and rewording the standard skip instruction, were also tested.

Errors of commission (which occur when respondents incorrectly answer questions they should have skipped) were significantly reduced in all of the experimental treatments, suggesting that the design changes improved respondents' perception and comprehension of the instruction. Errors of omission (which occur when respondents skip questions they should have answered) decreased for the Detection Treatment, but significantly increased for every other experimental treatment. Either type of error indicates respondent difficulty navigating the questionnaire, but their impact is different. Errors of omission result in missing data. Errors of commission increase respondent burden and frustration. The recommendation is to adopt the Detection method in mail questionnaires, since it significantly reduces both types of errors.

The residence instructions experiment involved the presentation of residence rules on Census 2000 short form. This research aimed to improve within-household coverage by rewording the roster instructions to make them understandable to respondents, by encouraging respondents to read them through appropriate placement and formatting, and by presenting the instructions to increase respondents' willingness to follow them.

The changes in format, presentation, and wording of the residence instructions resulted in a significantly higher response to the household count question (which serves as an important indicator of missing data and flags large household followup). The experimental panel also produced significantly fewer omissions among Hispanics in the low coverage stratum. The recommendation is to conduct additional testing of the graphical and wording changes that led to these improvements, to better understand their effects and to further improve the quality of household count data.

The race and Hispanic origin experiment compared the 1990-style race and Hispanic origin questions with the new questions in the Census 2000 short form. It examined the effects of changes mandated by the Office of Management and Budget to allow the reporting of more than one race and reversed the sequence of the race and Hispanic origin items. Other changes in format, categories, and wording were also introduced in Census 2000 and the net effects of all the changes on race and Hispanic reporting were analyzed.

Overall, the questionnaire revisions substantially improved the completeness of race and Hispanic origin reporting in mail short form questionnaires. In addition, Hispanics were less likely to report their race as Some Other Race, and more likely to report as White, in the 2000-style questionnaires. Although there were no apparent questionnaire effects on the fraction reported as Hispanic, there were effects on the reporting of detailed Hispanic origin groups. The 1990-style questionnaire obtained more detailed reports of Hispanic origin than the 2000-style questionnaire, probably due to the effects of question wording differences as well as examples. Unexpectedly, there were three times as many reports of the example groups for Asian and Pacific Islander groups in the 2000-style form, which did not list examples, as there were in the 1990-style form, where examples were listed. The experiment demonstrates that some questionnaire design changes made in Census 2000 resulted in substantial improvements in data quality, but that other changes had unintended consequences. The recommendations are to carefully pretest and field test all changes to the questionnaire and to conduct similar but larger replication studies in future censuses to evaluate the effects of questionnaire changes on the comparability of data from one census to the next.

The results of all three experiments point to interactions between question format and content, suggesting that the Census Bureau must attend to the complex relationships between format and meaning in self-administered questionnaires. These factors have been demonstrated here to have measurable effects on the data. These experiments demonstrate that format affects performance on branching instructions, affects response/nonresponse on the household count question and indirectly affects coverage and that format differences between 1990-style and Census 2000 forms affect race and ethnicity reporting.

Synthesis of Results from the Administrative Records Experiment in 2000 (AREX 2000) (Judson and Bye 2004)

The AREX 2000 was an experiment designed to gain information regarding the feasibility of conducting an administrative records census or the use of administrative records in support of conventional decennial census processes. In the AREX 2000, an administrative record census was defined as a process that relies primarily, but not necessarily exclusively, on administrative records to produce the population count and content of the decennial census short form, with a strong focus on apportionment and redistricting requirements. In addition to total population counts by state, the decennial census must provide counts of the voting age (18 and over) population by race and Hispanic origin for small geographic areas, currently in the form of census blocks.

Demographically, the AREX 2000 provided date of birth, race, Hispanic origin, and sex. Geographically, the AREX 2000 operated at the level of basic street address and corresponding Census block code. Unit numbers for multi-unit dwellings were used in certain address matching operations and one of the evaluations; but generally household and family composition were not captured. The design assumed the existence of a MAF and geographic coding capability similar to that available for Census 2000.

Two sites were selected that were believed to have a total of approximately one million housing units and a population of approximately two million persons. One site included Baltimore City and Baltimore County, Maryland. The other site included Douglas, El Paso, and Jefferson Counties, Colorado. The sites provided a mix of population and housing characteristics needed to assess the difficulties that might arise in conducting an administrative records census.

The AREX 2000 had two principal objectives: to develop and compare two methods for conducting an administrative records census, one that used only administrative records and a second that added some conventional support to the process to complete the enumeration and to explore the potential use of administrative records data for some nonresponding or unclassified households that occur in a conventional census.

A two-phase process accomplished the AREX 2000 enumeration. The first, or Top-down, phase involved the assembly of records from a number of national administrative record systems and unduplication of individuals within the combined systems. This was followed by computer geocoding of street addresses to the level of census block and two attempts to obtain and code physical addresses for those that would not geocode by computer. Finally, there was a selection of “best” demographic characteristics for each individual and “best” street address within the experimental sites.

The second, Bottom-up, phase of the AREX 2000 design was an attempt to complete the administrative-records-only enumeration by the correction of errors in administrative records addresses through address verification (a coverage improvement analogue) and by adding persons missed in the administrative records (a NRFU analogue). Considering the Top-down and Bottom-up processes as part of one overall design, the AREX 2000 can be thought of as a prototype for a more or less conventional census with the initial mailout replaced by a Top-down administrative records enumeration.

There were four principal limitations on the experiment.

- The administrative records source files were limited to those used in the creation of the Statistical Administrative Records System (StARS) 1999, which relied primarily on files for tax year 1998 and other files extracted early in calendar year 1999. These files neither exhausted the national-level administrative records that might have been available for the AREX 2000 nor were they the most timely with respect to April 1, 2000, Census Day for Census 2000.
- The number of experimental sites was small. Although it would not have been reasonable or realistic to attempt to mount this first administrative records experiment in a representative sample of geographic areas large enough to make national estimates, additional sites would have provided more confidence that the results were not idiosyncratic to the sites selected.
- There was no experimental variation in key design parameters such as the clerical and field operations and the address selection algorithm. Without some factorial or fractional

factorial structure, direct estimates of operational impacts of components, individually or in combination, were not possible.

- The measurement of race and Hispanic origin in administrative records at the national level is deficient. Attempts were made to improve the measurement through the use of certain statistical models, but the results were not entirely satisfactory.

There were four evaluations: Process, Outcomes, Household, and Request for Physical Address evaluations.

AREX 2000 Process Evaluation

Basic results from the AREX 2000 Process evaluation included:

There is a reporting lag of approximately one year between the AREX 2000 source files and the target date of April 1, 2000. The reporting lag impacted the interpretation of results.

In the Maryland test site, the machine geocoding rate was approximately 86 percent, while in the Colorado site the rate was approximately 80 percent. The clerical geocoding process added about 3 percent to the number of addresses geocoded in Maryland and about 5 percent to the number of addresses geocoded in Colorado.

