Census 2000 Evaluation B.3
April 3, 2003

Census Quality Survey to Evaluate Responses to the Census 2000 Question on Race: An Introduction to the Data

FINAL REPORT

This evaluation reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning. Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

Michael Bentley
Tracy Mattingly
Christine Hough
Claudette Bennett

Planning, Research, and
Evaluation Division
and Population Division

USCENSUSBUREAU

Helping You Make Informed Decisions

CONTENTS

LIST OF TABLES

EXECUTIVE SUMMARY

Data on race from most federal surveys currently reflect a collection methodology of asking respondents to mark only one category. Census 2000 was the first decennial census to ask respondents to "mark one or more races." Some users of the Census 2000 data on race 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 (OMB) has referred to this comparison as "bridging."

The objective of the Census Quality Survey (CQS) is to enable users to make comparisons between race data obtained using "mark one race" and "mark one or more races" methods. The CQS attempts to meet this objective by collecting race data using *both* methods from the *same* people. That is, each respondent in the sample was asked in separate interviews to report "one or more races" and to report a single race. Thus, the CQS can be used to evaluate how multiple race reporters respond when asked to report a single race. The data can be used to answer questions of the sort, "What fraction of people who report as 'White ***and*** Black' when asked to report one or more races, report as 'Black' when asked to report only one race?" The results can be applied to bridge the two methods by constructing statistical adjustments to race distributions obtained using one method to make them more comparable to race distributions obtained using the other.

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. This document, created to accompany the Census Quality Survey data file, provides the following information:

- the background on the reporting of Two or more races,
- the methods and limits of the Census Quality Survey,
- some results produced from the file, and
- a data dictionary (file layout) of the variables contained on the Census Quality Survey public-use data file.

Data file users are encouraged to read the Office of Management and Budget's 2000 "Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity" for more background on comparing racial data and bridging methods. This, and several related documents, is available at <http://www.whitehouse.gov/omb/inforeg/statpolicy.html#dr>.

The Census Quality Survey has 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 is split into two panels. Panel A received the "mark one or more races" instruction at the initial contact, whereas Panel B received the "mark one race" instruction first. During the second contact, or the re-contact, 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. In

addition, Census Quality Survey respondents were matched to their Census 2000 responses (with a match success rate of 86 percent) and these data can also be used to produce bridging parameters.

What were the housing unit response rates?

Initially about 27,500 housing unit addresses were designated to be in the sample for each panel. Of the eligible (occupied) addresses, 97 percent completed an interview in the initial contact. In the re-contact, sample housing units were only contacted if an initial contact questionnaire was completed. Of the eligible re-contact addresses, 87 percent completed an interview in Panel A and 94 percent completed an interview in Panel B.

Was the Census Quality Survey representative of Census 2000 data?

The results from the question on race suggest that each panel appears to be representative of Census 2000. Aggregated reporting of race among non-Hispanic respondents to the "mark one or more races" instruction closely resembles Census 2000 reporting of race for each panel. No race group appears to be significantly different from Census 2000 ($p < 0.1$ level) in either panel, including the Two or more races population. Reporting of race for Hispanic respondents is also similar to that in Census 2000, though in Panel A a smaller proportion of Hispanics chose White as a single race and a larger proportion chose Some other race compared with Census 2000 data.

Was reporting of race consistent between Census 2000 and the Census Quality Survey?

Only 40 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 ("mark one or more races" instruction). Similarly, only 41 percent of those in Panel B who reported two or more races in the census also reported Two or more races in the re-contact. The other 60 percent went on to report a single race. The generally low level of consistency in the reporting of Two or more races has several consequences, including:

- The effective sample size for computing bridging parameters is reduced and the parameters are sensitive to which data are used to compute them.
- The stability of bridging parameters may be unclear given the observed instability in reporting two or more races.

In contrast, 97 percent 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. For American Indian or Alaska Natives, Native Hawaiian or Other Pacific Islanders, and Some other race respondents, the reporting of race consistency ranges from 55 percent to 58 percent in Panel A, and 72 percent to 78 percent in Panel B.

How do people who reported Two or more races respond to a "mark one race" instruction?

We cross-tabulated the "mark one or more races" data collection contact with the "mark one race" contact to assess how individuals respond when asked to choose a single race for people for whom multiple races has been reported. Even with the "mark one race" instruction, a significant portion of respondents reported Two or more races. This portion was greatly reduced, though, when the followup race probe was used in the Panel A re-contact. Users of the data file will need to determine how best to treat these reluctant cases when computing bridging parameters. This treatment may depend on the particular purpose and uses of the file.

1. BACKGROUND

1.1 Evaluation objectives

Data on race from most federal surveys currently reflect a collection methodology of asking respondents to mark only one category. Census 2000 was the first decennial census to ask respondents to "mark one or more races." Some users of the Census 2000 data on race 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 (OMB) has referred to this comparison as "bridging."

The objective of the Census Quality Survey (CQS) is to enable users to make comparisons between race data obtained using "mark one race" and "mark one or more races" methods. The CQS attempts to meet this objective by collecting race data using *both* methods from the *same* people. That is, each respondent in the sample was asked in separate interviews to report "one or more races" and to report a single race. Thus, the CQS can be used to evaluate how multiple race reporters respond when asked to report a single race. The data can be used to answer questions of the sort, "What fraction of people who report as 'White ***and*** Black' when asked to report one or more races, report as 'Black' when asked to report only one race?" The results can be applied to bridge the two methods by constructing statistical adjustments to race distributions obtained using one method to make them more comparable to race distributions obtained using the other.

The primary goal is to improve comparisons between 1990 and Census 2000 race distributions at national and lower geographic levels. Other goals are to facilitate comparisons between data on race from Census 2000 and current Census Bureau surveys which instruct respondents to mark one race, and with data from the vital records system, which uses census data to calculate such indicators as birth and death rates.

Users of the data file are strongly encouraged to read OMB's 2000 "Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity[1]." This provides information on comparing racial data collected under the 1977 standards and the new standards which allow for the reporting of more than one race. The report also introduces various bridging tabulation methods, some of which do not require the use of auxiliary data such as that produced by the Census Quality Survey. This document, created to accompany the CQS public-use data file release, also includes information on the background and methods of the CQS, some results produced from the file, and a data dictionary (file layout) of the variables contained on the CQS public-use data file.

[1]The OMB report, and several related documents, is available on OMB's website at <http://www.whitehouse.gov/omb/inforeg/statpolicy.html#dr>.

1.2 The concept of race

Throughout U.S. history, the meanings, measurements, and categories defining racial groups have changed. Beginning with the first decennial census in 1790, the categories reflected both anthropological and biological connotations. Enumerators, based on their subjective view of the phenotypical features of respondents, determined the racial category to which individuals were assigned. Over the years, both the scientific and the popular understandings of race and the categories used to measure race have changed. As discussed in the 1997 Office of Management and Budget standards on collecting data on race and ethnicity, the current categories used to collect data on race are socio-political constructs and do not reflect any anthropological or genetic definitions. Additionally, the method by which these data are collected is based on self-identification.

Numerous studies have revealed that people may have several racial and ethnic identifications, and that these identifications may change over time and across circumstances, and thus inconsistent responses from interview to follow-up may occur. In 1974, Johnson found a lack of consistent reporting in ethnicity when 34 percent of people reporting had different ethnicity during a survey conducted in 1971 and 1972. Further, Hahn (1993) proposed that the more heterogeneous one perceives one's ancestry to be, the more one's self-identification is likely to change over time. He found that the proportion of people reporting different ancestries in initial and follow-up surveys increased in people reporting one (40 percent), two (58 percent), three (66 percent), and four (75 percent) ethnic backgrounds. Thus, it is difficult to conclude that ethnic identification is a fixed and singular characteristic for people in the United States. Instead, we recognize that self-identification of race and ethnicity is fluid for some people, and self-perception changes.

Because of the fluid nature of the concept of race, the reporting of race by some respondents may be influenced by minor deviations in question wording or mode of data collection. Any time lapse, modification to the questions, presence of an interviewer, or simple response variance can influence reporting of race patterns. Additionally, the introduction of the reporting of more than one race in Census 2000 may contribute to the fluidity in the reporting of race.

1.3 Reporting of race and ethnicity in the federal government

In response to legislative, programmatic, and administrative requirements in the federal government, the Office of Management and Budget (OMB) in 1977 issued the "Race and Ethnic Standards for Federal Statistics and Administrative Reporting," set forth in Statistical Policy Directive No. 15. These standards were used for more than two decades in decennial censuses, in national surveys of the population, and in data collections to meet statutory requirements. Data on race and ethnicity are needed to monitor equal access to housing, education, employment opportunities, and so forth for population groups that historically have experienced discrimination and differential treatment because of their race or ethnicity. The categories that were developed represent a social-political construct designed to be used in the collection of data on the race and ethnicity of major broad population groups in this country, and they are not

anthropologically or scientifically based. The four basic race categories specified in Directive No. 15 were: "*American Indian or Alaska Native*," "*Asian or Pacific Islander*," "*Black*," and "*White*." The two specified ethnic categories were: "*Hispanic origin*" and "*not of Hispanic origin*." The federal government treats race and Hispanic origin as two separate and distinct concepts. Hence, people of Hispanic origin may be of any race.

During the 1980's the standards in Directive No. 15 came under increasing criticism. Some individuals who reported data about themselves and various users of the data believed that the categories no longer adequately reflected the increasing racial and ethnic diversity of the population of the United States. As a result of these concerns, the OMB initiated a comprehensive review of Directive No. 15 and solicited comments from the public on the usefulness of the Directive.

After a lengthy review process, on October 30, 1997, the OMB issued revised standards that all federal agencies, beginning with Census 2000, were to use to collect, tabulate, and present data on race and ethnicity. Included in these standards is the identification of five racial categories: "*White*," "*Black or African American*," "*American Indian or Alaska Native*," "*Asian*," and "*Native Hawaiian or Other Pacific Islander*." For respondents unable to identify with any of these five race categories, the OMB approved a sixth category - "Some other race" (SOR) - on the Census 2000 questionnaire. The category "Some other race" was used in Census 2000 and is also used in a few other federal data collection activities. Respondents who provided write-in entries to the race question such as Moroccan, South African, Belizean, or a Hispanic origin (for example, Mexican, Puerto Rican, or Cuban) are included in the SOR category. The overwhelming majority (97 percent) of the 15.4 million people who reported SOR alone in Census 2000 were Hispanic. However, of all combinations of Two or more races that included SOR, only 59 percent were Hispanic. A large majority (90 percent) of the population identified as "SOR alone or in combination with one or more other races" was also Hispanic.

The 1997 standards also include changes in the terminology used for each group and the sequencing of the questions on race and Hispanic origin. In the 1990 census, the question on race preceded the question on Hispanic origin with two intervening questions. For Census 2000, the question on Hispanic origin was placed immediately before the question on race with a note to respondents to answer both questions. But perhaps the most profound change to the standards was asking respondents to mark one or more races. Many census data users, both governmental and non-governmental, need to understand how the Census 2000 race distributions compare with race distributions from other data sources where respondents are asked to report only one race.

1.4 Data on interracial families and reporting of race prior to Census 2000

Some of the impetus for the OMB to ask for the reporting of one or more races came from the increasing number of interracial marriages and "multiracial" births in the past three decades. Prior census data suggest that individuals from smaller racial population groups are more likely to form interracial unions with individuals outside their race than are individuals from the larger White and Black populations. Since the White population composes a large proportion of the

total population, most interracial marriages have one partner who is White; similarly, for most children with parents of different races, one parent is White.

The 1970 census identified about 321,000 interracial married couples. By 1980, the number had increased to about 1 million, and by 1990 to about 1.5 million interracial couples[2]. In 1990, all but 8 percent of these interracial couples included one spouse (or unmarried partner) who was White. In 14 percent of all interracial couples, the non-White spouse was Black; in 22 percent, American Indian or Alaska Native; in 31 percent, Asian or Pacific Islander; and in 25 percent, Other race (most of whom were of Hispanic origin).

Census data indicate that the number of children in interracial families grew from less than one-half million in 1970 to about two million in 1990. In 1990, for interracial families with one White partner, the other parent was Black for about 20 percent of all children, the other parent was Asian for 45 percent, and the other parent was American Indian and Alaska Native for about 34 percent.

Data from the National Center for Health Statistics on reporting of race on birth certificates indicate that the number of children of mixed racial parentage varies with the racial combinations of the parents involved. In 1968, for two percent of the births with at least one Black parent, the second parent was reported as White on the birth certificates (N=8,800). This percentage increased to 9 percent in 1994 (N=63,000). Analysis of the change in the number of births where one parent is Black and the other parent is of a different race is complicated by the increasing number of births for which the race of the second parent, usually the father, is not given on the birth certificate - 40 percent in 1994, compared with 24 percent in 1968.

Among births of American Indian and Alaska Native children, the percentages of births in which the second parent was listed as White was 28 percent (N=6,900) in 1968 and 45 percent (N=23,000) in 1994. Among births of Asian and Pacific Islander children, the percentages were 28 percent in 1968 and 26 percent in 1994.

1.5 Related studies on reporting of race for "multiracial" individuals

Since 1990, the Census Bureau has conducted three major studies to evaluate the feasibility of collecting data for people reporting more than one race. The first was the May 1995 Race and Ethnicity Supplement to the Current Population Survey, which was conducted jointly with the Bureau of Labor Statistics (BLS) (OMB, 1997a). This supplement was designed to test the effect of asking questions about race and Hispanic origin with and without a "multiracial" response option. The sample size for the May 1995 CPS Supplement was about 60,000 households. Data from this supplement indicated that, nationally, a little more than 1.5 percent of respondents identified as "multiracial." American Indians and Alaska Natives were more likely to report multiple races both using a separate "multiracial" response category and without

[2]The Census 2000 special report "Married-Couple and Unmarried-Partner Households: 2000" was released in February 2003 and is available online at <http://www.census.gov/prod/2003pubs/censr-5.pdf>.

a separate "multiracial" option. The proportions reporting in the White, Black, and Asian and Pacific Islander categories were not statistically different when the "multiracial" option was used.

The second study was the 1996 National Content Survey (NCS) (U.S. Census Bureau, 1996). The NCS, conducted from March through June 1996, was a mail survey of about 94,500 households representing about 95 percent of the country. Four of the 13 panels, each with about 6,000 households, were designed to evaluate the effect of adding a "multiracial" or "biracial" category and reversing the sequencing of the questions on race and Hispanic origin. The NCS contained a question on race that included a separate "multiracial" or "biracial" response category in two of the four panels. The proportion of respondents identifying themselves as "multiracial" was less than two percent. The addition of a "multiracial" category had no statistically significant effect on the percentage of respondents who reported as White, Black, American Indian, or Asian or Pacific Islander. However, the relatively small sample size in the NCS might not detect changes that were substantively important for small populations.

