Population Division

Identification of Hispanic Ethnicity in Census 2000: Analysis of Data Quality for the Question on Hispanic Origin

by

Arthur R. Cresce, Audrey Dianne Schmidley
and Roberto R. Ramirez

Working Paper No. 75

U S C E N S U S B U R E A U
Helping You Make Informed Decisions

Acknowledgements

Table of Contents

Identification of Hispanic Ethnicity in Census 2000:

Analysis of Data Quality for the Question on Hispanic Origin

1. Introduction

Following a custom observed since the Census Bureau first began collecting Hispanic data in the 1970 census, this paper will report on the quality of information collected in Census 2000.[1] As did the authors of previous papers, we will look at the quality of Hispanic data from a variety of viewpoints, and draw some overall conclusions and implications for future research.[2]

We will begin by reviewing: selected studies that evaluated the quality of 1990 census data; findings and conclusions agency analysts derived from special surveys that tested census data collection methods; and subsequent efforts by the agency to address issues uncovered during the review process and in preparation for Census 2000. Next, we will summarize the results of Census 2000 evaluative studies and discuss what we know about Hispanic data quality based on this recent research. Finally, we will provide an overall assessment of Hispanic question results and describe how we are addressing data quality issues for Census 2010 and the American Community Survey.

2. Planning for Census 2000

Public Law 94-311[3] directs the U.S. Census Bureau to collect information about individuals of Spanish origin and descent. A major source of race and ethnic data is the decennial census. Census data are used for policy purposes, and to organize, monitor, and evaluate various federal programs mandated by law such as the Voting Rights Act. State and local governments as well as businesses also use race and ethnic data for administrative and marketing purposes.

While the majority of tasks associated with census data development such as collecting, processing, disseminating, and evaluating data are carried out by the Census Bureau, the Office of Management and Budget (OMB) and the General Accounting Office (GAO), have oversight responsibilities that involve establishing data policy standards and/or

---

[1] The Census Bureau collected ethnic data before 1970 for selected groups. See Gibson and Jung (2002).

[2] See McKenney et al. (1985); Fernandez and Cresce (1986); McKenney et al. (1988); McKenney and Cresce (1992); Cresce et al. (1992); Cresce (2002); and del Pinal (2003).

[3] See also 29 U.S. Code Section 8.

evaluating census data outcomes. The Census Bureau also receives planning advice and dissemination support from external analysts in academic and federal, state, and local government settings.

Following each decennial census, the Census Bureau has evaluated the quality of the information collected in the previous census. This evaluation is important in its own right, but it also serves as a first step in planning for the next census. During the period following the 1990 census, analysts reviewed and evaluated data results, and fielded surveys designed to address shortcomings uncovered in the post-census evaluation.

In this section, we focus on the evaluation and review of 1990 Hispanic data and steps taken in preparation for Census 2000. We begin our discussion by noting the "final" changes implemented in the Census 2000 process. Next, we summarize actions taken by OMB in the 1990s based on the results of Census Bureau research and public feedback. We conclude the section with an overview of the individual evaluation and research steps undertaken between 1990 and 2000.

2.1. Questionnaire changes between the 1990 census and Census 2000.

In 1970, the Census Bureau began collecting decennial census information about the Hispanic population. Following experimentation with various questions, the agency concluded that a single self-identification question would produce the most accurate and reliable results.[4] Since 1980, the Census Bureau has used a single core question on the decennial form, but for a variety of reasons, over the past three censuses (1980, 1990, and 2000) the agency has modified the Hispanic question as well as the census questionnaire. (See Appendix A for examples of 1980, 1990, and 2000 questions).

Following the 1990 census, the Census Bureau ascertained that a) Hispanic item non-response; b) misreporting of 'Mexican-American' by non-Hispanics; and c) the collection of reliable Hispanic group detail continued to present a challenge. After research and consultation, the agency completed the design of the Census 2000 questionnaire. The final Census 2000 instrument differed from that used in the 1990 census in the following ways:

> • In 1990, the census questionnaire used a matrix format with people in the columns and questions on the rows. In contrast, the Census 2000 form used individual person blocks of space.

---

[4] The Hispanic population has also been identified as Spanish and Latino by its constituents. See del Pinal (1994).

• In 1990, the race question preceded the Hispanic question. In 2000, the Hispanic question came first.

• The Hispanic question instruction wording changed. The 2000 question added the term "Latino", dropped the word "origin" from "Hispanic origin", and changed the general instruction to "Mark the 'No' box if not Spanish/Hispanic/Latino."

• In 1990, the residual (write-in) response category, "Other Spanish/Hispanic" response category was accompanied by examples. This was not so in Census 2000.

2.2. Revisions to Statistical Policy Directive 15.

In July 1993, the Office of Management and Budget (OMB) which oversees the development of federal forms used to collect information from individuals and organizations in the United States, initiated the process of considering revisions to Statistical Policy Directive No. 15 – the federal government's written policy for classifying racial and ethnic groups.[5] Key census-related issues OMB addressed during the 1990 period included re-sequencing the race and Hispanic questions on the decennial census form and the use of the term 'Latino' in the Hispanic question.