For the Bottom-up method, administrative record addresses were computer matched to an April 2000 extract of the DMAF. About 80 percent of Maryland AREX 2000 addresses were computer matched to at least one DMAF address, while about 81 percent of Colorado administrative record addresses were computer matched to at least one DMAF address. A clerical review of the computer matching process added an additional 4 percent of addresses in Maryland and nearly 6 percent of addresses in Colorado by clerically matching addresses to the DMAF.

For administrative record addresses that did not match a DMAF, field address verification was performed. The field verification was originally designed for 100 percent verification, but due to Census 2000 demand, the field verification was reduced to a sample basis composed of 6,644 addresses. About 13 percent of the Maryland addresses were valid as listed, while an additional 12 percent were deemed valid after the lister made minor corrections. In Colorado, about 8 percent were valid as listed and an additional 30 percent were deemed valid after minor corrections by the lister.

Time constraints did not allow for a detailed person-by-person comparison between the results of the Bottom-up method and the Decennial Census, nor between the results of the Bottom-up and Top-down methods. Although a household match was conducted between the Bottom-up method and the census, it remains an open question whether the matched addresses in the Bottom-up method contain the same people as those identified in the census. Administrative Records Research should perform an evaluation using a detailed person-by-person comparison (micro-match) of the matched addresses within the census and Bottom-up methods.

Additionally, a detailed person-by-person comparison between the Bottom-up and Top-down methods should also be pursued with regard to person and address matches.

When the Administrative Record Experiment population tallies were produced and compared to the Census 2000 tallies, the results showed that for the Bottom-up method, the five test site county tallies, ranged from 96 percent to 102 percent of the Census 2000 population tallies. For the Top-down method, the range was 84-92 percent. Based on these results, recommendations include that administrative records continue to be tested and refined as a possible supplement for future census operations. Future refinement and improvements should, at a minimum, focus on the following areas:

- Improve the computer matching and rematching processes.
- Evaluate the impact of multiple MAF IDs on the DMAF. Multiple MAF Identifiers assigned to a single address and duplicate MAF IDs assigned to multiple addresses contributed to the difficulty in classifying addresses as matched, nonmatched, or possibly matched.
- Improve the availability of source data for the under 18 population.
- Evaluate the effectiveness of computer models used in the experiment.
- Conduct further research on address selection rules used to determine a person's "best address."
- Conduct a full-scale field address verification.

AREX 2000 Outcomes Evaluation

As expected, the Bottom-up coverage is much improved compared to the Top-down. This is largely due to the completion of the Top-down enumeration by using census data for nonmatched addresses, which simulates a followup to the administrative records enumeration. Specifically, the Bottom-up coverage of children (81-94 percent across the test sites) is substantially better than the Top-down (72-83 percent). Coverage of children is a particular weakness for administrative records used in the AREX 2000.

Adults in the Bottom-up are more or less uniformly overcounted (102-104 percent). The overcount of adults most likely is due to unrecorded deaths that occurred in the 12 months prior to Census Day, the lack of special populations operations in the AREX 2000 (e.g., a group quarters enumeration), and failure to unduplicate persons after adding census data for nonmatched addresses. The latter means that there is some duplication of children as well.

Detailed enumeration results focused mainly on a comparison of the Bottom-up enumeration with the Census 2000. The analysis did not include group quarters and, due to limitations in the administrative records sources, persons could not be reported with "multi" or "other" race. The

analysis progressed from large geographic areas to small geographic areas, beginning with the five test site counties and ending with census blocks within the sites. The evaluation incorporated a variety of methods to accomplish its objectives, including univariate and multivariate statistical analyses of the AREX 2000/Census 2000 differences, and spatial/ecological maps that examined the geographic distributions of key comparison measures. The outcomes evaluation tried to disentangle the influence of demographic change and AREX 2000 processing, coverage, and data quality issues, while presenting basic enumeration statistics.

At the county level, the Bottom-up process undercounted total population in all sites except Baltimore City. As with the total population, males and females were undercounted in all sites except Baltimore City, but the female undercounts were slightly greater than male undercounts. Age groups showed more variability with most groups undercounted. Generally the size of the undercounts increased with decreasing age, except for the 20-24 age group. These patterns did not appear to be site-specific. Overcounts for the oldest old and undercounts for the youngest persons suggest that much more timely birth and death information must be obtained. Also, the special enumeration requirements for populations such as college students, the military, and persons in nursing homes must be incorporated into administrative records processes.

Administrative records are not currently a good source of data for race and Hispanic origin and the models were not sufficient to correct their deficiencies. Blacks and Hispanics were undercounted when they were a large minority group and overcounted when they were not. American Indians and Alaskan natives were not well identified and the accuracy of Asian/Pacific Islander counts was uncertain.

Bottom-up tract-level total population results indicated a good correspondence between the AREX 2000 and Census 2000. The population counts of 70 percent of tracts were within 5 percentage points and 95 percent of the tracts were within 25 percentage points, though a sizable number of tracts had moderate and large undercounts. At the block-level, population counts were the least accurate. For the total population, 38 percent of blocks met the 5 percent criterion and about 85 percent of blocks met the 25 percent criterion.

A multivariate analysis of block differences showed that large undercounts were associated with such block characteristics as high population density, high rental rates, and large proportions of persons age 20-24. Large overcounts were associated with high vacancy rates, low population density, small proportions of persons under the age of 20, and large proportions of persons age 20-24 and age 65 and over.

AREX 2000 Household-level Analysis

The general goal of the household-level analysis was to assess how well households formed from administrative records matched those from Census 2000 addresses. The evaluation focused first, on the factors associated with AREX 2000 and Census 2000 addresses that were (computer) linked. Then, demographic comparisons were made between households at linked addresses.

There was a special focus on Census 2000 households that required a NRFU visit and Census 2000 unclassified (imputed) households.

The evaluation used both descriptive analyses and logistic regression analysis to assess the coverage and accuracy of AREX 2000 households. Descriptive analyses were performed for households in all five AREX 2000 counties and for the Census 2000 NRFU and imputed households in the test sites. A logistic regression model was developed to predict the probability of an accurate household match using address and AREX 2000 processing characteristics as predictors. Addresses with a high probability of correct demographic match between occupants might be candidates for administrative records substitution in the case of NRFU in a conventional census. In the following discussion the term “linked” is used to mean a matched address. The term “matched” is reserved for household demographic comparisons at linked addresses.

AREX 2000’s coverage of the census NRFU universe was not as good as its coverage of the overall universe. AREX 2000 housing units were linked with 70.9 percent of the census NRFU housing units, compared with 88.4 percent of the census responding housing units. For occupied NRFU housing units, the coverage rate was 76.7 percent. The AREX 2000 housing units were linked with 63.2 percent of households that were imputed to have people in them and 34.7 percent of those imputed to be vacant.

The AREX 2000 and the census counted the same number of people in the housing unit for 51.1 percent of the 889,638 linked households and AREX 2000 was within one of the census for 79.4 percent of the units. The 51.1 percent is effectively a ceiling on the percent of linked households that had exactly the same persons from AREX 2000 and Census 2000. Although errors in address linkage would account for some of the mismatched households, the deficiencies in administrative records cited earlier in this report (missing children, lack of special population operations, and the time gap between the administrative records extracts and Census Day) most likely account for the major part.