The third study was the 1996 Race and Ethnic Targeted Test (RAETT) (U.S. Census Bureau, 1997). The RAETT, conducted in the summer of 1996, targeted 112,000 households in areas with high concentrations of six specified racial or ethnic groups: White, Black, American Indian, Alaska Native, Asian or Pacific Islander, and Hispanic. Respondents in this multipanel experiment were able to report their "multiracial" identity in several ways, depending on which questionnaire they received. First, they could mark a response box labeled "multiracial" and choose to write-in specific races. Second, they could mark two or more boxes in the question on race, in response to an instruction to "mark one or more" or "mark all that apply." Additionally, some respondents provided multiple responses even when asked to "mark one race." No differences were detected in the percentages reporting solely as White, Black, or American Indian when given the option to report more than one race. However, the percentages of the population reporting as Alaska Native and Asian and Pacific Islanders were statistically different when given this option. Significant differences were also found in the single race categories of Asian and Pacific Islander and of American Indian and Alaska Native when the "multiracial" category was used. The percentage for whom race was reported as Asian and Pacific Islander as a single race was also statistically different when the "mark all that apply" option was used, but not when the "mark one or more" option was given.

In addition to these studies, data for respondents providing one or more responses to a question on race have been collected since 1976 by the National Center for Health Statistics (NCHS) through the National Health Interview Survey (NHIS). The NHIS is a multipurpose health survey, where information is obtained on a wide range of health and health care topics through computer-assisted personal interviews. The current sample size is about 40,000 housing units or about 100,000 people with over sampling for Blacks and Hispanics (Madans, 2000). About 1.6 percent of all responses in the 1993-1995 NHIS included multiple race responses. This proportion has remained fairly consistent since 1982.

The NHIS also includes a followup question for people who report more than one race to select a single race that "best" describes their race. Results from the 1997 NHIS suggest that there are reporting differences among the racial combinations. For example, about 82 percent of respondents who report as American Indian and White, select White when asked to select a single race. Among those who report Asian and White, about 49 percent report as Asian with about 13 percent reporting as multiple race. Among respondents who report as Black and White, about 50 percent report as Black, and the remaining 50 percent is about evenly split between White and multiple race.

The NHIS also found that the reporting of race is influenced by the age of the individual and the region of residence. Respondents who report a single race tend to be older than those who report multiple races. According to results from the 1997 NHIS, about 78 percent of people for whom race was reported as Black and White were less than 18 years of age, compared with about 34 percent of people who reported Black only (Lucas, 2000). A similar pattern was observed for Asian and Pacific Islanders. In terms of regional variation, about 55 percent of those who reported a single race of Black resided in the South in contrast to 25 percent of respondents who reported as Black and White. In general, the majority of the reporting of multiple races occurred in the West.

Other researchers (Goldstein and Morning 2000) have also suggested that family composition and income help explain reporting among people who report more than one race. Harris (2000) presents a comprehensive examination of people who reported two or more races in the National Longitudinal Study of Adolescent Health (ADD Health) survey. Harris used these data to examine the levels, as well as the socio-demographic characteristics, of racial identity for "multiracial youth" compared with their monoracial peers. Harris used three measures of racial identification: school race, home race, and parents' race. School race was defined as the race reported by students when administered the survey instrument at school. Home race is the race reported by the student while at home, and parents' race is the racial identity provided by the biological parents at home. Harris found that multiracial or multiple response individuals tended to be younger, more likely to be female, less likely to live with both biological parents, and more likely to live in racially diverse and urban areas in the West region of the country.

1.6 Background on methodological issues associated with reporting of race

The objective of this study is to provide users with a data file which permits cross-classification of responses using two measurement methods: the Census 2000 question on race that allows reporting more than one race, and a "mark one" question on race similar to that used in the 1990 census. Ideally, the two measurements would be closely comparable, both would be collected under identical conditions using comparable methods except for the difference in the "mark one" or "mark one or more" question instructions. However, due to practical constraints, this study could not control all methodological factors which may influence reporting of race. Differences in the methods or conditions under which the measurements were obtained may introduce systematic bias or variability in race reports. This section discusses certain methodological issues which influenced the design of the survey and may influence the data. These issues

should be taken into account by analysts using and interpreting the data, especially when comparing them with other sources of race data. Note that a detailed discussion of CQS-specific issues is provided in section 2 (Methods) and section 3 (Limits).

1.6.1 Conceptual confusion and "other race" reporting

One problem concerns the mismatch between survey and respondent categories for race. Although the racial classification system used by the statistical system works well for many respondents, there is evidence that many others, including many Hispanics, have difficulty understanding the categories and selecting an appropriate response to accurately reflect their identity (Gerber, de la Puente, and Levin, 1998). Some groups fail to find a category that expresses their own sense of race, the most important being a substantial fraction of Hispanics. In Census 2000, 6.6 percent of people reported as "Some other race," either alone or in combination with another race category (U. S. Census Bureau, 2001). Most were Hispanics, who did not find a Hispanic category listed and so marked "Some other race" and/or wrote in "Hispanic," "Latino," or a specific Hispanic group (e.g., Mexican, Salvadoran). In Census 2000 42.2 percent of Hispanics were identified as Some other race alone, compared to just 0.2 percent of non-Hispanics (see Attachment 4 for the full Census 2000 distribution of race). Many Americans see both Hispanicity and race as part of the same global concept. This view seems rooted in culture and is given authority by other race questions or representations of data which include "Hispanic" among the races.

In part due to conceptual confusion, "other race" reporting by Hispanics is vulnerable to the effects of methodological differences between surveys, including the effects of question order and interviewer behavior.

1.6.2 Order effects

When race is asked first in a self-administered questionnaire, many Hispanic respondents look for but do not find a category to describe themselves, and so report "other race" and/or write in a Hispanic group (Kissam, Herrera, and Nakamoto, 1993; Davis et al. 2001). The effect of question sequence on Hispanic reporting of race in self-administered mail questionnaires is well-documented (Bates et al., 1995; U. S. Census Bureau, 1996; 1997). For this reason (and because item nonresponse rates are reduced in the Hispanic-first sequence), OMB guidelines require that the Hispanic origin question precede the question on race, when they are asked as separate questions. An experiment conducted during Census 2000 showed that the fraction of Hispanics reporting as "White" is higher by about 10 percentage points, and the fraction reporting as "Some other race" is lower by about the same amount, when the Hispanic origin item is asked first compared to the reverse order (Martin, de la Puente, and Bennett, 2001).

For all of the panels in this study, Hispanic origin was asked before race. However, some data sources with which an analyst might want to compare these data use the reverse order, including the 1990 census.

1.6.3 Potential mode effects on reporting of race

Different modes of survey administration present particular problems for the race question which may affect the comparability of data. Mode differences in reporting of race may arise from the effects of mode on communication of the race categories and the "one or more" option, or from possible interviewer effects, among other factors.

Because the list is long (15 categories), and because the categories may not correspond to some respondents' own understandings of race, it is usually thought necessary to communicate all the categories respondents might choose. In a self-administered questionnaire, it is relatively easy to present a long list of categories below the question. In a survey administered in person by an interviewer, communicating the list of categories is more awkward: an interviewer may read them, or present a flashcard, or both. Sometimes interviewers may abbreviate the list or respondents may interrupt them before they have read all of the categories. If a response does not fit one of the categories, the interviewer must further classify the response into a category on the spot, or offer a choice of categories to ask the respondent to choose one or more which fit. Problems can arise when respondents do not find the category with which they identify (e.g., some Hispanics or Arabs), or reject categorization of their group within a predefined category (e.g., Haitians who do not consider themselves to be "Black, African American, or Negro"), or do not know in which category they belong (e.g., Central American Indians). Telephone interviews present an even greater challenge because the flashcard is not available to communicate the categories; usually the question is modified to use a branching structure, which may result in differences in response.

Another mode-related problem is that in personal interviews the answer to the question about race may seem self-evident to both interviewer and respondent, making the question awkward to ask in person. In censuses before 1980 enumerators recorded race based on their own observation (U.S. Census Bureau 1983), and this practice may still sometimes occur. In the 1980 census, Hispanics were far less likely to be reported as "other race" in personal visit reinterviews than in self-administered census questionnaires (McKenney et al., 1985). The explanation is that enumerators recorded as "White" many Hispanics who reported themselves as "Some other race" on their census questionnaires. Also, interviewers whose training emphasizes probing "other" responses often obtain lower rates of "Some other race" reporting than interviewers who were not trained to probe "other" responses (Raglin and Leslie, 2002).

Thus, the mode of interview affects the manner and completeness with which the race categories are communicated, and may well affect responses. In addition, there appears to be a good deal of interviewer variability in how the question is asked. One study of personal interviews found that interviewers made major changes to the race question in over 40 percent of the interviews that were behavior coded, usually by omitting categories (Smiley and Keeley, 1996). (See section 2.1.3.)

The related problem of communicating the intent of the question to elicit reports of one or more races may also be affected by mode. Census 2000 was the first census to ask for reports of one

or more races, and early cognitive testing (Gerber, de la Puente, and Levin, 1998) showed that respondents often did not realize they had the option of marking more than one box, even when they had just read the pertinent instruction aloud. People may not notice this novel feature because the Census 2000 mail form asked, "What is this person's race?" Unless the respondent read the instruction where the phrase "one or more" occurs the option is lost. In the CQS, the "one or more" option was communicated by the use of a flashcard in personal interviews and by a slight rewording of the question in telephone interviews. See sections 2.1, 2.6, and 3.2 for discussion of mode and related limits.

1.6.4 Resistance or reluctance to report one race

The design of this study required that we ask respondents to answer both a question asking for reports of one or more races and a question that requested one race. Some mail respondents report more than one race even when asked to report one (almost 1 percent did so in an experiment conducted during Census 2000, nearly half the number who reported more than one race when invited to do so; Martin, 2002). Evidence suggests that some respondents who identify with more than one race may resist providing only one race, resulting in missing data. The National Health Interview Survey has for some years accepted reports of more than one race, following up with a probe to obtain single race reports. Some groups have high nonresponse rates to the follow-up question asking for a single race (23 percent for White/Black and 13 percent of White/Asian and Pacific Islander respondents; Lucas et al. 1999). The CQS uses a split-panel experimental design to obtain single race reports by asking the "mark one race" question in a mail questionnaire or in a telephone reinterview, where interviewers could probe respondents who were reluctant to report a single race. (See sections 2.1 and 4.2.)

1.6.5 Proxy vs. Self Reporting of race

OMB has established the general principle that "self-identification is the preferred means of obtaining information about an individual's race and ethnicity, except in instances where observer identification is more practical (e.g., completing a death certificate)" (OMB, 1997b: 58785). As a practical matter, most household surveys and the census ask a household respondent to report the race (or races) each household member "considers himself/herself to be." The decennial census in last resort cases obtains information from non-household proxies (e.g., a landlord or a neighbor) whose knowledge of a person's racial self-identification may be limited. A household respondent's reports may be influenced by his or her own race, relationship to other household members, perception of the purpose of the data collection, and other factors. Within a household, different respondents may provide different responses to the question on race. This may be especially true for people who consider themselves as multiracial. In the CQS, every attempt was made to collect responses to both questions on race from the same household respondent, in order to reduce variability in reporting of race due to changes in respondent. In addition, unlike the census, non-household proxy reports were not accepted in the CQS (see sections 2.1.2 and 3.4 for further information).

2. METHODS

The CQS was designed to produce a data file that could be used as a bridge between "single" and "one or more races" distributions. The CQS has a nationally representative sample 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 is split into two panels. Panel A received the "mark one or more races" instruction at the CQS initial contact, whereas Panel B received the "mark one race" instruction. During the second contact, or the re-contact, each panel received the alternate instruction. See Table 1 for an overview of the data collection sequence[3]. 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.

Table 1. CQS data collection sequence: race instruction by panel.

	Data Collection Contact		
CQS Panel	**Census 2000 (April - August 2000)**	**CQS Initial Contact (June - August 2001)**	**CQS Re-contact (August - October 2001)**
A	"Mark one or more races"	"Mark one or more races"	"Choose one race"
B	"Mark one or more races"	"Mark one race"	"Choose one or more races"

The following sections provide detailed information on the data collection process (section 2.1), the CQS sample design (section 2.2), and linking the CQS data to Census 2000 data (section 2.3). The latter sections describe the weights assigned to the sample cases and a method for estimating variances for CQS estimates (section 2.4), the development of a tract-level contextual variable (section 2.5), methodological considerations for comparisons using the Census Quality Survey (section 2.6), and the application of quality assurance procedures (section 2.7).

2.1 Data collection

The methodology for the evaluation required that the sample households be contacted twice during the CQS survey to provide information on race. The evaluation required the administration of both a 1990 Census instruction to the question on race, that is, "mark one race," and the Census 2000 instruction to the question on race in a split panel design. A total sample of 55,000 addresses were selected, including households containing respondents who reported more than one race and households where all respondents reported only a single race in Census 2000.

2.1.1 Initial contact data collection

[3]Refer to Attachment 2 for the exact wording of Census 2000 and Census Quality Survey questions on race and Hispanic origin.

The sample households were mailed an initial questionnaire, which they received in June 2001. A second questionnaire was sent in early July 2001 to those households that did not return the first questionnaire. Nonresponse followup (NRFU) procedures similar to those used for Census 2000 were implemented for households that did not respond to the first and second mailings. Enumerators visited addresses that did not respond via mail and areas which were not in mailout/mailback enumeration areas. For the initial data collection, one panel (Panel A) of about 27,500 housing units (HUs) was enumerated using a questionnaire with the race and Hispanic origin questions identical to Census 2000 (with the wording "mark one or more races" for the question on race). The other panel (Panel B) of about 27,500 housing units was enumerated using the identical questionnaire, except the instruction to the question on race was to "mark one race." For personal visit interviews in the CQS initial contact, as in Census 2000, the enumerators used show cards to help communicate to the respondent the instructions and the categories for the questions on race and ethnicity. In the initial contact, about 54 percent responded by mail and the remainder were interviewed in personal visits in both panels.

The CQS survey rules called for enumerating people who lived or were staying at sample units on the date of the interview. These may or may not have been the sample people who lived at the sample address on April 1, 2000 (Census Day). Every effort was made to capture data for people who moved into the sample address (inmovers) and ascertain the previous address at which they were enumerated in Census 2000. However, no efforts were made to trace outmovers. That is, we did not ask information about people who had moved out of the sample addresses since Census 2000. In addition to the race and other short form questions, respondents were asked whether a census form had been filled out for the household and, if so, who completed the form. This information could be used to assess consistency of reporting when race was reported by the same or a different respondent.

In order to assist the matching process (see section 2.3), we also collected information on the address where each person in the household was living on Census Day. The relationship question was not asked in the initial contact questionnaires due to space limitations, since the census address item was included, but was asked in the re-contact interview. Note that, unlike Census 2000, the CQS did not permit non-household proxy respondents. That is, the form was intended to be "filled out by a person who lives at this address and is knowledgeable about the people living here now."

2.1.2 Re-contact data collection

Four to six weeks after the second mailout, the sample households that responded in the initial contact were then re-contacted by telephone to collect data on race from the alternate race question and other information on socio-demographic characteristics such as education and income. A reverse questionnaire design procedure was used to re-contact housing units that participated in the initial data collection. That is, housing units that participated in the initial data collection with the mark one or more races instruction (Panel A) were re-contacted by telephone and asked to report one race. Those housing units that first received the "mark one race" instruction in the initial contact (Panel B) were asked to "choose one or more races" in the

re-contact. For housing units for which there was no telephone contact, personal interviews were conducted to collect the re-contact information. More than 70 percent of the re-contact interviews were conducted by telephone.