In the fall of 1997, referring to research conducted by the Census Bureau following the 1990 Census (see below) as well as comments from stakeholders responding to a Federal Register Notice, OMB updated Directive No. 15. The revised directive required the Census Bureau to place the Hispanic question before the race question on the Census 2000 form, and to change the term "Hispanic" to "Hispanic or Latino" in the Hispanic question.

2.3. Evaluation and review of the 1990 census results.

Evaluative research conducted by the Census Bureau following the 1990 census detected several areas of concern with regard to the Hispanic question results: a) higher than expected allocation rates; b) misreporting in the "Mexican, Mexican Am., Chicano" and "Other Spanish/Hispanic" categories; and c) evidence of response problems with the "Yes, Other Spanish/Hispanic" category. Analysts uncovered these issues by comparing: a) 1980 and 1990 race and ethnic distributions; b) 1980 and 1990 imputation results; and c) 1990 results with results from the Content Reinterview Survey.[6]

---

[5] See http://www.whitehouse.gov/omb/fedreg/ombdir15.html

[6] See McKenney et al (1993) and Cresce et al (1992).

Comparing the 1980 and 1990 census results demonstrated that the 1990 short form produced a relatively higher computer allocation rate for the Hispanic item (10.0 percent) compared with the 1980 Hispanic short form (4.2 percent). On the other hand, the 1990 long form rate was 3.5 percent, much closer to that of 1980 long form rate (2.3 percent) *even though both the long and short forms used the same question.* The Census Bureau determined that field follow-up was the major factor in lowering the allocation rate in these instances.

Although an overall reduction in misreporting in the Mexican category occurred between the 1980 and 1990 censuses, Census Bureau analysts noted that an inordinate increase in misreporting had occurred in several states. One study observed that the 1990 edit and imputation procedures had increased the overlap between various racial groups and the Hispanic population which probably led to these unexpected results.[7] Specifically, this analysis suggested about 62 percent of the Black Mexicans reflected in the census results were created by the edit and imputation procedures.

Analysts also indicated that the "Other Spanish/Hispanic" population increased significantly between 1980 and 1990.[8] Although some of the change could be explained by population growth due to natural increase, the history of inconsistent reporting through content reinterview surveys in this category suggested further evaluation was necessary.

2.4. Preparing for Census 2000.
2.4.1 National Content Survey (NCS).
Following the decision to eliminate content-edit followup in Census 2000, the Census Bureau began to search for new ways to reduce non-response and improve the edit and imputation procedures used to assign missing information. During the 1990s, the agency fielded a number of test surveys designed to meet these goals. The results from the surveys provided a basis for evaluating the effects of various questions and questionnaire formats, as well as field operations procedures disclosed when the Census 2000 Operational Plan was unveiled in 1996.[9]

In April 1996, the Census Bureau conducted the National Content Survey (NCS) in

---

[7] Del Pinal (1994)

[8] Cresce (1992)

[9] The Operational Plan included a simplified questionnaire that eliminated and/or shortened instructions and/or removed examples including those used in the 1990 Hispanic question.

preparation for Census 2000. The primary purpose of the 1996 NCS was to test the feasibility of having respondents report more than one race - either by the use of a single "multiracial" category or a "mark all that apply" approach. Another major objective of the NCS was to test the effects of sequencing the Hispanic question before the race question.

Two versions of the Hispanic question were used in the survey questionnaire. One version was based on the 1990-style Hispanic question ("Is this person of Spanish/Hispanic origin?") with examples for "Other Spanish/Hispanic". The alternative "Census 2000" version ("Is this person Spanish/Hispanic/Latino") excluded these examples.

Evidence from the NCS supported the notion that the position of the Hispanic question on the Census 2000 questionnaire would be a major factor in narrowing the Hispanic item nonresponse rate. NCS results also supported the notion that the term 'Latino' was salient within the Hispanic population.

2.4.2. Race and Ethnic Targeted Test (RAETT).

In June 1996, the Census Bureau conducted the Race and Ethnic Targeted Test (RAETT). RAETT was designed to: a) evaluate further methods for allowing respondents to report more than one race, b) evaluate further the effect of alternative sequencing of race and Hispanic questions, and c) examine the feasibility of using a combined race and Hispanic question. RAETT obtained responses from the members of relatively small racial groups that might not have sufficient representation in a standard national probability sample (such as the 1996 NCS).

The NCS results, published by the Census Bureau in December of 1996, had indicated that the placement of the Hispanic question before the race question reduced nonresponse to the Hispanic question and did not increase nonresponse to the race question. However, the percent Hispanic based on the Census 2000-style question results appeared to differ from the results based on the 1990-style question, regardless of whether or not the Hispanic question appeared before or after the race question.