For linked NRFU housing units, AREX 2000 had the same numbers of persons for 37.0 percent of the units and was within one 69.3 percent of the time. Census 2000 NRFU housing units were more susceptible to the AREX 2000 deficiencies than responding units. In addition, enumeration errors in Census 2000 might have been higher for these units.

The regression analysis demonstrated a number of factors associated with greater probability of matched household demographics. These include single unit address rather than multi-unit, household with only one or two members, all household occupants over the age of 65, at least one White occupant, and no occupant with imputed race in the AREX 2000. The predictive power of the model was moderately strong. At a predicted probability of 0.5 or higher, the probability of a correct household match was about 72 percent. At a predicted probability of 0.8 or higher, the probability of a correct match increased to about 83 percent, but the proportion of addresses with predicted probability this high was only about 4 percent of all addresses. Evidently, the limitations in the data, particularly the administrative records cutoffs and poor race and Hispanic origin measurement, made household prediction quite difficult.

AREX 2000 Request for Physical Address Evaluation

The AREX 2000 Request for Physical Address Operation collected physical addresses (geocodable house number, street name, city and state) for individuals whose address was listed as a P.O. Box or other noncity-style address from six administrative records source files. Major components of the operation were to create an address file from administrative records where the mailing address was a P.O. Box or other noncity-style address, design a form and mail it to the addresses, requesting a physical address, clerically geocode the physical addresses to state, county and block, and key addresses and geocode information to a file for further analysis.

The mailing requesting a physical address included 138,653 individuals in the administrative records files. At the conclusion of the operation, 9,431 physical addresses were geocoded, of which 8,107 were geocoded to a test site county. While the initial plan for the operation called for incorporating the geocoded addresses into the Administrative Record Experiment files, the low return rate, combined with resource limitations, led to a decision to not update the files with the information. Instead, the results were contained in a separate evaluation report.

Where respondents returned physical address information, addresses were able to be geocoded to the state, county, and block level and were defined as being in or out of the test site. The timing of the mailing contributed to a low return rate.

Unintended recipients affected the results. An underlying phenomenon regarding the mailing is that the person who received the mailing and responded may not be the person that was linked to that P.O. box in the administrative records source files. This can be attributed to the time elapsed between the date of the administrative record source file data and the mailing of the Request for Physical Address materials and the typically high turnover rental rate of P.O. Boxes.

The design of the form impacted the results. Although the form generally suited the purpose of the Request for Physical Address operation, comments returned on the form suggest some areas of the form could have been expanded or reviewed to improve the quality of information that was returned.

Based on an analysis of the results of the Request for Physical Address Operation, recommendations include the following actions:

- Assess the impact of form design and timing of the mailing in a remailing to the original addressees.
- Assess the impact of the form design and timing of the mailing by using another universe of administrative record addresses.
- Assess the need for this type of operation and determine if collecting physical addresses on individuals who have no other address type in administrative record source files merit the cost of time and resources to launch a separate operation to gather these addresses.

Overall Administrative Records implications for 2010 Census planning include:

Continue to explore the possibility of using administrative records as a substitutions for NRFU in the 2010 Census. Although the results of the household-level analysis were not definitive due to the limitations on AREX 2000, they were sufficiently strong that research into the substitution of administrative records households for NRFU or unclassified households in a conventional census should continue. For NRFU households there is the potential for significant cost savings, and for unclassified households, the potential for greater accuracy than that provided by imputation.

There are other aspects of 2010 Census development in which administrative records might play a role. These include MAF improvements, development and testing of unduplication methods for 2010, subnational Demographic Analysis, and coverage measurement research.

Arrangements should be made to acquire administrative records on a timelier basis and to obtain some data sets that might fill some of the administrative records coverage gaps.

A research agenda for 2010 could include:

- Additional evaluation of the impact of clerical and field operations in AREX 2000.
- Person unduplication in the Administrative Records Experiment Bottom-up process.
- Repeating AREX 2000 with StARS 2000 data.
- Repeating the Household-level analysis using StARS 2000 data.
- Analysis of administrative records coverage gaps, in particular gaps related to persons in group quarters.
- MAF improvements using administrative records.
- Improving address linkage techniques.
- Enhancing Numident race and Hispanic origin data using Census 2000.
- Contributing to subnational Demographic Analysis.

Synthesis of Results from the Social Security Number, Privacy Attitudes, and Notification Experiment in Census 2000 (SPAN) (Larwood and Trentham 2004)

The SPAN Experiment was designed to assess the public's attitudes on privacy and confidentiality issues related to the notion of an administrative records census and to further examine how the notification of administrative records use and the request for a Social Security Number would impact census response rates and item nonresponse rates during Census 2000. The project included a survey and a panel component, enabling both attitudinal and behavioral responses to be evaluated.

The survey component (Study of Privacy Attitudes in 2000) was conducted by the Institute of Social Research at the University of Michigan and The Gallup Organization. This component gathered information on public attitudes regarding the census, its uses, trust and privacy issues, the Census Bureau's confidentiality practices, possible data sharing across Federal agencies, and the willingness to provide one's Social Security Number. Telephone surveys were conducted with two different samples of U.S. household residents in 1999 and 2000, before and after Census Day 2000. The major analyses included (a) comparisons of the responses to those of similar 1995 and 1996 public surveys commissioned by the Census Bureau to assess long-term attitudinal trends, (b) comparisons between 1999 and 2000 responses to examine any potential effects the census environment may have had upon public attitudes, and c) the assessment of how self-reported census media exposure by Census 2000 survey respondents may have impacted their responses. Respondents' addresses were also obtained to examine how predictive respondents' attitudes were of their behavior of actually returning the Census 2000 form. Relationships between respondents' attitudes, demographic information, exposure to census publicity, and response behavior were subsequently determined.

The panel component consisted of two studies examining respondents' behavioral responses to actual Social Security Number requests and/or public notification of administrative record use. The Social Security Number-Notification study evaluated the effects of the Social Security Number request and the notification of administrative records use upon mail response rates and form completeness. The Social Security Number-Validation study focused upon the accuracy of Social Security Numbers provided by respondents and examined the effect of the request and administrative records notification upon their validation rates. Both studies used data collected during Census 2000. Ten panels were designed with different experimental treatments. The experimental cover letters and forms were the official census forms received by the sampled households, in the standard sequence and timing. All Social Security Number requests were voluntary.

In brief, the results of the Survey of Privacy Attitudes in 2000 indicated that:

- The public has steadily increased its knowledge and awareness of the census, its uses, and laws related to confidentiality practices between 1995 and 2000. The Census 2000 publicity seemed to enhance the public's knowledge and endorsement to cooperate with the census.