The questions on both of the re-contact questionnaires were similar; only minor modifications were made in Panel A to probe for additional information in instances where respondents were reluctant to report a single race when asked to do so. During the re-contact, every effort was made to speak with the individual who completed the initial questionnaire. To facilitate this effort, address and name information from the initial questionnaires were transcribed to the re-contact questionnaires.

During the re-contact interview, respondents were asked to provide additional socio-demographic information such as veteran's status, educational attainment, household income, and language spoken at home. This information was thought to be relevant to the issue of differential race bridging parameters. In addition, each person's relationship to the householder was also included in the re-contact.

2.1.3 Cognitive interviews

About 150 cognitive interviews were conducted prior to data collection to provide insight into potential reporting problems associated with using two separate instruments and using a telephone re-contact to obtain additional information about reporting of race (Davis et al., 2001). Information collected from the cognitive interviews was used to design and develop the final questionnaires (Davis et al., 2002).

2.1.4 Editing and imputation of collected data

The procedures for editing CQS data on race and Hispanic origin were very similar to the procedures used for editing Census 2000 data on race and Hispanic origin. One major difference is that editing procedures for race and Hispanic origin in the CQS, unlike those used in Census 2000, did not impute for nonresponse except when Hispanic origin could be classified from responses to the question on race or race could be obtained from responses to the question on Hispanic origin.

2.2 Sample design

In this section we present the basic sample design decisions relating to the overall sample size, allocation of the sample to form type (Census 2000 short or long form), allocation of the sample to panel, and the geographic level for which the sample was designed to provide relatively

reliable estimates. We discuss the rationale for these choices and the methods used to operationalize these design decisions. Below is a brief summary of the sample design, as outlined in the following sections:

- The final sample size was approximately 50,000 interviewed HUs (see section 2.2.1).
- 25 percent of the sample was allocated to each of the four cells created by crossing form type (short, long) and panel (A or B) (section 2.2.2).
- Each state was treated as an independent sampling stratum (section 2.2.3).
- HUs with at least one individual who reported more than one race in Census 2000 were oversampled (90 percent of the initial sample) (section 2.2.4).
- Four distinct sampling strata were identified within each state (section 2.2.5).

2.2.1 Designated sample size

The initial designated sample size for this evaluation was 55,000 HUs. It was anticipated that approximately 10 percent of the selected housing units would be vacant when the CQS data collection phase was completed. Thus, the sample size eligible for interviews was expected to be about 50,000 housing units. The sampling frame consisted of all units in the Hundred Percent Detail File (HDF), which is a tabulation geography file containing the 100 percent detail data and is the source for creating the Public Law 94-171 (redistricting) counts. Note that the group quarters population was not included in the CQS universe.

2.2.2 Sample allocation to panel and form type

Panel refers to the two distinct methodologies used to collect the data (see section 2.1). As there was no *a priori* information to favor one data collection methodology over the other from a precision or validity criterion, the optimum allocation method was to assign half of the sample (27,500 HUs) to each panel.

In discussions concerning the design, methodology, and analysis of the CQS, it was suggested that selection of some proportion of the sample from among HUs designated to receive a long form questionnaire in Census 2000 could provide valuable background information for internal analysis purposes[4]. Again, based on no *a priori* information that the evaluation would be compromised in any major way (e.g., substantially reduced mail response) by selecting long form HUs, half of the sample was allocated to each form type in order to increase the availability of such background information. The CQS initial contact questionnaire only included short form data items, though, and the re-contact questionnaire had a few long form questions.

2.2.3 Sample allocation to census division and state

[4]Due to disclosure limitations, these long form data cannot be included on the public-use micro-data file.

The sample was first allocated to the nine census divisions[5] proportional to the square root of the Census 2000 count of individuals reporting more than one race. In an effort to provide for the possibility of producing state level estimates for those states containing relatively few Two or more races individuals, we also allocated the division level sample to states within each division proportional to the square root of the state's Two or more races population. Further, the division sample sizes were rounded to the nearest 100 and the state sample sizes were rounded to the nearest 50 with the restriction that the minimum state sample size was 300 HUs. In order to allow for an expected 10 percent vacancy rate during data collection, the designated sample sizes were increased by 10 percent. Thus, the sample design treats each state as an independent sampling stratum in anticipation that the sample allocated to each was sufficient to produce relatively reliable state level estimates. Of course, the data can always be cumulated to higher levels of geography (i.e., census division) if increased reliability is necessary.

2.2.4 Allocation to housing units containing Two or more races individuals vs. all others

In order to maximize the likelihood of contacting HUs in CQS that contained individuals reporting more than one race, the majority of the sample HUs were selected from among HUs containing at least one individual who reported multiple races in Census 2000. In order to ensure that every HU (occupied or vacant) enumerated in Census 2000 had a chance to be included in the sample, HUs that were vacant and HUs that included only people who reported a single race in Census 2000 were allocated a disproportionately lower fraction of the sample. In addition, the HUs containing all single race people in Census 2000 may have contained Two or more races individuals in the CQS if the household enumerated in Census 2000 moved and was replaced by a household containing Two or more races individuals. For the CQS design, 90 percent of the sample was allocated to the Census 2000 universe of HUs containing Two or more races individuals in order to maximize the yield of such people from the CQS sample.

Allocation of the sample to the 57 possible combinations of race from Census 2000 was accomplished as follows[6]. There were five combinations of "Some other race" (SOR) and one of the five OMB race categories (e.g., "White ***and*** SOR"), ten combinations consisting of two OMB race categories (e.g., "White ***and*** Black"), and 42 categories consisting of three or more race categories. Proportional allocation of the sample to the 57 combinations would result in more than 42 percent of the sample being allocated to the SOR combinations, including more than 30 percent to the combination "White ***and*** SOR." Most SOR responses are Hispanic ethnicities, so

---

[5]A division is a grouping of states within a census geographic region, established by the Census Bureau for the presentation of census data. The current nine divisions are intended to represent relatively homogenous areas that are subdivisions of the four census geographic regions. For a description of each of the divisions, see Attachment 1 for the public-use data file layout.

[6]The six race categories of Census 2000 (White, Black, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Some other race) can be put together in 57 possible combinations of two, three, four, five, or six races. Refer to American Factfinder Census 2000 Summary File 1 Table P3 (Race) for a complete population count of all single race and Two or more races combinations.

we decided to focus on the OMB race combinations[7]. Thus, in an effort to provide greater reliability for the combinations consisting of two OMB categories excluding SOR, it was decided to sample the SOR combinations at one-third the rate of the other combinations. This resulted in 18 percent of the sample being distributed to the SOR combinations, in contrast to 42 percent using proportional allocation[8].

2.2.5 Four sampling strata identified

In order to allocate state level sample sizes consistent with the design decisions outlined earlier in section 2.1, it was necessary to define four sampling strata (SS) as follows:

- **SS=1:** Consists of HUs containing ALL single race people OR Two or more races people where the Two or more races code is based on the Census 2000 edit and imputation process. This stratum defines 88.1 percent of the total Census 2000 universe.

- **SS=2:** Consists of HUs containing one or more individuals having a race code consisting of SOR plus another OMB category (e.g., "White ***and*** SOR"). This stratum defines 1.1 percent of the total Census 2000 universe.

- **SS=3:** Consists of HUs containing one or more individuals having a race code consisting of two or more OMB categories including the SOR category if the response consists of three or more races. This stratum defines 1.8 percent of the total Census 2000 universe.

- **SS=4:** Consists of HUs enumerated as vacant in Census 2000. This stratum defines 9.0 percent of the total Census 2000 universe.

Ten percent of the overall sample size was allocated to SS=1 and SS=4. The remaining 90 percent of the sample was allocated to SS=2 and SS=3, and sample HUs in SS=2 were selected at one-third the rate of those in SS=3.

2.3 How sample cases were matched to Census 2000 data

[7]Census Bureau research shows that the reporting of "Some other race" (SOR) is highly related to how Hispanics report race. Many responses to race are "ethnic" terms. Therefore, the Census Bureau developed a method which was called the "90 Percent Rule" to reclassify ethnic responses in the race question into an OMB race category. The method is empirically based using 1990 Census sample data as reported and not imputed. Single ancestry responses (which are primarily ethnic responses) were cross-tabulated by race responses. If 90 percent or more of respondents of a specific ancestry group selected a particular race, then that race was assigned to respondents who reported that particular ethnic response in the race question in Census 2000. If less than 90 percent of respondents in 1990 selected any particular race category, then SOR was assigned. (For more information on the 90 percent rule, see Census 2000 Decision Memorandum No. 106, July 6, 2000).

[8]Note that on the public-use data file a recoded race variable was created in which Some other race was removed from all Two or more races responses (see Attachment 1).

In order to make comparisons between Census 2000 and CQS race data, we linked census records to CQS records. This controls for inmovers and outmovers since we sampled HUs from Census 2000 and not individuals. This procedure also provided each panel with another "mark one or more races" contact that could be used to produce bridging parameters. The linkage process matched a record in the Hundred Percent Census Unedited File (HCUF) to records in the CQS file by comparing various fields such as: first name, last name, middle initial, suffix, sex, date of birth, age, street name, and zip code.

2.3.1 Record linkage software

The Vality Integrity™ (website: http://www.vality.com/news/vality/) record linkage software was used to match the CQS and Census 2000 data. The record linkage software generates an agreement weight and a disagreement weight for each match field from: (1) the probability the fields agree given that the record pair *is* a match, and (2) the probability the fields agree given that the record pair *is not* a match. A composite weight is generated for each record by adding the agreement and disagreement weights from the comparison for each match variable. Weight cutoffs, to indicate whether a composite weight score was a match or not, were set based on review of record pair listings generated during the development stages.

2.3.2 Four phase matching process

The first phase of the matching process involved matching the CQS data from the initial contact to the HCUF reference file. The reference file contains all HCUF person-level records for each selected housing unit from the sampling results. Six passes were processed as part of the reference phase, with different match fields and parameter settings for each pass. The second phase matched the CQS residuals (people who did not match to the reference file) to HCUF state files. Prior to the third matching phase, name and address corrections from the CQS re-contact were used to update the data from the final linked initial contact file. All records were then matched again to the HCUF files, using the same process, match parameters, and cutoffs as in the initial contact match. Finally, phase 4 involved the linking of the updated CQS residuals to HCUF state files.

2.3.3 Matching results

Eighty-six percent of the total CQS person records were matched to their respective Census 2000 data. That is, out of 155,137 records on the data file, 133,086 records have Census 2000 race data. Although non-matched individual records are included on the data file, we removed non-matched cases from all analyses that are provided in this evaluation report (see sections 4.2, 4.3, and 4.4). The race and Hispanic origin distributions for matched and non-matched people was found to be similar, though.

2.4 Variance estimation and weighting adjustments

2.4.1 Variance estimation

Given that the CQS design is not a simple random sample, but rather a stratified clustered design, CQS variances should be calculated using methods developed for complex survey designs. Using the unadjusted simple random sample variance will underestimate the CQS estimates' variances and can result in making the determination that differences are significant, when in fact they are not. The standard errors and variances that appear in this report were calculated using the stratified jackknife replication approach using the VPLX software available on the Census Bureau website (Fay, 1998).

2.4.2 Weighting adjustments

Each person record in the CQS data contains two weights to use in creating estimates from the CQS data. The first weight, Z_WGT1, is the inverse of the probability of selection, and has a nonzero value on all records. The second weight, Z_WGT2, has the same value as Z_WGT1 for all cases coming from sampling strata 2 and 3. Cases selected in strata 1 and 4 have a Z_WGT2=0. The second weight was created to be used specifically in forming estimates of the population reporting two or more races in one or more of the contacts and their single race response. Due to the sample design of the CQS, strata 1 and 4 cases were assigned very large weights compared to strata 2 and 3, since we oversampled in strata 2 and 3 and highly undersampled in strata 1 and 4. Table 2 shows the distribution of the sample across the four sampling strata and the extreme variability in the weights by stratum.

Table 2. Distribution of sample people across sampling strata.

	Stratum 1	Stratum 2	Stratum 3	Stratum 4
Unweighted Number of Sample People	Panel A=6,032 Panel B=5,970	Panel A=11,134 Panel B=11,004	Panel A=60,109 Panel B=60,185	Panel A=352 Panel B=351
Portion of Sample	8%	14%	77%	1%
Average Weight (Z_WGT1)	20,925	173	51	17,910
Minimum Weight	654	18	5	654
Maximum Weight	69,004	577	198	69,004

Ordinarily, one would not expect to have individuals from stratum 1 reporting more than one race, since by definition stratum 1 is made up of housing units where no members reported more than one race in Census 2000. But, due to the effect of inmovers and the instability of reporting more than one race, we do have a few sample cases from stratum 1 who reported more than one race when given the option (see limits section 3.5). When these cases were included in the analysis of people reporting more than one race, they had a sizable effect on the weighted total

and the variances. After much deliberation and outside consultation, we decided to create a second weight, which gave a weight of zero to all respondents in strata 1 and 4. This had a dramatic effect on reducing the variances, increased the reliability of the estimates, and eliminated the effects of several outlying weights. **Z_WGT2 should be used to evaluate how Two or more races individuals respond when asked to report a single race. Z_WGT1 should be used for other statistical analyses, including estimating the race and Hispanic origin distributions.**

2.5 Tract-level contextual variable

In order to enhance the subnational analyses, we added a single contextual variable to the data file, using Census 2000 tract level data. The file contains the variable Z_NHWTRC, which indicates whether an address is from a tract where the concentration of the non-Hispanic White population is either above or below the state median. This will be useful in producing separate bridging parameters for the areas with a low versus a high concentration of non-Hispanic Whites.

Although there were a large number of contextual variables that we considered adding to the file, we were confined to just one variable in order to limit the risk of disclosure. Some of the other considerations for tract density variables were: other race variables, Hispanic origin, educational attainment, poverty, tenure, unemployment, military experience, and age. After reviewing the associations of all the possible variables with responses to a "mark one race" instruction, we concluded that Z_NHWTRC was the most useful and most closely associated with the major race combinations.

2.6 Methodological considerations for comparisons using Census Quality Survey

The different approaches to the collection of data on race using a "mark [choose] one race" and "mark [choose] one or more races" instruction are subject to different methodological limitations. The CQS design supports six comparisons using the data collection contacts shown in Table 1 that users of the public-use data file may choose from to allocate responses of more than one race to a single race. Users may compare the results with data sources which asked for only reports of one race or conduct other methodological or substantive analyses on the reporting of race.

No one large scale sample survey incorporates all of the ideal data collection methods for meeting the objective of the CQS. However, as many components as possible of the ideal criteria of a large scale sample survey are met given the timing, cost and complexity of the data collection involved. To understand the characteristics of the ideal method, six criteria were identified to be used to assess the merits of possible methodologies for the CQS. The criteria and proposed design options are given in Attachment 3. After internal Census Bureau discussions and consultation outside experts, the methods used in Panels A and B were judged to be those that most closely approximate the ideal criteria.