- Long-term survey trends showed increases in the public's belief that the Census Bureau actually protects data confidentiality; however no changes were shown in the public's trust in the Census Bureau to keep data confidential between 1999 and 2000, suggesting no effect by the census publicity upon public attitudes related to confidentiality issues.
- General privacy concerns showed a very small, yet statistically significant, decline between 1999 and 2000; however long-term trends show small increases in public concerns about personal privacy and the loss of control over personal information. The proportion who viewed the census as an invasion of privacy did not change between 1999 and 2000.
- Trends revealed that increasing percentages express disapproval towards data sharing or providing one's Social Security Number. Around 45 percent in 1999 and 2000 stated that it would bother them "a lot" if their census information was shared, a significant increase from prior years. Expressed willingness to provide one's Social Security Number declined from 68 percent in 1996 to 55 percent in 1999, with no change in 2000.
- Relationships were revealed between Census 2000 survey respondents' attitudes and self-reported exposure to census-related media. Those exposed to both positive and negative media were more knowledgeable about the census, considered it more important, and were more likely to endorse an obligation to cooperate with the census than those with no media exposure. The only negative exposure group had similar responses to those with both positive and negative media exposure, while more differences were shown between the only positive exposure group and those who reported exposure to both types of census-related media.
- Attitudes were shown to predict respondents' behavior, with high privacy concerns, negative views on the Census Bureau's confidentiality practices, disapproval of data sharing, and a lack of willingness to provide Social Security Numbers, being reliable negative predictors of whether respondents returned their Census 2000 forms and provided mailing addresses that could be used to determine the return status of their forms. Using reported demographics, non-White respondents were shown to be less likely to return their forms.

The Social Security Number-Notification panel study results revealed that:

- The Social Security Number request for one or all household members decreased mail response rates, yet the decreases were smaller than expected. Specifically, results suggested that the Social Security Number request for all household members would decrease response by 2.1 percent in high census coverage areas and 2.7 percent in low census coverage areas compared to no request. The difference between the drop in response rates of the high and low coverage areas was not statistically significant.
- The Social Security Number request for all household members was associated with more missing data, yet there was no effect shown for Person 1.

- Taken together, specific and general notification of administrative record use was shown to decrease mail response. Separately, however, specific notification did not demonstrate the predicted stronger effects than the general notification. Furthermore, there was not sufficient evidence to conclude that notification further discouraged response in the presence of a Social Security Number request compared to notification alone.
- Notification was not shown to affect item nonresponse rates, whether the two notification types were grouped together or examined separately. Further, there were lower responses to the Social Security Number item for Person 1 when the request was made without notification (contrary to prediction). This occurred regardless of whose numbers were requested (Person 1 only versus all household members) and regardless of the notification type. Also, there were no individual effects upon form completeness by type of notification.

Finally, the Social Security Number-Validation panel study results showed that:

- There was a high degree of accuracy for the provided Social Security Numbers, with an overall match rate of 94.8 percent between the provided numbers and Census Numident file (provided by the Social Security Administration). Only 5.2 percent of the reported Social Security Numbers were considered invalid.
- The valid Social Security Number rates for high and low coverage areas revealed a small, but statistically significant, 2.4 percent difference between the accuracy rates of respondents' reported numbers within the two coverage areas (high, 95.2 percent, and low, 92.8 percent).
- The valid Social Security Number rates for Person 1 were not affected by whether a Social Security Number request was made for Person 1 only or all household members. Person 1 valid rates were high across the panels (about 96-97 percent). Results also revealed patterns of decreasing validation rates for Person 2, Person 3, and so on through Person 6 among the panels that requested numbers for all household members. Nevertheless, their valid rates, were high with a range of over 95 percent to the lowest rate of 80.2 percent for Person 6.
- Notification of administrative records use had no effect upon the validation rates of provided Social Security Numbers for Person 1. Also, there were no differences between the valid rates of those who received the specific notification type versus the general notification type.

Based upon the findings of the three studies, the following recommendations were made:

- Design research that further explores public attitudes on privacy, confidentiality, and trust in the Census Bureau, and tests more effective ways to address these issues in future publicity efforts.

- Assess the potential impact of September 11, 2001 (and the extra security concerns that followed) upon public attitudes.
- Conduct qualitative research with members of targeted population segments that show lower mail response rates, less acceptance of data-sharing, and less willingness to provide Social Security Numbers, to better understand their perspectives and reservations.
- Design research to identify other, currently unknown barriers to census responses (besides the attitudes and demographics identified in the present analysis) to ultimately reduce them.
- Conduct more research on the effect of Social Security Number requests upon response behavior that further examines: the characteristics of households that provide and do not provide numbers, the accuracy of households reconstructed from administrative records, and the effect of having and not having the number in household reconstruction.
- Perform research that focuses upon the Social Security Number requests of all household members to identify factors other than attitudes (e.g., practical barriers), that may contribute to the nonresponse rates of Social Security Number requests, and develop new techniques that may overcome these non-attitudinal factors.
- Design research to further examine the effect of general and specific notification upon response behavior by considering other interpretations of how they may be viewed (e.g., justifications), and by developing new methods that further establish the relationship between notification treatment conditions and behavior. Future research also needs to assess if providing information on the use of Social Security Numbers does not markedly decrease response rates and improves validation rates, as this may change future censuses.
- Develop research to assess the cumulative nonresponse to Social Security Number requests (i.e., unit nonresponse, item nonresponse, and invalid rates) to obtain an indicator of the extent to which matching to administrative records could take place.
- Conduct a cost/benefit analysis that fully assesses all implications, should the Census Bureau consider asking census respondents for Social Security Numbers in future decennial censuses. Future research could also document the use of other identifiers that are used to link files with fewer costs.

Synthesis of Results from the Response Mode and Incentive Experiment in 2000 (RMIE) (Caspar 2004)

The RMIE investigated the impact of three computer assisted data collection techniques on the response rate and data quality in Census 2000. The three techniques were CATI, Internet, and IVR. Households participating in the study were randomly assigned to six panels and to a control group. The households in the six panels were given the choice of providing their Census 2000 data via the usual paper forms or by one of the alternate computer-mediated response modes. Half of the panels were offered an incentive, a telephone calling card good for 30 minutes of calls, for using the alternate response mode.

In addition, the experiment included a nonresponse component designed to assess the effects of a promised incentive and alternative response mode options on response among a sample of census households who failed to return their census forms by April 26, 2000. The intent of the nonresponse component was not to test incentives or response mode options as possible nonresponse conversion techniques for the census. Rather, the experiment was designed to test the effect of these factors on response among a group representing those who are traditionally difficult to enumerate.

A final component of the experiment involved interviewing households assigned to the Internet mode (both with and without the incentive) who opted to complete the traditional paper census form to determine why these households did not use the Internet.

Results from the initial mailout portion of the RMIE showed:

- CATI brought about a small but statistically significant improvement in the overall response rate. It also had a low item nonresponse rate.
- The Internet mode yielded relatively high data quality. The benefits of this data collection method may outweigh the costs.
- The implications of this study are complex for the use of the IVR technology. Data quality was the lowest for this mode. Respondents appeared to dislike lengthy surveys with this method and some respondent sub-groups (mixed race respondents and Hispanics) were more likely to report confusion with the task.
- The calling card was very effective in promoting the use of the alternative response mode. However, rather than encouraging more households to participate, the incentive tended to redirect households that would have responded by mail to the alternate computer-mediated response mode. This effect may be partially attributable to the colorful inserts in the household mailing that directed attention to the calling card.

Results from the nonresponse component of the RMIE showed that:

- CATI elicited the highest response from Census nonrespondents (7.8 percent) followed by the IVR Questionnaire (4.8 percent) and the Internet (3.7 percent). This comparison is confounded by the fact that Internet access may be especially problematic for this target population.
- Respondents to the IVR mode are significantly younger and reside in households with, on average, fewer people than both mail and Computer Assisted Telephone Interview respondents. Computer Assisted Telephone Interview respondents are disproportionately Black, with more households residing in low coverage areas compared to Internet respondents.
- The calling card incentive increased response to the alternative modes by 1.9 percent across all response modes.
- Person 1 in households receiving the incentive due to alternative response mode participation tended to be younger than Person 1 in households not receiving the incentive.
- Contrary to past research, the increase in response due to the incentive is not statistically different in areas with high concentrations of Black and Hispanic populations and renters (1.9 percent) from other areas (2.0 percent).
- When total response to an experimental second mailing is considered, no significant incentive effect remains. That is, when mail responses are included as respondents, the incentive group (13.8 percent) is no more likely to respond than the nonincentive group (13.2 percent). Similar to the initial mailout experiment, it appears that the incentive merely redirects responses that would have otherwise been obtained by mail to alternative modes.
- Irrespective of the experimental treatments, around 13 percent participation was obtained from cases that did not initially return the questionnaire or returned the questionnaire late. Replacement questionnaires were not included in the second mailing, implying that respondents who returned a mail form used their original questionnaire.

Finally, results from the Internet Usage Survey indicated that:

- Approximately 63 percent of the Internet Usage Survey sample reported having access to the Internet. Thus, access does not appear to be a major reason why these census respondents did not opt to complete their census form via the Internet.
- Nearly half of the Internet Usage Survey respondents reported they were unaware that the Internet was an option for completing their census forms.

- Among respondents who were aware of the Internet option, 35 percent reported that they believed the paper census form would be easier to complete. Other reasons for not using the Internet included: no access to a computer, concerns about privacy, forgot the Internet was an option, and insufficient knowledge of the Internet.
- Respondents reported that an incentive to complete the census via the Internet would have encouraged them to use this alternative mode. About 41 percent of respondents who were not offered the incentive or were unaware of the offer said they would fill out their census form via the Internet if they were offered a 30 minute calling card. Another 9 percent indicated they would do it for a 60 minute calling card, and an additional 12 percent would be willing if a 90 minute calling card was offered.

Based on the RMIE, the following were the overall results:

- The Internet is an attractive alternative data collection mode for the decennial census.
- The use of an incentive was an effective means of promoting the use of the alternative response modes. However, some of this effect may be attributable to the use of the insert which drew the respondent's attention to the availability of the alternative mode.
- Data quality was improved for the CATI mode (as compared with mail). However, this mode entails substantial cost investments for hardware, software, and programmer and interviewer time.
- Without significant improvements in the voice-user interface, the IVR technology is probably not a viable alternative for Census 2010.
- The use of alternative response modes did not increase overall response rates to the census.

Results from the RMIE suggest several areas worthy of future research:

- Research is needed to determine the best ways to present the response mode alternatives, as it appears that some respondents assigned to the no-incentive treatments did not read the letter that accompanied their paper census form informing them of the alternative mode option. The use of a colorful mailing insert, irrespective of whether an incentive is offered may be enough to attract respondents to an alternative census mode. However, this information cannot be determined from the data obtained from this experiment.
- Research is needed to determine whether recent advances in speech recognition software can improve the voice user interface to increase data quality and eliminate some of the dissatisfaction voiced by respondents who answered the IVR Questionnaire satisfaction survey.

- The choice of incentive should be revisited. Based on the number of respondents who never used their calling card once they were activated, it appears that the card may not have been a powerful incentive.

A Demonstration of the Operational Feasibility of the American Community Survey (Griffin and Obenski 2001)

The ACS is indispensable to the successful reengineering of the 2010 Census design. To meet the challenges of rapid demographic and technological change and in response to stakeholders' requests, Census Bureau managers have concluded that the design of the decennial census must be simplified and long form data collection must be more timely. Therefore, rather than occurring as part of the decennial census, collection of demographic and socioeconomic data will be ongoing throughout the decade via the ACS.

The formal program to develop and test the ACS began in 1994. Since then, the program's name, size, and scope have been evolving in preparation for full implementation in 2003. The full set of testing and developmental activities starting in 1994 are referred to as the ACS development program in this report. Key results will be documented in a series of reports.

This initial report focuses on the desirability and operational feasibility of full implementation of the ACS. In 2000, ACS methods were employed in 36 counties. A nationwide survey – the Census 2000 Supplementary Survey (C2SS) – was conducted in an additional 1,203 counties. The primary purpose of the C2SS was to demonstrate the operational feasibility of collecting long form data at the same time as, but in a separate process from, Census 2000.

The successful implementation of the C2SS during Census 2000 demonstrated that full implementation of the ACS is operationally feasible. Operational feasibility means that C2SS planned tasks were executed on time, within budget, and that the data collected met certain basic Census Bureau quality standards. Despite competition from Census 2000 for resources and lack of experience with a nationwide workload, staffing was sufficient, operations were carried out as anticipated, and observed response rates were high. To determine whether a nationwide implementation would adversely affect operational performance, comparisons were made between the 1999 and 2000 results from 36 counties. Based on the results of the C2SS, managers are confident that the full ACS can be successfully implemented nationwide in 2003.

Specific findings include:

- The ACS will improve planning and simplify the 2010 decennial census design.
- Implementing the ACS, supported by the MAF/TIGER modernization, will potentially improve decennial census coverage.
- The ACS has been designed to collect detailed demographic and housing information.

- The ACS development program provides current, timely information essential for governing.
- Communities continue to benefit from the ACS development program.
- The ACS development program is improving the federal statistical system.
- Workload projections for the C2SS were achieved.
- An effective strategy allowed the C2SS to collect much of the survey data by mail.
- TQA and Edit Followup activities were completed, although staffing limitations caused some delays.
- The QA process for data entry ensured accurate capture of mail returns.
- Computer Assisted Telephone Interviews were conducted on schedule, successfully reducing the workload for personal visit followup.
- Personal visit interviewing was completed on schedule, with a high rate of response.
- Total survey response rates remained high in 2000.
- Timely release of C2SS data products occurred as expected.

Nationwide implementation of the ACS via the C2SS in Census 2000 was operationally successful and demonstrated that full implementation of the ACS is operationally feasible. Based on these findings, the Census Bureau should fully implement the ACS in 2003.

Ethnographic Studies

There are three studies included in this experimental category.