The comparison methods, along with a discussion of the more important survey methodological issues associated with each comparison, follow. **The Census Bureau has not evaluated the**

quality of the data in these various comparisons and hence cannot recommend one comparison over another as a source of bridge data. In Table 3 we describe the sample design elements and methodological limitations for each of the six comparisons that the CQS data on race will support. The user should consider carefully (and analyze, if possible) the potential methodological factors which may affect the use of the data as a source of bridge parameters. Please refer to Appendix C, "The Bridge Report: Tabulation Options for Trend Analysis," of OMB's 2000 "Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity" for a further description of the criteria by which different "bridging" methods should be evaluated.

Table 3. Methodological limitations for each of the six comparisons using CQS data.*

Bridging Comparison	Similar Modes of Data Collection	Relatively Shorter Time Between Contacts	Design Incorporates "Same Respondent"	Relatively Larger Sample Size Available for Analysis
Panel A (1) Initial vs. Re-contact w/ probe (2) Initial vs. Re-contact w/o probe		✓	✓	
Panel A (3) Census vs. Re-contact w/ probe (4) Census vs. Re-contact w/o probe				✓
Panel B (5) Census vs. Initial Contact	✓			✓
Panel B (6) Re-contact vs. Initial Contact		✓	✓	

* If a check mark (✓) exists in a given cell in Table 2, then that methodological consideration is a favorable aspect for that particular bridging comparison.

One factor that should be considered in the use of the comparisons is that the Panel A re-contact attempted to obtain a single race response using an instruction to the respondent to "choose one race." A subsequent followup probe question was asked if the respondent nevertheless chose more than one race for a given person in the household. This followup question was asked to obtain a single race response. Even so, a significant portion of respondents still reported Two or more races. Users of the public-use data file may produce estimates with or without the followup probe data using the Panel A re-contact, depending on their particular purpose.

Another methodological consideration for bridging data on race is whether or not the two data contacts had similar modes of data collection. The CQS initial contact and re-contact data collection modes differed. The initial contact used mailout/mailback and personal visit methods, while the re-contact questionnaire was administered by telephone or personal visit for households with no telephone information. The specific effects of mode differences on the estimates are unknown. The only comparison with similar modes for the two data contacts is that comparing Census 2000 data to the CQS initial contact for Panel B, since the initial contact replicated most of the Census 2000 data collection methods and used the Census 2000 questionnaire with only

minor modifications. (See section 2.1 for more detail on the data collection, and section 3.2 for further discussion on the limits of the different collection methods.)

A third methodological consideration is whether or not there was a comparable time frame between the two data collections. The initial contact and re-contact data were collected relatively close in time. This probably reduced such factors as attrition due to households moving, high response variance, and the likelihood that real changes occurred in how people respond to the race question. On the other hand, the CQS initial contact occurred 15 to 18 months after Census 2000, introducing unknown effects in how respondents report race. Note that, in addition to the time lag, the comparison using Census 2000 and the CQS re-contact for Panel A has the CQS initial contact between the two comparative data collection points.

Another methodological consideration for bridging is whether or not the sample design enabled contacting the same household respondent. The initial contact and re-contact questionnaires were designed to collect data, to the extent possible, from the same household respondent. Address and name information from the initial contact questionnaires were transcribed to the re-contact questionnaires. Comparisons which include Census 2000 data on race have an increased likelihood of having different household respondents in the two data collections. Note that when data is collected from the same respondent in a relatively short time frame, the latter responses may be somewhat dependent on earlier responses due to recall and conditioning effects (see section 3.3).

A final consideration that may affect the choice of the comparison to be used is the number of sample individuals available in sampling strata 2 and 3 to estimate the "bridging parameters." (Refer to section 2.2 for information on the sample design and section 2.4.2 for information on weighting adjustments.) A considerably larger sample is available from the population reporting Two or more races for the two comparisons that use Census 2000 data. Thus, users will want to weigh the increased reliability of estimates that will occur using these comparisons against the potential biasing effects of the survey methods and issues discussed previously. In addition, single race information is missing for about one-quarter of individuals for whom Two or more races were previously reported, either because of nonresponse or respondents gave another multiple race response (see Tables 13 and 14).

2.7 Applying quality assurance procedures

We applied quality assurance procedures throughout the creation of this report. They encompassed how we determined evaluation methods, created specifications for project procedures and software, designed and reviewed computer systems, developed clerical and computer procedures, analyzed data, and prepared this report.

3. LIMITS

This section outlines the operational and qualitative limits of the evaluation.

3.1 Census Quality Survey could not repeat the Census 2000 environment

Many factors associated with the decennial census can affect responses to the question on race. Ideally, we want to collect CQS data in that same environment. However, several major elements make the census unique and nearly impossible to replicate. One factor is the Census Bureau's extensive Partnership and Marketing Program to promote Census 2000. This program included numerous census partnerships, nationwide paid advertising, special methods to encourage response from direct mail, as well as a media public relation's campaign and many other promotions and special events.

3.2 Different collection methods were used in the initial contact and the re-contact

The initial CQS data collection attempted to replicate, as much as possible, the usual census enumeration techniques. This included mailout/mailback of survey forms and personal enumeration for addresses where no questionnaire was returned by mail and for areas of the country where the mailout/mailback method was not appropriate. The re-contact, however, used telephone and personal visit interviewing techniques. Switching from mail to telephone interviewing may cause potential mode differences in the responses to the race question.

3.3 Conditioning effects between initial contact and re-contact

Response to the question on race using either instruction can be influenced by the response to the first race question when both questions are asked of the same respondent in a relatively short time frame. That is, the order in which the questions are asked can have an effect. Ideally, we do not want the measurement of one question on race (e.g., "mark one race") to be modified by the earlier presence of the other question on race (e.g., "mark one or more races"). For example, it is possible that some respondents received the "mark one or more races" instruction first (in Panel A) and thus were very reluctant to answer only one race in the "choose one race" re-contact.

3.4 One household member reports race and Hispanic origin for all members

Household surveys such as the CQS often rely on a single person to respond for all people in a HU. To control for this limitation, we attempted to have the same respondent across all data collections in order to increase the consistency of responses. We asked each respondent to recall who completed the Census 2000 questionnaire, and then used a name matching process - similar to the linkage process described in section 2.3 - to identify whether or not we had the same respondent as in Census 2000. About 59 percent of the respondents were found to be the same in the CQS initial contact as in Census 2000. Every attempt was also made to have the same person respond in both the initial contact and the re-contact. However, we do not have a reliable

measure to identify if we interviewed the same respondent from the initial contact to the re-contact.

3.5 Effects of movers on size of sample reporting Two or more races

The CQS was designed to oversample HUs containing people reporting Two or more races in the census, but we did not control for families who moved into or out of the sample units. Because of this, many HUs in sampling strata 2 and 3 that were expected to have at least one respondent reporting Two or more races did not (see section 4.3). In a much smaller set of cases, individuals from within the single race sampling stratum reported more than one race. Some of these deviations can be explained by general race fluidity and others may be unexplained.

3.6 Possible error associated with linking Census 2000 data to Census Quality Survey data

We cannot be absolutely certain that we correctly linked all CQS individuals to their respective Census 2000 data. This means that some respondents may actually be "false matches." Also, we did not match some respondents. Even so, given the observed racial distributions as seen in the results (section 4), the impact on the results of this evaluation appears to be minimal.

4. RESULTS

In this section, we discuss the final housing unit response rates, the representativeness of the CQS sample, the general consistency in reporting of race, and some tabulations from comparing the responses of the "mark one or more races" instruction to those of the "mark one race" instruction.

4.1 Housing unit response rates

Initially about 27,500 addresses were designated to be in the sample for each panel. Of these, approximately 1,770 turned out to be vacant HUs in each panel. Of the remaining eligible addresses, 97 percent completed a CQS interview in the initial contact. Fewer than 1 percent refused to be interviewed. Other types of noninterviews, including blank returned forms, accounted for 1.8 percent and 1.9 percent of the eligible HUs in Panel A and Panel B, respectively. See Table 4 for further information on the initial contact data collection outcomes.

Table 4. Initial contact data collection outcomes.

Housing Unit Status for Initial Contact	**Panel A ("Mark one or more races")**		**Panel B ("Mark one race")**	
	Number of HUs	Percent	Number of HUs	Percent
Unit Response	24,976	97.1	24,967	97.0
Refusal	225	0.9	220	0.8
Other Noninterview	462	1.8	485	1.9
Unknown Outcome	62	0.2	59	0.2
Total	25,725	100.0	25,731	100.0

In the CQS re-contact, sample HUs were only contacted if an initial contact questionnaire was previously received. Note that the unit response figures in Table 4 do not equal the totals in Table 5 because some initial contact respondents were classified as vacant or out-of-scope in the re-contact. Of the eligible re-contact addresses, 86.9 percent completed a CQS interview in Panel A and 94.2 percent completed an interview in Panel B. The reason for the significant discrepancy is mostly due to the disproportionately high number of unknown outcome cases in Panel A (8.2 percent). The outcome information for these HUs, which were concentrated predominately in Panel A, was erroneously missing and the HUs were not included in the re-contact workload. The impact appears to be minimal given the observed similarity of the initial contact and re-contact race distributions as shown in Tables 6 and 7.

In addition, fewer than 1 percent refused to be interviewed in each panel. Other types of noninterviews, including blank returned forms, accounted for 4.4 percent and 4.0 percent of the eligible HUs in Panel A and Panel B, respectively. See Table 5 for further information on the re-contact data collection outcomes.

Table 5. Re-contact data collection outcomes.

Housing Unit Status for Re-contact	Panel A ("Choose one race")		Panel B ("Choose one or more races")	
	Number of HUs	Percent	Number of HUs	Percent
Unit Response	21,341	86.9	23,160	94.2
Refusal	136	0.6	153	0.6
Other Noninterview	1,072	4.4	979	4.0
Unknown Outcome	2,001	8.2	289	1.2
Total	24,550	100.0	24,581	100.0

4.2 Racial distributions and representativeness of the Census Quality Survey

Analytical results can be biased if the interviewed sample is not representative of the population of interest. Table 6 indicates that aggregated reporting of race among non-Hispanic CQS respondents to the "mark one or more races" instruction closely resembles Census 2000 reporting of race for each panel.[9] No race group is significantly different from those in Census 2000 ($p < 0.1$ level) in either panel, including the Two or more races population, for the contacts where respondents were asked to "mark one or more races" (as highlighted in the tables). A few minor differences exist between Panel A and Panel B, such as the percentage of Blacks in the initial contact of Panel A (10.3 percent) compared to the Panel B re-contact (12.2 percent), however these differences are not significant.

Note that the standard errors associated with the race data in Tables 6 and 7 are shown in parentheses. One-hundred percent data items from the census, such as race and ethnicity, have no standard errors associated with them since a decennial census is an enumeration of the entire population. But, since the CQS used only a sample of the population, standard errors were calculated for each of the estimates for the Census 2000, CQS initial contact, and CQS re-contact data on race.

[9]Refer to Attachment 4 for Census 2000 data on race for both the non-Hispanic and Hispanic populations.

Table 6. Distribution of race using Census 2000, CQS initial contact, and CQS re-contact data (for non-Hispanics only).*

Race	Panel A				Panel B		
	Census 2000**	CQS Initial Contact**	CQS Re-contact	CQS Re-contact w/ probe	Census 2000**	CQS Initial Contact	CQS Re-contact**
White	83.0 (1.12)	82.0 (1.16)	82.2 (1.25)	83.0 (1.22)	81.4 (1.26)	81.3 (1.26)	80.8 (1.35)
Black or African American	10.6 (0.96)	10.3 (0.96)	11.5 (1.09)	11.6 (1.09)	12.5 (1.18)	12.3 (1.17)	12.2 (1.22)
American Indian and Alaska Native	0.6 (0.19)	0.6 (0.14)	0.6 (0.13)	0.7 (0.15)	0.7 (0.16)	0.8 (0.17)	0.7 (0.17)
Asian	3.8 (0.61)	3.9 (0.61)	3.8 (0.62)	3.8 (0.62)	3.4 (0.53)	3.2 (0.48)	3.8 (0.66)
Native Hawaiian and Other Pacific Islander	< 0.1 (0.01)	0.1 (0.02)	0.1 (0.02)	0.1 (0.02)	0.1 (0.03)	0.1 (0.03)	0.1 (0.03)
Some other race	0.2 (0.10)	0.4 (0.14)	0.2 (0.07)	0.2 (0.07)	0.2 (0.14)	0.8 (0.27)	0.3 (0.09)
Two or more races	1.8 (0.18)	2.0 (0.33)	1.4 (0.32)	0.4 (0.13)	1.6 (0.16)	1.0 (0.20)	1.6 (0.24)
Missing or Uncodable	NA	0.7 (0.20)	0.3 (0.09)	0.3 (0.09)	NA	0.5 (0.14)	0.5 (0.15)

NA = not applicable

* The data in Table 6 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The estimates were produced using the inverse of the original sampling probabilities with no adjustment (Z_WGT1) and the standard errors are shown in parentheses.

** The respondents in these columns were given the "mark [choose] one or more races" instruction.

As shown in Table 6, the percentage of Panel A non-Hispanic respondents who reported two or more races in the re-contact, when asked to report only one race, was 1.4 percent initially. This was not statistically different from the 2 percent who did so in the initial contact. The figure decreased to 0.4 percent after these respondents were given the followup race probe. This indicates that there is a sizable proportion of people who will persistently report Two or more races when asked to report only one race. Note that the followup probe was only intended to be asked of Panel A individuals who responded with Two or more races, or used a response such as "multiracial" or "biracial," in the original re-contact race question. The re-contact with followup probe race variable is identical to the original re-contact race variable, except for people who were eligible for and reported a single race to the probe.

In Table 7 we present the race distributions for Hispanics in the CQS sample. Here, Panel B also appears to be representative of Census 2000 in that the race distribution for the re-contact looks very similar to the Census 2000 data for Panel B respondents matched to the census. But, a few differences exist in the race distribution for Panel A's initial contact compared to Census 2000. While a smaller proportion of Hispanics chose White as a single race in the initial contact, it appears that a larger proportion chose "Some other race." Also, the percentage of Hispanics in Census 2000 who reported their race as White differs significantly between Panel A and Panel B ($p < 0.1$ level). These results can be explained by some outlying original sample weights (as discussed in section 2.4.2) which skewed the Hispanic Census 2000 race distribution for Panel A somewhat. However, the distributions are all very similar when the most extreme weights are reduced.