Complex Households and Relationships in the Decennial Census and in Ethnographic Studies of Six Race/Ethnic Groups (Schwede 2003)

Household structure has changed dramatically in the past fifty years. In the 1950s, the most prevalent household type was the “married couple with children” household, also known as the “Ozzie and Harriet” household. As a result of important trends since that time—such as increases in immigration, changing migration streams now coming predominantly from Latin America and Asia, rather than from Europe, increases in divorce, remarriages, blended families and cohabiting couples, and children living with them, and increases in grandparent-maintained households and nonrelative households—household structure has diversified, and this type of family is no longer the modal type of household. Recognizing that household structure is

changing and that it varies among different race/ethnic groups and over time, the Census Bureau funded exploratory ethnographic research to learn more about non-nuclear, or complex, households and to identify ways to improve enumeration of them.

This study identified and described complex households in selected ethnic groups in the U.S. The specific ethnic groups studied include Korean immigrants in Queens, New York, Latino immigrants in central Virginia, African Americans in southeastern Virginia, rural non-Hispanic Whites in western New York, Navajo Indians on an Arizona reservation, and Inupiat Eskimos, known as the Inupiat, in Alaska. These studies cover five of the six main race and ethnic categories mandated by the Office of Management and Budget for use in federal data collections.

This ethnographic research project had three aims. The first was to explore the range and functioning of complex households within different ethnic groups. The second was to examine how well the response categories of the decennial relationship question capture the emerging diversity of household types in this country. The third aim had three components: to assess how well census methods, questions, relationship categories, and household composition typologies describe the emerging diversity of household types, to suggest revisions to the relationship question and response categories for the 2010 Census test cycle, and to call for new research.

“Complex household” is a research category, not an official Census Bureau type of household. For the purposes of this study, a complex household is defined as a non-nuclear family household, including nonrelatives, such as roommates and unmarried partners; more distant relatives not listed on the census form, such as nephew/niece, cousin, brother-/sister-in-law; persons shared across households, such as children in joint custody arrangements and persons tenuously attached to more than one housing unit, and more than one family sharing a housing unit.

The report includes an introduction that gives background information on four topics. The first is a description of the purposes for including the relationship question on the census form. The second is a short history of how the relationship categories have changed in decennial censuses since 1970 to reflect ongoing changes in society. The third is a comparison of relationship questions and categories among major Census Bureau surveys and censuses and the fourth background topic includes an explanation of how the household type variable is constructed from the relationship question data and a brief overview of the five basic household types included in ongoing Census Bureau publications on household structure.

The overall project was designed to have experienced ethnographers already immersed in six different race/ethnic communities conduct exploratory ethnographic studies of complex households using the same methods and the same core questions at the same time in the late spring of 2000. Twenty five complex household interviews were conducted in each ethnic community between May and July of 2000, as soon after Census Day (April 1, 2000) as possible without adversely affecting ongoing Census Bureau NRFU interviews. A new African American study was commissioned in 2002 as a result of methodological issues with the first study done in 2000.

Five major themes have been identified in this research. They include: 1) issues with the relationship question and the household type variable; 2) cultural, linguistic, and nationality differences with census concepts, methods, and procedures; 3) conceptual differences in the definition and application of the key census concept of “household;” 4) mobility patterns and respondents’ conceptions of who is a household member that may not match the fundamental census residence rule concept of “usual residence;” and 5) fear and mistrust of the government and pledges of confidentiality.

The method of asking for relationships only with respect to Person One has three limitations identified in this study. First, interrelationships among other persons in the household can be masked and not be identifiable either from the census form itself or in the data produced. Second, the classification of household type may change, sometimes dramatically, depending on who is listed as Person One, possibly distorting the distribution of household types that are used in developing programs, implementing the poverty definition, and allocating funding. Third, Person Two may not be the biological parent of a coresident child. There is a way to overcome these problems resulting from collecting relationships to Person One only. Recommendations include development and testing of an individual-level question, along the lines of the England census form question to identify all interrelationships in the household.

The number and types of relationships that are specified in stand-alone response categories set limits on the types of complex households that can be identified. The relationship categories used by the Census Bureau reflect the relationships in society deemed most important to specifically delineate at the time of each census as well as norms for household composition. These categories express relationships based on kinship, marriage, and cohabitation, and on economic (e.g., housemate/roommate) and/or legal ties (e.g., adopted child, foster child). These categories can and do change over time, reflecting changes in society. In decennial censuses since 1970, the number of relationship categories printed on the form has increased, improving the ability to identify variation in lineally extended (multigenerational) households. Additional categories are needed to enable us in the future to identify laterally extended households (with brother-in-law/sister-in-law, nephew/niece, uncle/aunt, and cousin, as these types of relatives appear to be more common among non-White ethnic subpopulations, which are growing rapidly. Results from this ethnographic study of complex households in six race/ethnic groups identify both *lineal* and *lateral* extended family households.

The absence of definitions or instructions for cohabitators on choosing proper relationships for partners may lead to inconsistencies in marking “husband/wife” and “unmarried partner” congruent with the official definitions, possibly reducing the quality of data on relationships and affecting the distribution of household types. There are three factors that may contribute to this: the first has to do with the categories on the form not fully operationalizing the official concepts (e.g., should common law partners mark husband/wife or unmarried partner?); the second has to do with different cultural interpretations of the meanings and connotations of “husband/wife” and “unmarried partner” in some subpopulations; and the third has to do with social desirability and consequent unwillingness of some cohabiting persons to mark “unmarried partner” on a questionnaire. The extent that respondents mark categories other than “unmarried partner,” may result in inaccurate counts of married couples relative to other families.

Relationship categories are not always mutually exclusive. When more than one relationship category can be marked, the choice of one or the other may cause household type to vary and sometimes to be masked.

The second major theme running across the component race/ethnic studies is that there are cultural, linguistic, and perhaps nationality differences with census concepts, methods, and procedures that need to be identified, explored and taken into account when developing forms, methods, training, and procedures. Most of the body of this ethnographic research report was devoted to ethnographic descriptions of the six ethnic groups included in the complex households study. Some of the important cultural, linguistic, and nationality differences identified in the body of this report that may affect the accuracy of counts and household data include: Latino naming customs, Navajo matrilineal kinship system and different kinship terms, Inupiaq customs of grandparents formally or informally adopting their grandchildren, and translation issues in developing foreign language versions of the census form.

The third major theme crosscutting these studies is a mismatch between the census definition of “household” and the definitions of respondents in different ethnic and cultural groups that may lead to miscounting and misclassification of household types. The Census Bureau definition basically says that a household consists of all of the people who live in one housing unit. The number of households therefore equals the number of occupied housing units. In this study, many Navajo and Inupiaq respondents did not identify households in terms of shared physical structure, but rather on the basis of sharing of domestic functions. Emotional closeness is also a key component in determining who is part of one’s household. The ethnographers documented cases of “households without walls” where persons from more than one housing unit identify themselves as one household as well as the converse: people sharing one housing unit who consider themselves to be separate households. This ambiguity in the boundaries of “household” has been documented by anthropologists, sociologists, economists, and others.