Table 7. Distribution of race using Census 2000, CQS initial contact, and CQS re-contact data (for Hispanics only).*

Race	Panel A				Panel B		
	Census 2000**	CQS Initial Contact**	CQS Re-contact	CQS Re-contact w/ probe	Census 2000**	CQS Initial Contact	CQS Re-contact**
White	57.6 (4.56)	37.0 (4.48)	39.3 (5.04)	41.6 (5.02)	42.7 (4.88)	35.9 (4.51)	42.0 (4.63)
Black or African American	1.4 (1.11)	0.8 (0.38)	1.5 (0.66)	1.6 (0.66)	3.4 (1.93)	1.9 (0.92)	2.4 (1.78)
American Indian and Alaska Native	1.9 (1.52)	2.1 (1.56)	4.1 (2.30)	4.2 (2.30)	2.4 (1.62)	0.1 (0.02)	0.1 (0.02)
Asian	0.3 (0.20)	0.3 (0.21)	0.3 (0.22)	0.3 (0.22)	0.4 (0.35)	0.1 (0.04)	0.1 (0.03)
Native Hawaiian and Other Pacific Islander	< 0.1 (0.02)	< 0.1 (0.01)	0.1 (0.03)	0.1 (0.03)	1.3 (1.27)	0.1 (0.05)	1.5 (0.64)
Some other race	32.7 (4.37)	47.0 (4.86)	48.2 (5.08)	48.8 (5.06)	45.4 (4.99)	50.9 (4.83)	44.1 (4.59)
Two or more races	6.1 (0.92)	6.5 (1.81)	3.6 (1.16)	0.4 (0.15)	4.5 (0.48)	3.9 (2.13)	6.3 (1.75)
Missing or Uncodable	NA	6.2 (3.48)	3.0 (1.47)	3.0 (1.47)	NA	7.1 (2.56)	3.4 (1.54)

NA = not applicable

* The data in Table 7 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The estimates were produced using the inverse of the original sampling probabilities with no adjustment (Z_WGT1) and the standard errors are shown in parentheses.

** The respondents in these columns were given the "mark [choose] one or more races" instruction.

As with the non-Hispanic data, about 0.4 percent of the Hispanics did not report a single race in the re-contact even after the followup probe. In general, unless a probing question is asked, it appears that about half of all Two or more race respondents do not give a single race response. Nonetheless, the data suggest that the race distributions do not change much with the followup probe results.

As shown in Table 8, the weighted percentage of respondents who reported that they were of Hispanic origin was approximately 13 percent in both the initial contact and the re-contact. This figure is slightly higher than the CQS respondents reported in Census 2000, but the difference is not significantly different. In addition, the data on Hispanic origin was statistically similar between Panel A and Panel B.

Table 8. Distribution of Hispanic origin using Census 2000, CQS initial contact, and CQS re-contact data.*

Hispanic Origin	Panel A			Panel B		
	Census 2000**	CQS Initial Contact**	CQS Re-contact	Census 2000**	CQS Initial Contact	CQS Re-contact**
Hispanic	11.1 (0.92)	12.9 (1.03)	13.2 (1.10)	11.1 (0.98)	13.3 (1.02)	13.3 (1.04)

* The estimates in Table 8 were produced using the inverse of the original sampling probabilities with no adjustment (Z_WGT1) and the standard errors are shown in parentheses.
** The respondents in these columns were given the "mark [choose] one or more races" instruction.

4.3 Consistency in reporting of race (Census Quality Survey data from "mark one or more races" instruction)

Of the 1.8 million non-Hispanic[10] people in Panel A reporting Two or more races in Census 2000, only 40 percent also reported Two or more races in the CQS initial contact (see Tables 9 and 10). The other 60 percent reported a single race. This has a significant effect on the number of sample cases available for any analysis of the population reporting Two or more races, since the design of Panel A relies on the initial contact data for identifying respondents who reported more than one race. Instead of having an unweighted sample size of 17,124, Panel A only has 10,013 individuals reporting Two or more races in the initial contact. Even so, the weighted total of people who reported more than one race was similar in Census 2000, 1.8 million, and the CQS initial contact, 2.0 million.

[10]Note that the data presented in Tables 9-14 of the Results section are shown for illustrative purposes and are intended to provide control totals for users of the data file to ensure that they are using the variables correctly. Data for comparable Hispanic tables are provided in Attachment 5.

Also, the inconsistency with reporting race has a small effect in the opposite direction for the unweighted sample. While 36,817 respondents in Panel A reported a single race in Census 2000, 1,978 of these individuals reported Two or more races in the CQS initial contact.

Table 9. Consistency in reporting Two or more races for non-Hispanics for Panel A.*

Census 2000 Race	CQS Initial Contact ("Mark one or more races")		
	Single race	Two or more races	TOTAL
Single race	96,987,813 n= 34,839	1,286,746 n= 1,978	98,274,559 n= 36,817
Two or more races	1,089,924 n= 9,089	**724,686** **n= 8,035**	1,814,610 n= 17,124
TOTAL	98,077,737 n= 43,928	2,011,432 n= 10,013	100,089,169 n= 53,941

* The data in Table 9 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the weighted data were obtained using the inverse of the original sampling probabilities with no adjustment (Z_WGT1).

As Table 9 shows, only 40 percent (724,686/1,814,610) of Panel A respondents reporting Two or more races in Census 2000 also reported Two or more races in the CQS initial contact ("mark one or more races" instruction). Of this 40 percent, though, more than 85 percent (621,015/724,686) reported the same Two or more races in both Census 2000 and the CQS initial contact (e.g., reporting "White ***and*** Asian" at both contacts). Note that when we limited our analysis to the HUs who were identified as having the same respondent in Census 2000 and the CQS initial contact, the results were similar.

In contrast, as shown in Table 10, approximately 97 percent to 98 percent of individuals who reported a single race of White, Black, or Asian in Census 2000 reported the same race in the CQS initial contact. Of the people from whom their race was reported as American Indian or Alaska Native (AIAN), Native Hawaiian or Other Pacific Islander (NHOPI), or Some other race (SOR) in Census 2000, only 55 percent to 58 percent go on to report the same race in the initial contact. The AIAN and NHOPI populations generally have a higher proportion of individuals reporting Two or more races, which may contribute to the lower consistency in the reporting of a single race.

Table 10. Detailed consistency in reporting of race for non-Hispanics for Panel A.*

Census 2000 Race	CQS Initial Contact ("Mark one or more races")							
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more	TOTAL
White	**81,931,427**	130,599	172,418	34,919	666	135,776	792,989	83,198,794
Black	71,372	**10,144,296**	1,061	39,007	437	4,372	131,766	10,392,311
AIAN	37,373	1,015	**339,370**	413	90	338	211,260	589,859
Asian	2,489	353	889	**3,665,710**	15,699	10,592	97,859	3,793,591
NHOPI	1,478	17	0	90	**20,152**	20	14,554	36,311
SOR	41,530	35,062	960	1,412	450	**145,961**	38,318	263,693
Two or more	628,407	116,519	95,289	152,934	31,721	65,054	**724,686**	1,814,610
TOTAL	82,714,076	10,427,861	609,987	3,894,485	69,215	362,113	2,011,432	100,089,169

* The data in Table 10 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the weighted data were obtained using the inverse of the original sampling probabilities with no adjustment (Z_WGT1).

The consistency in reporting of race is similar between Panel A and Panel B. In Panel B, only about 41 percent (565,422/1,391,183) of the individuals who reported Two or more races in Census 2000 went on to also report Two or more races in the CQS re-contact (see Tables 11 and 12). Of this 41 percent, though, more than 76 percent (434,470/565,422) reported the same Two or more races in both Census 2000 and the CQS initial contact. Of the single race reporters, a small portion reported Two or more races in the re-contact. This inconsistency in reporting Two or more races has less of an effect on the Panel B sample size, given that for Panel B one can use the census data for identifying multiple race reporters without having to rely on the CQS. Also, in Panel B, the weighted total of people who reported more than one race was similar in Census 2000 (1.4 million) and the CQS re-contact (1.5 million).

Table 11. Consistency in reporting Two or more races for non-Hispanics for Panel B.*

Census 2000 Race	CQS Re-contact ("Choose one or more races")		
	Single race	Two or more races	TOTAL
Single race	89,881,179 n= 32,848	935,610 n= 1,476	90,816,789 n= 34,324
Two or more races	825,761 n= 8,994	**565,422** **n= 7,148**	1,391,183 n= 16,142
TOTAL	90,706,940 n= 41,842	1,501,032 n= 8,624	92,207,972 n= 50,466

***** The data in Table 11 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the weighted data were obtained using the inverse of the original sampling probabilities with no adjustment (Z_WGT1).

In contrast to the lower consistency in reporting Two or more races, as shown in Table 12, approximately 97 percent to 99 percent who reported a single race in Census 2000 of White, Black, or Asian reported the same race in the re-contact. Among the individuals who reported that their race was American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, or Some other race, 72 percent to 78 percent reported the same race in the re-contact. Note that this percentage is much higher than the 55 percent to 58 percent who reported the same race in the Panel A initial contact and in Census 2000 (from Table 10). This result is interesting in that Panel A initial contact statistics are based on the same data collection methods as Census 2000, while Panel B used telephone interviewing for the re-contact; yet Panel B has the higher consistency in reporting a single race.

Table 12. Detailed consistency in reporting of race for non-Hispanics for Panel B.*

Census 2000 Race	CQS Re-contact ("Choose one or more races")							
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more	TOTAL
White	**74,562,989**	172,545	81,577	215,986	46,302	188,638	430,121	75,698,158
Black	13,447	**10,652,593**	8,725	307	1,118	2,509	321,107	10,999,806
AIAN	88,288	1,183	**508,101**	1,466	4,504	253	102,209	706,004
Asian	56,139	934	227	**3,040,520**	3,652	2,763	19,990	3,124,225
NHOPI	120	53	15	469	**49,539**	98	12,913	63,207
SOR	6,604	3,176	307	2,532	646	**162,854**	49,270	225,389
Two or more	474,088	119,103	46,072	133,512	29,200	23,786	**565,422**	1,391,183
TOTAL	75,201,675	10,949,587	645,024	3,394,792	134,961	380,901	1,501,032	92,207,972

* The data in Table 12 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the weighted data were obtained using the inverse of the original sampling probabilities with no adjustment (Z_WGT1).

4.4 Tabulations of "mark one race" responses by specific combinations of "mark one or more races" responses

In preparation for developing "bridging parameters" using the CQS data, Table 13 shows some results from comparing the Two or more races responses for Panel A with the single race chosen in the re-contact. For Panel A respondents, we compared responses to the initial contact ("mark one or more races" instruction) to the race responses from the re-contact ("choose one race") that used the follow-up race probe for those reporting Two or more races. Even with the followup, a significant portion of respondents still reported Two or more races. The level of this trend varies between each of the specific race pairs.

Table 13. Tabulations for individuals reporting Two or more races in Panel A (for non-Hispanics only).*

CQS Initial Contact ("Mark one or more races")	CQS Re-contact ("Choose one race") with followup probe									
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more (same)	Two or more (different)	Missing	TOTAL
White - Black	**12,499**	**35,591**	34	14	0	18,315	21,611	1,943	15,214	105,222
White - AIAN	**64,664**	909	**34,465**	198	0	1,333	10,411	1,346	15,774	129,101
White - Asian	**64,519**	511	46	**42,472**	524	6,956	32,301	3,721	23,984	175,034
White - NHOPI	**4,244**	48	0	33	**5,932**	269	1,201	489	1,260	13,476
White - SOR	**22,736**	338	0	0	35	**3,294**	1,173	679	4,378	32,634
Black - AIAN	43	**11,217**	**2,913**	0	0	492	2,521	1,005	2,690	20,880
Black - Asian	11	**8,049**	6	**1,442**	0	524	1,646	749	1,889	14,317
Black - NHOPI	23	**1,575**	0	194	**511**	111	0	185	0	2,598
Black - SOR	700	**11,080**	0	628	0	**1,612**	200	787	535	15,541
AIAN - Asian	701	57	**627**	**775**	52	39	137	58	112	2,559
AIAN - NHOPI	0	0	**20**	0	**178**	0	31	0	29	258
AIAN - SOR	648	38	**217**	0	0	**70**	0	57	288	1,318
Asian - NHOPI	482	225	0	**6,318**	**11,698**	113	2,486	539	3,039	24,900
Asian - SOR	1,147	364	518	**5,758**	0	**1,274**	645	1,269	260	11,234
NHOPI - SOR	0	1,095	0	405	**650**	**114**	0	938	54	3,257
Three or more	10,721	10,390	2,421	4,681	12,637	7,791	4,436	8,049	11,096	72,221
TOTAL	183,138	81,487	41,267	62,919	32,217	42,307	78,800	21,813	80,603	624,550

* The data in Table 13 were restricted to people who were identified as Two or more races in the initial contact and where the races were not imputed for those matched to Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the data were restricted to cases in sampling strata 2 and 3 (Z_WGT2).

Table 14 shows some race tabulations for Two or more races individuals in Panel B by the race reported in the initial contact. For Panel B respondents, we compared the Census 2000 response ("mark one or more races" instruction) to the race responses from the initial contact ("mark one race"). Some of the general results are similar to those in Table 13.[11]

Table 14. Tabulations for individuals reporting Two or more races in Panel B (for non-Hispanics only).*

Census 2000 Race ("Mark one or more races")	CQS Initial Contact ("Mark one races")									
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more (same)	Two or more (different)	Missing	TOTAL
White - Black	**18,305**	**48,769**	117	107	63	21,841	39,935	6,126	1,862	137,126
White - AIAN	**133,746**	484	**54,952**	54	0	2,807	33,070	3,283	2,170	230,566
White - Asian	**53,311**	317	262	**66,680**	675	12,585	65,682	11,093	940	211,546
White - NHOPI	**8,725**	0	0	432	**7,515**	860	4,627	1,910	535	24,604
White - SOR	**130,574**	1,139	378	5,076	35	**25,297**	4,117	4,233	663	171,512
Black - AIAN	131	**24,944**	**5,094**	0	0	651	4,493	1,672	943	37,927
Black - Asian	295	**9,522**	89	**4,890**	32	1,555	6,075	1,619	106	24,183
Black - NHOPI	0	**2,888**	0	0	**580**	181	22	130	0	3,801
Black - SOR	2,402	**39,352**	55	1,195	269	**8,824**	2,235	2,206	0	56,537
AIAN - Asian	23	90	**1,057**	**5,913**	26	686	1,056	586	20	9,458
AIAN - NHOPI	151	0	**172**	0	**478**	104	115	0	0	1,021
AIAN - SOR	2,037	64	**1,572**	631	0	**537**	0	201	0	5,041
Asian - NHOPI	595	175	0	**12,421**	**9,459**	35	9,002	3,527	330	35,543
Asian - SOR	1,446	541	324	**26,337**	2,748	**3,259**	2,082	5,971	23	42,730
NHOPI - SOR	310	0	0	0	**600**	**232**	0	0	0	1,142
Three or more	11,304	15,222	2,111	7,957	9,993	7,831	7,380	20,072	344	82,215
TOTAL	363,354	143,506	66,184	131,693	32,475	87,285	179,890	62,629	7,936	1,074,952

* The data in Table 14 were restricted to matched people who were identified as Two or more races in Census 2000 and where the races were not imputed - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the data were restricted to cases in sampling strata 2 and 3 (Z_WGT2).

[11]Note that the data presented in Tables 9-14 of the Results section are shown for illustrative purposes and are intended to provide control totals for users of the data file to ensure that they are using the variables correctly. Data for comparable Hispanic tables are provided in Attachment 5.

ACKNOWLEDGMENTS

The authors of this report would like to thank the rest of the members of the B.3 Analysis Team for their assistance in developing and reviewing the CQS data file and this report:

Julie Bouffard
Arthur Cresce Jr.
Jorge del Pinal
Joan M. Hill
David Hubble
Nicholas Jones
Elizabeth Martin
Lance Moore
James Poyer
Amy Symens Smith
Henry Woltman

REFERENCES

American Association for Public Opinion Research (2000) *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, Ann Arbor, Michigan: AAPOR.

Anderson, M. (2000) *Encyclopedia of the U.S. Census*, Congressional Quarterly: Washington, DC, p. 313.

Bates, N., Martin, E. A., DeMaio, T. J., and de la Puente, M. (1995) "Questionnaire effects on measurements of race and Hispanic origin," *Journal of Official Statistics*, Vol.11 p. 433-459.

Berning, M. (2002) *Matching Process for the Census Quality Survey to Evalaute Responses to the Census 2000 Question on Race*, U.S. Census Bureau internal document.

Davis, D.K.; Blair, J.; Crawley, E.L.;Craig, K.M.; Rappaport, M.S.B.; Baker, C.A. & S.L. Hanson (2001) "Census 2000 Quality Survey Instrument Pilot Test," Development Associates, Inc. March 2001.

Davis, D. K.; Blair, J.; Penelziek, M. & N. Garas (2002) "Census Quality Survey Telephone Interview Behavior Coding Study," Development Associates, Inc. September 24, 2002.

del Pinal, Jorge; Martin, Elizabeth; Bennett, Claudette & Art Cresce (2002) "Overview of Results of New Race and Hispanic Origin Questions in Census 2000," 2002 Proceedings of the American Statistical Association, Survey Research Method Section [CD-ROM], Alexandria, VA: American Statistical Association.

Fay, Robert E. (1998) "VPLX: Program Documentation," U.S. Bureau of the Census internet site, http://www.census.gov/sdms/www/vwelcome.html.

Gerber, E., de la Puente, M., and Levin, M. (1998) *Race, Identity, and New Question Options: Final Report of Cognitive Research on Race and Ethnicity*. Unpublished Census Bureau report, Feb. 24, 1998.

Goldstein, Joshua R. and Ann J. Morning. (2000) "The Multiple Race Population of the United States: Issues and Estimates," Proceedings of the National Academy of Sciences. Vol. 97, No. 11, p. 6230-5.

Grieco, Elizabeth M. (2001) "An Evaluation of Bridging Methods Using Race Data from Census 2000," Currently under review for publication with Population Research and Policy Review.

Grieco, Elizabeth M. and Rachel C. Cassidy. (2001) "Overview of Race and Hispanic Origin: 2000 - Census 2000 Brief," U.S. Census Bureau.

Hahn, Robert (1993) *Rethinking "Race" and Ethnicity" for Public Health Surveillance*. Paper presented at American Anthropological Association Meeting. November 1993. Washington, DC

Harris, David R. (2000) *Does It Matter How We Measure? Racial Classification and the Characteristics of Multiracial Youth*, Paper Presented at the Levy Institute Conference on Multiraciality: How Will the New Census Data Be Used? (Draft).

Harris, David R. and Justin L. Thomas. (2001) *The Educational Costs of Being Multiracial: Evidence from a National Survey of Adolescents*, Presented at the Annual Meetings of the Population Association of America.

Harrison, Roderick J. (2000) *Inadequacies of Multiple Response Race Data in the Federal Statistical System*, Paper Presented at the Levy Institute Conference on Multiraciality: How Will the New Census Data Be Used? (Draft).

Johnson, Charles E. (1974). *Consistency of Reporting of Ethnic Origin in the Current Population Survey*. Washington, DC: Bureau of the Census; U.S. Department of Commerce, Technical Paper No. 31.

Jones, Nicholas and Amy Symens Smith. (2001) *The Two or More Races Population: 2000 - Census 2000 Brief*, U.S. Census Bureau.

Kissam, E., Herrera, E., and Nakamoto, J. M. (1993) *Hispanic Response to Census Enumeration Forms and Procedures*. Report prepared by Aguirre International for the Census Bureau.

Lucas, Jacqueline. (2000) "OMB's New Classification of Race and Ethnicity and its Potential Impact on Health Research," National Center for Health Statistics. Presentation to the Census Bureau B.3 Working Group, August 29, 2000.

Lucas, J. W., Parker, J. D., and Kingston, R. (1999) *Black, White, and Shades of Gray: The Impact of the New Standards for Race and Ethnicity Data Collection on Assessing Race and Health Outcomes in Epidemiologic Research*. Paper presented at the annual meeting of the American Public Health Association, Nov. 8-11, Chicago IL.

Madans, Jennifer H. (2000) *Monitoring the Nation's Health Amid Changing Standards for the Collection of Data on Race and Ethnicity: The NCHS Experience*, National Center for Health Statistics. Paper Presented at the Levy Economics Institute Conference on "Multiraciality: How Will the New Census Data Be Used?" (Draft). Annandale on Hudson, New York. September 22-23, 2000.

Martin, E. (2002) *Questionnaire Effects on Reporting of Race and Hispanic Origin: Results of a Replication of the 1990 Mail Short Form in Census 2000*. Final report of Census 2000 Alternative Questionnaire Experiment. U.S. Census Bureau.

Martin, E. A., de la Puente, M., and Bennett, C. (2001) "The Effects of Questionnaire and Content Changes on Responses to Race and Hispanic Items: Results of a Replication of the 1990 Census Short Form in Census 2000." *Proceedings of the American Statistical Association (Section on Survey Research Methods Section)*. Alexandria VA: ASA.

McKenney, N. R., Fernandez, E. W. and Masamura, W. T. (1985) "The quality of the race and Hispanic origin information reported in the 1980 census." *Proceedings of the American Statistical Association (Section on Survey Research Methods Section)*, pp. 46-50.

Office of Management and Budget (1977) "Race and Ethnic Standards for Federal Statistics and Administrative Reporting," Statistical Policy Directive No. 15. pp. 37-38. May 12, 1977.

Office of Management and Budget. (1994) "OMB Announces Review of Federal Racial and Ethnic Categories and Related Public Hearings." June 9, 1994.

Office of Management and Budget. (1995) "Standards for the Classification of Federal Data on Race and Ethnicity," Federal Register Notice (8/28/95, Vol. 60, No. 166). http://www.whitehouse.gov/omb/fedreg/race-ethnicity.html.

Office of Management and Budget (1997a) "Recommendations from the Interagency Committee for the Review of the Racial and Ethnic Standards to the Office of Management and Budget Concerning Changes to the Standards for the Classification of Federal Data on Race and Ethnicity," Federal Register Notice (7/9/97, Part II).

Office of Management and Budget. (1997b) "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity," OMB Revised Statistical Policy Directive No. 15, Federal Register Notice (10/31/97, Part II), p. 58782-58790.

Office of Management and Budget. (2000) "Federal Register Notice: Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity," Federal Register Vol. 66, No. 10, p. 3829-31 (with three appendices). http://www.whitehouse.gov/omb/inforeg/statpolicy.html#dr.

Raglin, D. A. and Leslie, T. F. (2002) "How Consistent is Race Reporting Between the Census and the Census 2000 Supplementary Survey?" *Proceedings of the American Statistical Association (Section on Survey Research Methods Section)*, pp. 2826-2831.

Smiley, R., and Keeley, C. (1997) *1996 Community Census--ICM Behavior Coding Report*. 1996 Community Census Results Memorandum No. 23. Washington DC: Census Bureau.

Tucker, C.; McKay R.; Kojetin, B.; Harrison, R.,;de la Puente, M.; M. Stinson, L., & E. Robinson (1996) "Testing Methods of Collecting Racial and Ethnic Information: Results of the Current Population Supplement on Race and Ethnicity". Bureau of Labor Statistics, BLS Statistical Notes No. 40.

U.S. Census Bureau. (1989) *200 Years of Census Taking: Population and Housing Questions, 1790-1990*. November 1989.

U.S. Census Bureau (1996) *Findings on Questions on Race and Hispanic Origin Tested in the 1996 National Content Survey*, Population Division Working Paper No. 16.

U.S. Census Bureau (1997) *Results of the 1996 Race and Ethnic Targeted Test*, Population Division Working Paper No. 18.

U.S. Census Bureau (2000) *Coding Ethnic Origin Write-in Responses Given to the Census 2000 Question on Race*, Census 2000 Decision Memorandum No. 106, July 6, 2000.

U.S. Census Bureau (2001a) "Table 4. Difference in Population by Race and Hispanic or Latino Origin, for the United States: 1990 to 2000." http://www.census.gov/population/cen2000/phc-t1/tab04.pdf. Internet release date: April 2, 2001.

U. S. Census Bureau. (2001b) *Census 2000 Overview Brief*. http://www.census.gov/prod/2001pubs/c2kbr01-1.pdf, Table 9.

ATTACHMENT 1

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
1	5	5	**Z_ID**	Sample ID Number	00001 - 49957	This household level identification variable was randomly assigned to each housing unit.
7	8	2	**Z_PNUM**	Person Number	01 - 24	
10	19	10	**Z_WGT1**	Original Sample Weight	This is the initial mailout weight; that is, the inverse of the probability of selection.	The data user should use Z_WGT1 when examining race consistency or general race distributions. Note that this field is numeric with no leading zeros and two decimal places.
21	30	10	**Z_WGT2**	Bridging Weight	For cases where Z_STRAT eq '2' or '3', Z_WGT2 = the original sample weight. For cases where Z_STRAT eq '1' or '4', Z_WGT2 = 0.	The data user should use Z_WGT2 when examining race bridging, or allocation parameters, to a single race. Note that this field is numeric with no leading zeros and two decimal places.
32	32	1	**Z_PANEL**	Panel	1 = Panel A 2 = Panel B	

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
34	34	1	**Z_STRAT**	Sample Stratum	1 = Consists of HUs containing ALL single race persons pre-edit/allocation, or HUs enumerated as vacant in Census 2000. 2 = Consists of HUs containing one or more persons having a race code consisting of SOR plus ONE other OMB category. 3 = Consists of HUs containing one or more persons having a race code consisting of two OMB races, or three or more of any race (including SOR).	
36	37	2	**Z_STFIPS**	State FIPS Code	01 - 56 01 = AL, 02 = AK, 04 = AZ, 05 = AR, 06 = CA, 08 = CO, 09= CT, 10 = DE, 11 = DC, 12 = FL, 13 = GA, 15 = HI, 16 = ID, 17 = IL, 18 = IN, 19 = IA, 20 = KS, 21 = KY, 22 = LA, 23 = ME, 24 = MD, 25 = MA, 26 = MI, 27 = MN, 28 = MS, 29 = MO, 30 = MT, 31 = NE, 32 = NV, 33 = NH, 34 = NJ, 35 = NM, 36 = NY, 37 = NC, 38 = ND, 39 = OH, 40 = OK, 41 = OR, 42 = PA, 44 = RI, 45 = SC, 46 = SD, 47 = TN, 48 = TX, 49 = UT, 50 = VT, 51= VA, 53 = WA, 54 = WV, 55 = WI, 56 = WY	
39	40	2	**Z_STPOST**	State Postal Abbreviation	AL - WY	

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
42	42	1	**Z_DIV**	Census Division	1 = Northeast 2 = Middle Atlantic 3 = South Atlantic 4 = East South Central 5 = East North Central 6 = West South Central 7 = West North Central 8 = Mountain 9 = Pacific	
44	44	1	**Z_REGION**	Census Region	1 = Northeast 2 = South 3 = Midwest 4 = West	
46	46	1	**Z_IMODE**	Data Collection Mode, Initial Contact	1 = Mailout/Mailback 2 = Personal Visit	
48	48	1	**Z_RMODE**	Data Collection Mode, Re-contact	1 = Telephone 2 = Personal Visit 9 = No re-contact data	
50	50	1	**Z_IHISP**	Hispanic Origin, Initial Contact	0 = Non-Hispanic 1 = Hispanic 9 = Missing, unknown	
52	52	1	**Z_RHISP**	Hispanic Origin, Re-contact	0 = Non-Hispanic 1 = Hispanic 9 = Missing, unknown	

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
54	54	1	**Z_CHISP**	Hispanic Origin, Census	0 = Non-Hispanic 1 = Hispanic 9 = Missing, unknown	
56	56	1	**Z_SEX**	Sex	1 = Male 2 = Female 9 = Missing, unknown	Note that Z_SEX is a composite edit from the Census, Initial Contact, and Re-contact.
58	59	2	**Z_AGE**	Age Category	00 = 00-04, 05 = 05-09, 10 = 10-14, 15 = 15-17 18 = 18-19, 20 = 20-24, 25 = 25-29, 30 = 30-34 35 = 35-39, 40 = 40-44, 45 = 45-49, 50 = 50-54 55 = 55-59, 60 = 60-64, 65 = 65-69, 70 = 70-74 75 = 75-79, 80 = 80-84, 85 = 85+ 99 = Missing, unknown	Note that Z_AGE is a composite edit from the Census, Initial Contact, and Re-contact.
61	61	1	**Z_EDU**	Educational Attainment	1 = Not a HS Graduate 2 = High School Graduate - diploma or GED 3 = Some college, no degree 4 = Associate Degree 5 = Bachelors Degree 6 = Graduate or professional Degree 8 = Not in universe (age <25 or unknown) 9 = Missing, unknown	Note that Z_EDU is a composite edit from the Census and Re-contact.

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
63	64	2	**Z_RREL**	Relationship to Householder, Re-contact	01 = Householder 02 = Spouse 03 = Natural born child 04 = Adopted child 05 = Stepchild 06 = Brother or sister 07 = Parent 08 = Grandchild 09 = Other relative 10 = Roomer/boarder or foster child 11 = Housemate/roommate 12 = Unmarried partner 13 = Other nonrelative 14 = Institutional GQ person 15 = Noninstitutional GQ person 99 = Missing, unknown	Note that the respondent is generally assumed to be the householder, with some exceptions introduced during the editing process.
66	67	2	**Z_CREL**	Relationship to Householder, Census	(see description for Z_RREL above)	
69	69	1	**Z_MILIT**	Veteran's Status	1 = No military service 2 = Served in military, previously or currently 8 = Not in universe (age <18 or unknown) 9 = Missing, unknown	Note that Z_MILIT is a composite edit from the Census and Re-contact.

ATTACHMENT 1

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
71	71	1	**Z_INCOME**	Household Income	1 = Less than $10,000 2 = $10,000 - $24,999 3 = $25,000 - $34,999 4 = $35,000 - $49,999 5 = $50,000 - $69,999 6 = $70,000 - $99,999 7 = $100,000 or more 9 = Missing, unknown	Note that Z_INCOME is a composite edit from the Census and Re-contact.
73	74	2	**Z_IRACE**	Race, Initial Contact	01 - 63 99 = Missing, unknown	Refer to the Race Code List (Footnote 4) for a complete description of race codes 01 - 63.
76	77	2	**Z_RRACE**	Race, Re-contact	01 - 63 99 = Missing, unknown	See Footnote 4.
79	80	2	**Z_R2RACE**	Race, Re-contact with Followup Probe	01 - 63 99 = Missing, unknown	Note that for the majority of the records, Z_R2RACE = Z_RRACE. The followup race probe was only asked of Panel A individuals who responded with Two or more races, or used a response such as "multiracial" or "biracial", in the initial re-contact race question. See Footnote 4.
82	83	2	**Z_CRACE**	Race, Census	01 - 63 99 = Missing, unknown	See Footnote 4.