The fourth major overarching theme in all six race/ethnic studies is that mobility patterns can lead to ambiguity between the household membership status from the perspective of the respondents and the official membership status according to the census residence rules. The ethnographers identified the following mobility patterns for households in this study: long-distance cyclical mobility to and from Navajo and Inupiaq households for temporary wage labor jobs; cross-national cyclical mobility between households in Latin America and Latino households in Virginia for wage labor jobs; seasonal cyclical mobility for subsistence activities among the Inupiat or to escape cold winters among rural Whites called “snowbirds;” mobility for purposes of higher education, found in most of the samples; frequent movement of children among households for the Navajo and Inupiat for schooling and other purposes and for joint custody among rural Whites; cyclic movement of elderly persons between their own houses and their relatives’ houses (Navajo) and among households of adult children (African American); sporadic movements of tenuously attached persons (African American); and temporary ad hoc moves of indeterminate length into the houses of sick and/or elderly relatives who can no longer manage for themselves (rural Whites and Inupiat). The nature and duration of such moves as well as the anchor household respondents’ interpretation of who is a household member may cause ambiguities in determining where a person should be counted in the census. It is

sometimes not easy to apply census residence rules to determine where mobile persons should be counted.

The fifth and final theme was fear and/or mistrust of government and its pledges of confidentiality. This theme ran through the recruitment, completion of mock census forms by respondents, and interviewing in the Navajo, immigrant Korean, immigrant Latino, and African American ethnographic studies. This may relate to discussions of correlation bias in coverage evaluations resulting from persons being missed in both the census and in the coverage followup study.

Key recommendations are made in five areas: revisions to and pretesting of the relationship question; new research on the relationship question and household type; language and translation issues; outreach and training; and new ethnographic research related to coverage and residence rules.

Recommended Revisions to and Pretesting of the Relationship Question

- Expand the number and precision of response categories in the relationship question to reflect the growing cultural diversity of this country and its household composition.
- Add niece/nephew, aunt/uncle, cousin, brother-in-law/sister-in-law, and grandparent as specific response categories for relatives in the relationship question to better reflect the range of complex households.
- Add “child of unmarried partner” as a specific nonrelative response category to obtain more accurate information on the numbers and types of unmarried households with children.
- If consistency between the census short form and the ACS form is a priority for 2010, modify the new ACS question to match the decennial short form, not vice versa.
- Design research and conduct semi-structured interviews on an expanded list of relationship terms, develop new wording for relationship terms, particularly for persons in custodial care, and conduct cognitive testing.
- Conduct a split-panel test using three alternative versions of the relationship question and response categories in the 2005 National Census Test.

Recommendations for New Research on the Relationship Question and Household Type

- Design new research to develop and test an individual-level question on a mailout form to identify all interrelationships in the household, not just relationship to Person One.
- Design quantitative and qualitative research to assess how accurately the relationship categories of “husband/wife” and “unmarried partner” differentiate married couple, male

householder, and female householder household types by comparing with actual marital status.

Recommendation on Language and Translation Issues

- Increase the scope and size of the new “Language Research” and “Translating Demographic Surveys” projects to identify linguistic, cultural, cognitive, and methodological issues in developing foreign language versions of census and survey forms and develop and test improved foreign language forms.

Recommendations on Outreach and Training

- Expand outreach efforts and develop new outreach messages to immigrant Koreans, immigrant Latinos, and immigrants from other countries to maintain and improve coverage levels of the foreign born in the post-September 11th era.
- Develop special training modules for enumerators on American Indian reservations and in Alaska Native areas that identify cultural factors that may affect the way respondents interpret and answer census and survey questions and provide instructions and procedures on how to help respondents “translate” their answers into the appropriate Census Bureau categories.
- Plan and conduct targeted ethnographic research in other American Indian and Alaska Native tribal areas to identify cultural-specific factors that may affect the quality and comparability of data with other ethnic groups and develop enumerator training guidelines to address these factors.

Recommendations for New Ethnographic Research Related to Coverage and Residence Rules

- Plan and conduct ethnographic studies of household composition, residence rules, and coverage by race/ethnic groups in conjunction with the 2006 Census Test.
- Develop and conduct research to identify and assess reasons persons in different ethnic groups and of different ages might be missed in both the census and in subsequent followup coverage studies to reduce correlation bias in coverage estimates.
- Conduct research on Latino naming customs and what surnames they write on census forms to assess the extent to which Latinos vary in which surname they record on the last name line and identify possible effects of variation on matching and duplication and omission rates with non-Hispanics.
- Plan and conduct new research on persons who have more than one residence and/or more than one P.O. box to identify factors determining where they wish to be counted and why.

Generation X Speaks Out on Civic Engagement and the Decennial Census: An Ethnographic Approach (Crowley 2003)

This research was part of a larger program of ethnographic research, known as *Ethnography for the New Millennium*, conducted by the Census Bureau. The purpose of this project was to develop social science insights that may help improve Census Bureau survey response rates and population counts.

The premise of this research was that response to surveys is motivated by a respondent's sense of civic responsibility. The main goal of this study was to investigate shared attitudes among Generation X about civic engagement and community involvement, government in general, and decennial census participation in particular. Participants of this study were drawn from "hard to reach" respondent populations, such as ethnic minorities, lower socioeconomic classes, immigrants, and alienated young adults who are all members of the birth cohort Generation X. For the purpose of this research, Generation X is defined as persons aged 21 to 32, that is, respondents born during the years 1968-1979. Various studies define Generation X differently by age, with some analyses categorizing persons born in 1961 as the cohort's oldest members, while others use a younger upper boundary to demarcate the age group. Only in hindsight will the boundaries for this cohort become clearer. The wider Generation X populace, according to past studies, tends to be apathetic about community and political involvement and disillusioned with government. If Generation X respondents in this study shared such attitudes as their wider Generation X counterparts do, then the Census Bureau faces another major obstacle in reaching out to them. This apathy and disillusionment with government also compounds existing enumeration barriers identified by past ethnographic research and may have short and long term implications for survey nonresponse issues, undercoverage challenges, privacy and confidentiality concerns, and effective outreach campaigns.

Research findings and recommendations are based on 150 semi-structured, individual ethnographic interviews, ten focus groups, a paper-and-pencil survey, and participant observation activities in diverse settings such as American Indian Pow Wow ceremonies, coffee bars, community demonstrations, class rooms, pool halls, job sites, and bowling alleys. The scope of this research was nationwide and was conducted in Oregon, Illinois, Florida, Texas, Maryland, Virginia, and Washington, DC. Recruitment for this research was nonrandom, and primarily by means of snowball sampling. Recruitment targeted 25 African Americans, 14 first-generation Afro-Caribbean Immigrants (Haitian and Jamaican), 20 American Indians (on and off reservations), 19 Southeast Asians (Cambodian, Laotian, and Vietnamese), 59 Hispanics (Mexican, Cuban, Puerto Rican, and Nicaraguan) and 13 non-Hispanic White Americans. Respondents were primarily working class adults with levels of education that ranged from high school dropouts to those in pursuit of PhDs.