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
85	85	1	**Z_CRFLAG**	Census Allocated Race Flag	0 = As reported 1 = Code changed through consistency edit 3 = Classified from race response in Hispanic question 4 = Allocated from within household 5 = Allocated from hot deck 9 = Missing, unknown	
87	87	1	**Z_CHFLAG**	Census Allocated Hispanic Origin Flag	0 = As reported 2 = Multiple response given a unique Hispanic or Non-Hispanic code 3 = Assigned Hispanic from race code 4 = Allocated from within household 5 = Allocated from hot deck (surname used) 6 = Allocated from hot deck (surname not used) 9 = Missing, unknown	
89	90	2	**Z_IRSO**	Race, Initial Contact (SOR removed)	01 - 63 99 = Missing, unknown	Note that for each of: Z_IRSO, Z_RRSO, Z_R2SO, Z_CRSO, Some other race (SOR) has been removed from all Two or more race responses. The single race SOR remains. See Footnote 4.
92	93	2	**Z_RRSO**	Race, Re-contact (SOR removed)	01 - 63 99 = Missing, unknown	See Footnote 4.
95	96	2	**Z_R2RSO**	Race, Re-contact with Followup Probe (SOR removed)	01 - 63 99 = Missing, unknown	See Footnote 4.

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
98	99	2	**Z_CRSO**	Race, Census (SOR removed)	01 - 63 99 = Missing, unknown	See Footnote 4.
101	101	1	**Z_SMRSP**	Same Respondent, Census to Initial Contact	0 = Not the Same or unknown 1 = Same	Note that Z_SMRSP is based on respondent self-response of which household member filled out the Census form.
103	103	1	**Z_NHWTRC**	Tract-level, Non-Hispanic White	0 = 'Low' 1 = 'High'	Note that Z_NHWTRC was computed based on the state-level medians of the proportion of Non-Hispanic Whites in a tract.

*** Footnote 1: For several of the variables, we used a "composite edit."**
This edit incorporated an algorithm which was designed to reconcile discrepant or missing responses among the CQS Initial Contact, Re-contact, and Census 2000.

*** Footnote 2: Missing or unknown values for Census 2000 variables indicate that the respondent was not matched to their Census 2000 data.**

*** Footnote 3: The character "Z" as a prefix to each of the names is only used for internal purposes to distinguish between the final and original variables.**
Users may name the variables as they see fit.

*** Footnote 4: Race Code List**

1 = White		42 = White; Black; AIAN; Asian	
2 = Black		43 = White; Black; AIAN; NHOPI	
3 = American Indian or Alaska Native (AIAN)		44 = White; Black; AIAN; SOR	
4 = Asian		45 = White; Black; Asian; NHOPI	

Census Quality Survey (CQS) Public-use Datafile Layout

BEG	END	LEN	VARIABLE	DESCRIPTION	VALUE DESCRIPTION	EXPLANATORY NOTE
5 = Native Hawaiian or Other Pacific Islander (NHOPI)					46 = White; Black; Asian; SOR	
6 = Some other race (SOR)					47 = White; Black; NHOPI; SOR	
					48 = White; AIAN; Asian; NHOPI	
7 = White; Black				22 = White; Black; AIAN	49 = White; AIAN; Asian; SOR	
8 = White; AIAN				23 = White; Black; Asian	50 = White; AIAN; NHOPI; SOR	
9 = White; Asian				24 = White; Black; NHOPI	51 = White; Asian; NHOPI; SOR	
10 = White; NHOPI				25 = White; Black; SOR	52 = Black; AIAN; Asian; NHOPI	
11 = White; SOR				26 = White; AIAN; Asian	53 = Black; AIAN; Asian; SOR	
12 = Black; AIAN				27 = White; AIAN; NHOPI	54 = Black; AIAN; NHOPI; SOR	
13 = Black; Asian				28 = White; AIAN; SOR	55 = Black; Asian; NHOPI; SOR	
14 = Black; NHOPI				29 = White; Asian; NHOPI	56 = AIAN: Asian; NHOPI; SOR	
15 = Black; SOR				30 = White; Asian; SOR		
16 = AIAN; Asian				31 = White; NHOPI; SOR	57 = White; Black; AIAN; Asian; NHOPI	
17 = AIAN; NHOPI				32 = Black; AIAN; Asian	58 = White; Black; AIAN; Asian; SOR	
18 = AIAN; SOR				33 = Black; AIAN; NHOPI	59 = White; Black; AIAN; NHOPI; SOR	
19 = Asian; NHOPI				34 = Black; AIAN; SOR	60 = White; Black; Asian; NHOPI; SOR	
20 = Asian; SOR				35 = Black; Asian; NHOPI	61 = White; AIAN; Asian; NHOPI; SOR	
21 = NHOPI; SOR				36 = Black; Asian; SOR	62 = Black; AIAN; Asian; NHOPI; SOR	
				37 = Black; NHOPI; SOR		
				38 = AIAN; Asian; NHOPI	63 = White; Black; AIAN; Asian; NHOPI; SOR	
				39 = AIAN; Asian; SOR		
				40 = AIAN; NHOPI; SOR		
				41 = Asian; NHOPI; SOR		

ATTACHMENT 2

Exact Wording of Census 2000 and CQS Questions on Race and Hispanic Origin

Mail questionnaire for Census 2000 (Both Panels) and CQS Initial Contact Panel A:

Note: Please answer BOTH Questions 8 and 9. [Questions 7 and 8 for Census 2000]

8. **Is this person Spanish/Hispanic/Latino?** Mark the "No" box if **not** Spanish/Hispanic/Latino.

☐ No, not Spanish/Hispanic/Latino ☐ Yes, Puerto Rican
☐ Yes, Mexican, Mexican Am., Chicano ☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino -- Print group.

9. **What is this person's race? Mark one or more races** to indicate what this person considers himself/herself to be.

☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native -- Print name of enrolled or principal tribe.

☐ Asian Indian	☐ Japanese	☐ Native Hawaiian
☐ Chinese	☐ Korean	☐ Guamanian or Chamorro
☐ Filipino	☐ Vietnamese	☐ Samoan
☐ Other Asian -- Print race.		☐ Other Pacific Islander -- Print race.

☐ Some other race -- Print race.

Mail questionnaire for CQS Initial Contact Panel B:

Note: Please answer BOTH Questions 8 and 9.

7. **Is this personSpanish/Hispanic/Latino?** Mark the "No" box if **not** Spanish/Hispanic/Latino.

☐ No, not Spanish/Hispanic/Latino ☐ Yes, Puerto Rican
☐ Yes, Mexican, Mexican Am., Chicano ☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino -- Print group.

8. **What is this person's race? Mark one race** to indicate what this person considers himself/herself to be.

☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native -- Print name of enrolled or principal tribe.

☐ Asian Indian	☐ Japanese	☐ Native Hawaiian
☐ Chinese	☐ Korean	☐ Guamanian or Chamorro
☐ Filipino	☐ Vietnamese	☐ Samoan
☐ Other Asian -- Print race.		☐ Other Pacific Islander -- Print race.

☐ Some other race -- Print race.

ATTACHMENT 2

Exact Wording of Census 2000 and CQS Questions on Race and Hispanic Origin

Enumerator questionnaire for Census 2000 (Both Panels) and CQS Initial Contact Panel A:

(ENUMERATOR NOTE: It is important to answer BOTH questions 4 and 5 and show Cards 1 and 2.)[12]
[Questions 5 and 6, and Cards B and C for Census 2000]

4. Are any of the persons that I have listed Mexican, Puerto Rican, Cuban, or of another Hispanic or Latino group?

☐ No, not Spanish/Hispanic/Latino ☐ Yes, Puerto Rican
☐ Yes, Mexican, Mexican Am., Chicano ☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino -- **What is this group?**

5. Now choose one race for each person. Which race does each person consider himself/herself to be?

☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native -- **What is the name of (your/...'s) enrolled or principal tribe?**

☐ Asian Indian	☐ Japanese	☐ Native Hawaiian
☐ Chinese	☐ Korean	☐ Guamanian or Chamorro
☐ Filipino	☐ Vietnamese	☐ Samoan
☐ Other Asian -- Print race.		☐ Other Pacific Islander -- **What is this race?**

☐ Some other race -- **What is this race?**

Enumerator questionnaire for CQS Initial Contact Panel B:

(ENUMERATOR NOTE: It is important to answer BOTH questions 4 and 5 and show Cards 1 and 2.)

4. Are any of the persons that I have listed Mexican, Puerto Rican, Cuban, or of another Hispanic or Latino group?

☐ No, not Spanish/Hispanic/Latino ☐ Yes, Puerto Rican
☐ Yes, Mexican, Mexican Am., Chicano ☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino -- **What is this group?**

5. Now choose one or more races for each person. Which race or races does each person consider himself/herself to be?

☐ White
☐ Black, African Am., or Negro
☐ American Indian or Alaska Native -- **What is the name of (your/...'s) enrolled or principal tribe?**

☐ Asian Indian	☐ Japanese	☐ Native Hawaiian
☐ Chinese	☐ Korean	☐ Guamanian or Chamorro
☐ Filipino	☐ Vietnamese	☐ Samoan
☐ Other Asian -- Print race.		☐ Other Pacific Islander -- **What is this race?**

☐ Some other race -- **What is this race?**

[12]For personal visit interviews, Census Bureau enumerators use show cards to help communicate to the respondent the instructions and the categories for the questions on race and ethnicity.

ATTACHMENT 2

Exact Wording of Census 2000 and CQS Questions on Race and Hispanic Origin

Telephone and personal visit questionnaires for CQS Re-contact Panel A:

6a. **[Are you/is** Person #**] Spanish, Hispanic or Latino?**
☐ Yes (- Ask 6b)
☐ No (- Ask 7a)

6b. **Which of the following groups [do you/does** Person #**] belong to:**
(Read list below. Accept multiple responses.)
☐ Mexican, Mexican American, or Chicano? ☐ Puerto Rican? ☐ Cuban?
☐ Some other Spanish, Hispanic, or Latino group? (- 6c)**What group is that?**

7a. **Now, I'd like you to tell me what race [you consider yourself to be] [he/she considers himself/herself to be]. Please choose one of the following 6 race categories:**
(Continue reading list even if respondent breaks in with an answer. Accept multiple responses if offered.)
☐ White?
☐ Black, African American, or Negro?
☐ American Indian or Alaska Native? (- Ask 7b)
☐ Asian? (- Ask 7c)
☐ Native Hawaiian or Other Pacific Islander? (- Ask 7e)
☐ Some other race? (- Ask 7g)

7b. (If American Indian or Alaska Native) **What is the name of [your/his/her] enrolled or principal tribe?**

7c. (If Asian) **Which of the following Asian groups [are you/is he/she]?**
☐ Asian Indian? ☐ Japanese?
☐ Chinese? ☐ Korean?
☐ Filipino? ☐ Vietnamese?
☐ Some other Asian group? (- 7d) **What is that group?**

7e. (If NHOPI) **Which of the following Native Hawaiian or Other Pacific Islander groups [are you/is he/she]?**
☐ Native Hawaiian? ☐ Samoan? ☐ Guamanian or Chamorro?
☐ Some other Pacific Islander group? (- 7f) **What group is that?**

7g. (If some other race) **What other race group is that?**

8a. (Did respondent offer more than one race in 7a or 7g OR use a term such as multiracial/biracial in 7g?)
☐ Yes (- Ask 8b) ☐ No (- Ask 9)

8b. **When asked to choose only one race from the original list I read you, what [do you/does** Person #**] usually answer?**

8c. **Why is that?**

ATTACHMENT 2

Exact Wording of Census 2000 and CQS Questions on Race and Hispanic Origin

Telephone and personal visit questionnaires for CQS Re-contact Panel B:

6a. **[Are you/is** Person #**] Spanish, Hispanic or Latino?**
☐ Yes (- Ask 6b)
☐ No (- Ask 7a)

6b. **Which of the following groups [do you/does** Person #**] belong to:**
(Read list below. Accept multiple responses.)
☐ Mexican, Mexican American, or Chicano? ☐ Puerto Rican? ☐ Cuban?
☐ Some other Spanish, Hispanic, or Latino group? (- 6c)**What group is that?**

7a. **Now, I'd like you to tell me what race or races [you consider yourself to be] [he/she considers himself/herself to be]. Please choose one or more of the following 6 race categories:**
(Continue reading list even if respondent breaks in with an answer.)
☐ White?
☐ Black, African American, or Negro?
☐ American Indian or Alaska Native? (- Ask 7b)
☐ Asian? (- Ask 7c)
☐ Native Hawaiian or Other Pacific Islander? (- Ask 7e)
☐ Some other race? (- Ask 7g)
(Mark all races mentioned by respondent.)

7b. (If American Indian or Alaska Native) **What is the name of [your/his/her] enrolled or principal tribe?**

7c. (If Asian) **Which of the following Asian groups [are you/is he/she]?**
☐ Asian Indian? ☐ Japanese?
☐ Chinese? ☐ Korean?
☐ Filipino? ☐ Vietnamese?
☐ Some other Asian group? (- 7d) **What is that group?**

7e. (If NHOPI) **Which of the following Native Hawaiian or Other Pacific Islander groups [are you/is he/she]?**
☐ Native Hawaiian? ☐ Samoan? ☐ Guamanian or Chamorro?
☐ Some other Pacific Islander group? (- 7f) **What group is that?**

7g. (If some other race) **What other race group is that?**

7h. (Did respondent offer more than one race?)
☐ Yes (- Ask 7b-7g as appropriate - make sure ALL categories that were marked were covered)
☐ No (- Ask 9a)

ATTACHMENT 3
(Date written: 7/27/2000)

Methodological Criteria and Proposed Sample Design Options

Criteria for the Ideal Measurement for the Evaluation of Responses to the Census 2000 Question on Race

Research Objective

The objective of the research is to produce data that will improve users' ability to make comparisons between Census 2000 data on race that allowed the reporting of one or more races, and data on race from other sources that allow only single race responses. The primary goal is to improve comparisons of 1990 and 2000 Census race distributions, at national and lower levels of geography. Other goals are to facilitate comparisons between Census 2000 and Census Bureau surveys which instruct respondents to mark one race, and with data from the vital records system, which uses census data to calculate such indicators as birth and death rates.

The intent is to provide users with a data file which cross-classifies race responses using two measurement methods: the Census 2000 question on race that allows reporting more than one race, and a question on race comparable to that used in the 1990 Census. This data file will enable users to develop "bridging" methods, including modeling single race distributions using Census 2000 data in order to make historical comparisons with previous censuses or with single race distributions from other sources.

The Ideal Survey Method

For various reasons, the ideal conditions for collecting data to meet the objective do not exist. But just the same, we believe that understanding the characteristics of the ideal method would be useful in comparing alternative options for conducting the evaluation. With this in mind, we identified 6 criteria to be used for the purpose of assessing the merits of proposed options.

So, ideally, the measurement of responses to a "mark one or more" instruction for the question on race and responses to a "mark one" instruction would be collected in such a way as to meet the criteria on the following page.

ATTACHMENT 3
(Date written: 7/27/2000)

Methodological Criteria and Proposed Sample Design Options

6 Criteria for Assessing Design Options: Description of Criteria

1. **Simultaneously:** Real changes in how people report race may occur as time passes between the two data collections (possibly more so for persons marking more than one race). Also, due to other survey methods factors, respondents may report different races in the two measurements. Ideally, we want to collect both responses at the same time under the same general survey conditions.