This research suggested that the following factors may contribute to decennial noncompliance and undercoverage errors:

Non-citizenship status or unstable immigration status

- The distinction between citizens and non-citizens is not nearly as important as the distinction between undocumented immigrants and documented individuals in an attempt to convince respondents to comply with Census enumeration efforts.
- Individuals with an insecure immigration status were much less likely to trust the government and specifically less likely to respond to the census. This research demonstrated that respondents with irregular immigration statuses are unlikely to directly cooperate with the census.
- Immigrant-centered community-based organizations are important conduits that attract distrustful and growing undocumented populations in the U.S.

Respondents not knowing about or understanding the decennial census

- Generation X respondents in this study value the decennial census as important and worthwhile. While most respondents were familiar with the census by name, most respondents were unclear about the roles and functions of the decennial census. However, most new immigrants and African American respondents in the study were significantly less informed about the census.
- Respondents either expressed a complete lack of faith in the notion of Census Bureau privacy and confidentiality statements or were ambivalent about such Census Bureau promises. In spite of perceived risks, most respondents were still willing to divulge personal information.

Increased levels of distrust among respondents towards the government

- Skepticism and mistrust towards the government was pervasive among this group of respondents. Respondents' past negative experiences and interactions with federal bureaucracies influenced their overall negative attitudes towards the government.
- Although respondents in this study possessed unfavorable attitudes towards the government, derogatory views were not extended towards the Census Bureau.
- Respondents were still willing to comply with decennial enumeration efforts because they believed the social importance and benefits of the census outweigh distrustful attitudes held towards the government.
- Respondent distrust of police and law enforcement agencies is a link that serves as a deterrent to Census participation.

Miscellaneous findings

- Minority Generation X respondents in this study generally did not confirm many of the negative stereotypes that surround all Generation X cohort members.
- Family (including family values and family relationships) is seen as the most important social institution among this age cohort; family seems to represent the most stable sense of community for Generation X respondents.
- Religious activities were extremely important among all respondent groups in this study with the exception of non-Hispanic White Generation X respondents (a small percentage of the respondents in this study). For non-citizen immigrant respondents in the study, the church and its parishioners represent their 'local' community.

The recommendations that follow are based on sound, empirical, ethnographic research provided to the Census Bureau to help make informed planning and policy decisions with regards to the 2010 decennial census. Key recommendations are:

- Downplay the national nature of the census; emphasize making a difference through Census participation in one's local community without over promising benefits.
- Stress benefits of the decennial census to families; promote the census as a family activity that will also help one's ethnic community.
- Continue to partner with church and faith-based organizations to reach special population groups, especially immigrants. Recruit and hire church youth groups for outreach and enumeration work.
- Conduct further research to determine if the Census Bureau should only emphasize the distinction between citizens and *non-citizens* in its advertisements and outreach efforts to various respondent groups or if the Census Bureau should also emphasize that personal information is confidential for *documented* and *undocumented residents*.
- Continue to focus on all undocumented immigrant population groups in educational and outreach campaigns.
- Publicize that the decennial census is an easy way in which a generation can give back to the community while empowering the community.
- Collaborate with Immigration and Naturalization Services to incorporate one or two decennial census questions on the Naturalization Civics Exam sponsored by the Department of Justice, Immigration and Naturalization Services.

Privacy Schemas and Data Collection: An Ethnographic Account (Gerber 2003)

This ethnographic research examined a broad range of respondent concerns about privacy and confidentiality. It examined the factors that respondents take into consideration when they are asked to reveal information about themselves across a variety of venues. The aim of this research was essentially descriptive.

A total of 120 interviews were carried out. Thirty-nine interviews were carried out in Phase I with respondents who had participated in at least one CPS interview. An additional 81 interviews were carried out in Phase II, with respondents recruited by local organizations (many of which had been partnership groups in Census 2000) and by other contacts. Over both Phases, 37 non-Hispanic White, 21 African American, 17 American Indian, 14 Asian, 3 Pacific Islander, 23 Hispanic, and 5 respondents who self-identified as more than one race were interviewed.

Semi-structured research protocols were designed to be administered by a team of ethnographers. The interview used flexible probes. The topics included debriefing about Census 2000 and CPS participation, experiences with other data collections, privacy attitudes, and a series of vignettes. These vignettes served to expand the set of circumstances under discussion to include things of particular interest to the research.

A descriptive model for understanding how respondents decide whether to divulge information was created. This model included three main parts: an assessment of the sponsor of the questions, an assessment of whether the questions are relevant to some legitimate purpose of the sponsor, and an assessment of risks and benefits of divulging information.

- In assessing sponsorship, respondents want to approve of the sponsoring agency.
- Respondents are also aware that sponsorship may be misrepresented; thus, the authenticity of the data collection is an issue for respondents. This makes telephone mode interviews highly unpopular, because it is impossible to be sure of a caller's true identity.
- Respondents form expectations of what questions are legitimate for the sponsor to ask, based on their understanding of the nature and purpose of the survey and the sponsoring organization. Questions that go beyond this framework of expectation may be perceived as intrusive.
- Respondents evaluate the risks and benefits of providing information. All respondents are familiar with exchanging information to receive particular benefits, for example, providing information to lenders or to social service agencies to receive services. Respondents are also motivated by altruistic benefits, such as providing information to the decennial census to enable services for a local area. They also may see participation in surveys and censuses as a way of bringing a group or a point of view greater attention. This is called "having one's voice heard." It was a powerful motivation for Latinos and American Indians to participate in the census.

- Respondents also worry about the risks of divulging information about themselves and their families.
- Respondents are also concerned about government agencies, such as police departments, the Internal Revenue Service, the Immigration and Naturalization Service, and, for American Indian respondents, the Bureau of Indian Affairs. These agencies are not seen as benign.
- Because respondents believe that information is freely shared between agencies, despite any assurances of confidentiality, if they have something to hide, they are reluctant to provide it to any government agency.

This model of how respondents decide to reveal information in censuses and surveys is widely shared in all groups. However some probable differences did emerge. One difference was between more and less technologically sophisticated respondents. Technologically sophisticated respondents were more comfortable with providing information on the Internet, and felt more able to deal with any potential problems that might occur. Simultaneously, however, such respondents often did not believe that it was possible for any institution to completely assure privacy or confidentiality to persons providing information. Differences in privacy sensitivities also emerged for groups that had negative experiences with the Federal Government or with particular agencies of government, (e.g., legal immigrants who experienced difficulty in crossing the border).

Recommendations based on this research include:

- It is not possible to create a list of items that will always or never be considered private because privacy judgements are situational.
- Be careful how disparate topics are combined in one survey setting because the sense of intrusiveness of questions is situational.
- Include the idea of having one's voice heard in motivational material for minority groups.
- Describe the Census Bureau's internal controls on the handling of data in explanations of confidentiality.
- Provide good, understandable explanations of why data are needed and how they will be used because respondents' comfort with questions rests on their assessment of the sponsor's legitimate right to know the information requested.