2. **Independently:** Response to the question on race using either instruction can be influenced by the response to the other race question when both questions are asked in the same survey instrument. In addition, the ordering of the questions can have an effect. Ideally, we do not want the measurement of one question on race to be modified just by the presence of the other question on race.

3. **In the Census 2000 Environment:** There are many factors associated with Census 2000 that can affect responses to the question on race. A major factor is the Bureau's extensive promotion and partnership program. Ideally, we want to collect the evaluation data in that environment, as well.

4. **Using Comparable Methods:** The survey methods used to collect the data can affect how people respond to the question on race. This includes the data collection mode (i.e., mail, personal visit, telephone, etc.) and the wording and ordering of the questions. Ideally, we want to collect the "mark one or more" instruction to the question on race using Census 2000 methods and the "mark one" instruction using 1990 Census methods.

5. **From the Same Respondent:** Providing responses to the question on race for yourself and other household members can be a sensitive issue. Reporting of race for other household members by a household respondent can be influenced by the respondents race, relationship to other household members, and other factors. Within a household, different respondents may provide different responses to the question on race. This may be especially true for people who consider themselves as multiracial. Ideally, we want to collect responses to both questions on race from the same household respondent.

6. **A Representative and Sufficient Sample:** Analytical results can be biased if the selected or "interviewed" sample is not representative of the population of interest. Using Census 2000 data, we have control over the selection of the sample households. An initially representative sample of households will be selected. Of more concern is the ability to collect responses from all individuals in our sample households. Ideally, we want to collect responses for both questions on race from all individuals in our selected

households. Nonresponse to the questions and other factors that depend on the design option used can introduce nonsampling error biases. In addition, we want to collect sufficient data for each of the reported race combinations of interest to insure that the estimates have sufficient reliability to support the research objectives.

Proposed Design Options

Several design options have been proposed for collecting information for the questions on race that instruct respondents to "mark one or more" or "mark one" race category. A brief description of each option follows.

Option 1: New Collection with Both Measurements in the Same Instrument

In a new data collection that replicates Census 2000 data collection methods, collect both measurements in the SAME questionnaire. A questionnaire with both a "mark one or more" and a "mark one" instruction to the question on race would be mailed out or dropped off at the sample addresses. For households not returning the questionnaire by mail, a personal visit interview would be conducted using a questionnaire containing both questions on race. A telephone item nonresponse follow-up operation might be needed for questionnaires returned by mail with a missing or invalid response to the "mark one" question on race.

Option 2: Match Census 2000 to One or More Current Surveys

Collect both measurements independently in the same general time frame and link responses by matching households and individuals. Currently, the Census Bureau's demographic surveys use a "mark one" instruction for the question on race. Specifically, this option consists of matching Census 2000 individuals to individuals interviewed in Census Bureau surveys conducted during February through May of 2000, the same general time frame as the Census 2000 data collection. (Note, even if this option is not selected for this evaluation, the match will be done for other evaluation purposes.)

Option 3: One New Collection with New Follow-up Interview Collection

ATTACHMENT 3
(Date written: 7/27/2000)

Methodological Criteria and Proposed Sample Design Options

In two separate stages collect new data using both questions on race. The first stage is a new data collection that replicates census data collection methods (i.e, mailout/drop-off with an in-person nonresponse follow-up operation). The second stage is a new data collection in which a follow-up interview is conducted (by telephone or in-person otherwise) with the household respondent in the first stage. The second stage is conducted a few weeks after the first stage is completed. (Issues of the sequence of the measurements and whether first stage responses are provided to the respondent at the second stage are being discussed.)

Option 4: Match Census 2000 Responses to a "Mark One" Measurement

In two separate stages, collect both measurements. The first stage, the "mark one or more" measurement, is the actual Census 2000 data. The second stage is a new data collection that replicates census data collection methods and uses a "mark one" question on race. The second stage collection requests the Census 2000 household respondents to respond to the second stage data collection. The individuals in the second stage data collection are matched to Census 2000 data records and their responses to the two questions on race are compared.

Option 5: Match Census 2000 to a Personal Visit "Mark One" Measurement

Identical to Option 4, except the second stage measurement ("mark one") is collected via a personal visit interview that is designed to include additional questions to maximize the potential for matching individuals to Census 2000 data records. Key design features include providing the interviewers with the names of the individuals enumerated at the sample address in Census 2000 and provision to collect "census day residence" information for whole household movers. Also, an attempt is made to identify and collect the data from the Census 2000 household respondent.

Option 6: Two New Independent Collections for Each Measurement

Select an independent sample to collect each measurement. Both the "mark one or more" measurement and the "mark one" measurement would replicate the census data collection methods. Unlike the other options, this option does not attempt to create a data file that contains both measurements for the same individuals, rather it would attempt to use race distributional differences between the two measurements to meet the research objective.

ATTACHMENT 3
(Date written: 7/27/2000)

Methodological Criteria and Proposed Sample Design Options

Discussion of the Options and Additional Issues

The proposed options meet the six criteria to different degrees of success. Based on the discussions at the June 26th meeting and subsequent Census Bureau discussions the most viable options have been identified. The other options were deemed to have at least one critical failure in meeting the criteria.

"Viable" Options: 3 and 4

"Failed" Options: 1, 2, 5, and 6

Option 1: It is believed that substantial, but unmeasurable, interactions will take place between the collected data for both measurements with both race questions in the same self-response instrument.

Option 2: Even though we will still perform this match for other evaluations, it will not provide a sufficient number of households with persons reporting more than one race to produce reliable estimates for our purposes.

Option 5: High cost per case associated with conducting all interviews in person would reduce the effective sample size. Effect of all data being collected using in-person interviewing on the "mark one" measurement is also of concern.

Option 6: Estimating the parameters of interest may not be possible. High costs associated with larger sample sizes needed to achieve comparable reliability.

Though everything above is still open for discussion, including any additional thoughts or suggestions of alternative options, we have specific issues to address within Options 3 and 4 to work out before we can make a decision on the "best" data collection method option.

Issues within Option 3

- Ordering of the "mark one or more" and "mark one" measurements

  Should the first stage be the "mark one" measurement with the followup being the "mark one or more" measurement or vice versa?

- Dependent or independent interview in second stage follwup

  Should the household respondent at the second stage be reminded of their first stage response to the race question?

ATTACHMENT 3
(Date written: 7/27/2000)

Methodological Criteria and Proposed Sample Design Options

Points to keep in mind:

- Desire to have "mark one or more" measurement race distribution be consistent with Census 2000 race distribution.
- Desire to have "mark one" measurement uninfluenced by "mark one or more" measurement.
- Desire to reduce impact of response variance on the utility of the data file.

Issues within Option 4

- Effect of whole household movers on our ability to match to Census 2000 data records and the consequences for the "representative" criteria.

- Effect of different household respondents and our inability to determine if they are different.

- Modifications to the data collection instruments (mailback/enumerator) needed to maximize the match rate.

Table 15. Census 2000 distribution of race.

Race	Non-Hispanic Percent	Hispanic Percent
White	79.0 %	47.9 %
Black	13.8	2.0
American Indian & Alaska Native	0.8	1.2
Asian	4.1	0.3
Native Hawaiian & Other Pacific Islander	0.1	0.1
Some Other Race	0.2	42.2
Two or More Races	1.9	6.3

* The data in Table 15 represent the total U.S. population in Census 2000, which includes people living in households, as well as in group quarters. Note that the CQS universe is slightly different, as it was restricted to the household population. In Census 2000, 2.8 percent of all individuals enumerated lived in group quarters.

ATTACHMENT 5

Hispanic Data Tabulations

Table 16. Consistency in reporting Two or more races for Hispanics for Panel A.*

Census 2000 Race	CQS Initial Contact ("Mark one or more races")		
	Single race	Two or more races	TOTAL
Single race	8,459,619 n= 4,322	543,615 n= 516	9,003,234 n= 4,838
Two or more races	392,662 n= 2,348	112,703 n= 1,126	505,365 n= 3,474
TOTAL	8,852,281 n= 6,670	656,318 n= 1,642	9,508,599 n= 8,312

* The data in Table 16 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the weighted data were obtained using the inverse of the original sampling probabilities with no adjustment (Z_WGT1).

Table 17. Detailed consistency in reporting of race for Hispanics for Panel A.*

Census 2000 Race	**CQS Initial Contact ("Mark one or more races")**							
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more	TOTAL
White	**2,679,570**	652	1,083	1,039	185	2,470,120	297,402	5,450,051
Black	555	**32,742**	157	0	0	10,312	116,572	160,338
AIAN	1,660	43	**164,332**	76	198	3,915	2,193	172,417
Asian	598	36	167	**24,105**	203	39,579	2,599	67,287
NHOPI	631	5	0	41	**1,275**	341	995	3,288
SOR	845,374	37,415	42,414	565	812	**2,099,419**	123,854	3,149,853
Two or more	219,221	12,262	8,049	7,829	883	144,418	**112,703**	505,365
TOTAL	3,747,609	83,155	216,202	33,655	3,556	4,768,104	656,318	9,508,599

* The data in Table 17 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the weighted data were obtained using the inverse of the original sampling probabilities with no adjustment (Z_WGT1).

ATTACHMENT 5

Hispanic Data Tabulations

Table 18. Consistency in reporting Two or more races for Hispanics for Panel B.*

Census 2000 Race	CQS Re-contact ("Choose one or more races")		
	Single race	Two or more races	TOTAL
Single race	9,485,906 n= 4,296	555,757 n= 420	10,041,663 n= 4,716
Two or more races	336,453 n= 2,285	91,160 n= 802	427,613 n= 3,087
TOTAL	9,822,359 n= 6,581	646,917 n= 1,222	10,469,276 n= 7,803

* The data in Table 18 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the weighted data were obtained using the inverse of the original sampling probabilities with no adjustment (Z_WGT1).

Table 19. Detailed consistency in reporting of race for Hispanics for Panel B.*

Census 2000 Race	**CQS Re-contact ("Choose one or more races")**							
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more	TOTAL
White	**2,844,747**	5,322	1230	430	43,604	1,451,297	233,311	4,579,941
Black	4,043	**203,652**	0	18	182	102,554	3,186	313,635
AIAN	15,208	167	**3,217**	0	52	174,330	44,531	237,505
Asian	1,712	0	92	**4,598**	0	38,957	2,136	47,495
NHOPI	430	0	0	3006	**3,473**	202	138,329	145,440
SOR	1,704,146	44,513	1,479	170	34,524	**2,798,551**	134,264	4,717,647
Two or more	171,677	14,271	9,562	5,318	3,232	132,393	**91,160**	427,613
TOTAL	4,741,963	267,925	15,580	13,540	85,067	4,698,284	646,917	10,469,276

* The data in Table 19 were restricted to matched people who did not have an imputed race in Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the weighted data were obtained using the inverse of the original sampling probabilities with no adjustment (Z_WGT1).

ATTACHMENT 5

Hispanic Data Tabulations

Table 20. Tabulations for individuals reporting Two or more races in Panel A (for Hispanics only).*

CQS Initial Contact ("Mark one or more races")	CQS Re-contact ("Choose one race") with followup probe									
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more (same)	Two or more (different)	Missing	TOTAL
White - Black	**1,007**	**2,145**	0	0	0	2,029	1,165	291	1,529	8,165
White - AIAN	**5,560**	0	**2,927**	0	0	1,802	825	729	2,150	13,991
White - Asian	**5,160**	171	0	**2,504**	0	1,446	1,743	691	3,098	14,813
White - NHOPI	**628**	0	0	0	**615**	97	92	112	125	1,669
White - SOR	**29,601**	307	55	18	0	**21,425**	1,372	464	12,091	65,333
Black - AIAN	0	**745**	**76**	0	0	447	0	30	404	1,702
Black - Asian	0	**742**	6	**0**	0	56	366	19	0	1,191
Black - NHOPI	0	**0**	0	0	**0**	0	0	0	38	38
Black - SOR	144	**5,277**	405	0	0	**2,826**	1,246	665	2,184	12,747
AIAN - Asian	111	0	**640**	**110**	0	401	0	19	58	1,339
AIAN - NHOPI	0	0	**0**	0	**595**	0	0	0	0	595
AIAN - SOR	543	0	**974**	0	0	**3,125**	198	126	553	5,519
Asian - NHOPI	92	0	0	**438**	**1,053**	343	230	113	282	2,551
Asian - SOR	1,366	0	0	**1,458**	0	**2,586**	1,758	411	217	7,795
NHOPI - SOR	64	0	0	0	**687**	**159**	0	20	38	968
Three or more	3,264	1,964	185	951	4,321	6,402	839	1,541	2,510	21,976
TOTAL	47,540	11,352	5,269	5,479	7,269	43,142	9,833	5,231	25,276	160,391

* The data in Table 20 were restricted to people who were identified as Two or more races in the initial contact and where the races were not imputed for those matched to Census 2000 - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the data were restricted to cases in sampling strata 2 and 3 (Z_WGT2).

ATTACHMENT 5

Hispanic Data Tabulations

Table 21. Tabulations for individuals reporting Two or more races in Panel B (for Hispanics only).*

Census 2000 Race ("Mark one or more races")	CQS Initial Contact ("Mark one races")									
	White	Black	AIAN	Asian	NHOPI	SOR	Two or more (same)	Two or more (different)	Missing	TOTAL
White - Black	**3,212**	**2,745**	181	0	0	4,450	3,015	1,560	1,093	16,255
White - AIAN	**10,194**	0	**3,196**	0	0	5,521	2,848	1,326	924	24,009
White - Asian	**3,584**	0	194	**1,880**	172	2,249	3,494	1,114	198	12,885
White - NHOPI	**766**	0	0	0	**245**	507	102	165	11	1,796
White - SOR	**143,138**	125	416	53	20	**120,320**	21,447	1,314	11,913	298,746
Black - AIAN	226	**770**	**201**	0	0	744	728	593	12	3,274
Black - Asian	0	**148**	0	**0**	0	41	88	251	0	528
Black - NHOPI	0	**482**	0	0	**0**	114	0	0	0	596
Black - SOR	3,071	**9,907**	0	18	0	**7,525**	2,593	722	3,364	27,199
AIAN - Asian	65	0	**0**	**295**	0	196	258	414	0	1,228
AIAN - NHOPI	97	0	**0**	0	**41**	0	0	0	0	138
AIAN - SOR	2,195	55	**1,924**	0	0	**7,160**	88	392	301	12,115
Asian - NHOPI	92	0	0	**822**	**372**	333	629	580	0	2,828
Asian - SOR	1,659	0	0	**6,642**	0	**1,308**	1,385	1,301	528	12,827
NHOPI - SOR	244	0	120	0	**0**	**1,767**	0	405	312	2,849
Three or more	3,269	1,683	489	818	1,215	5,902	2,801	4,835	576	21,588
TOTAL	171,811	15,916	6,721	10,528	2,066	158,135	39,478	14,974	19,232	438,862

* The data in Table 21 were restricted to matched people who were identified as Two or more races in Census 2000 and where the races were not imputed - that is, only those cases where the final edited race was "as reported," or where the code was changed "through consistency edit." The CQS initial contact Hispanic origin response was used. Additionally, the data were restricted to cases in sampling strata 2 and 3 (Z_WGT2).