Although there were no other statistical differences owing to sample size, the proportion of the 'Other Hispanic" sample in the Census 2000-style panels was larger regardless of whether or not the Hispanic question appeared before or after the race question. The reason for the apparent discrepancy was not clear (i.e., did adding "Latino," dropping "origin," or eliminating examples have an effect). Agency analysts peer reviewing the results recommended further investigation using multivariate analysis to see if the interaction between using a multiracial category and sequencing of the questions might

explain the observed effects, but no further research was conducted because the effects could not be decomposed using the available sample.

Regarding the testing of a combined race/ethnic question, the purpose of the RAETT was to determine the effects of collecting information about race, Hispanic origin, and ancestry in a combined two-part question. The RAETT tested two versions of a combined question (Figure 4). Both provided response boxes that conformed to the extant OMB race groups and for "Some other race." Both also included a write-in line for American Indian or Alaska Native tribe affiliation. One version (Panel E) included a "multiracial" category and requested that only one checkbox be marked, while the other (Panel F) did not include a "multiracial" category but included an instruction to "mark one or more" checkboxes. Both were followed by Part B of the question, which asked respondents to report their "ancestry or ethnic group" on write-in lines provided for this purpose. One objective of the ancestry write-in was to determine how detailed Asian or Pacific Islander and Hispanic groups would be reported.

Nonresponse to each combined question was significantly lower than nonresponse to corresponding separate Hispanic origin and race questions (included in the RAETT for comparison) in the other test panels. Furthermore, both of the combined questions elicited high levels of multiple response in the Hispanic targeted sample (i.e. both race and Hispanic origin were provided). And, in comparison with the separate questions, when all the 'Hispanic' responses to the combined questions (either alone or in combination with any other response (race or ethnic) were added together (which is the proper comparison) there was no statistical difference in the percent reporting Hispanic origin.[10]

In about 11-13 percent of the responses to the RAETT combined questions, the respondent selected 'Hispanic' but did not include a response in the ancestry component of the question. Because of this relatively high rate of nonresponse to the ancestry component, the percent amounts of specific Hispanic origin groups obtained from the combined questions were consistently lower than those obtained from the separate Hispanic origin questions.[11]

In summary, RAETT determined that a combined race and Hispanic origin question produced lower nonresponse rates than did separate questions. Reporting rates in the

---

[10] U.S. Census Bureau, Population Division Working Paper No. 18. "Results of the 1996 Race and Ethnic Targeted Test" pages I-18 to I-19.

[11] U.S. Census Bureau, Population Division Working Paper No. 18. "Results of the 1996 Race and Ethnic Targeted Test" pages I-21.

'alone' categories of Hispanic origin for the combined question appeared to be lower than comparable results from the separate questions. However, when Hispanic responses in combination with the race responses were included, the percent Hispanic for the combined panels was not statistically different from the percent using a separate question. Finally, information about detailed Hispanic groups appears to have been reduced by the combined question approach.

In May 1997, the Census Bureau published the 1996 RAETT results. Using the Census 2000-style Hispanic question, RAETT showed that asking the Hispanic question before the race question reduced item non-response. The combined race and Hispanic question produced conflicting results, leading to a decision by the agency not to pursue this approach for Census 2000. On one hand, nonresponse was lower in the combined question than for the separate race and Hispanic questions. In addition, the mixed question generated results for the major race groups and the total Hispanic population that were comparable to those derived from separate race and Hispanic questions.[12] On the other hand, the combined question could not produce totals for detailed Hispanic groups that were comparable to those from a separate Hispanic question.

Following the decision by OMB to revise Directive 15, in April 1998 the Census Bureau submitted the list of specific questions to be included in Census 2000. The Hispanic question was to be placed before the race question and be the essentially the same as that used in the 1996 RAETT (with no examples for "Other Spanish/Hispanic/Latino" and the Hispanic question would include the term "Latino" but exclude the term "origin").

3. Results from Census 2000

Census 2000 results show the Hispanic population grew from 22.4 million to 35.3 million between 1990 and 2000, or 57.9 percent (Table 1). This change outpaced that observed between 1980 and 1990 when the Hispanic population increased by 53.4 percent on a smaller base.[13]

Census 2000 data also reveal that during the 1990s, the Hispanic population dispersed beyond traditional population centers. Although States with generally large Hispanic populations, such as California and Texas experienced substantial growth (42.6 percent and 53.7 percent, respectively), other States such as North Carolina and Virginia also experienced significant growth (393.9 percent and 105.6 percent, respectively).

---

[12] OMB classification. See Directive No. 15.

[13] Hobbs and Stoops (2002).

Because Census 2000 represented the first time that write-in responses for detailed Hispanic groups were coded on a 100-percent basis (write-in responses had been coded on a sample basis in 1990), information for detailed Hispanic groups such as Dominican and Salvadoran were available at the block level. *Sample data* from both censuses, indicate that the population of most Hispanic groups grew between 1990 and 2000 (Table 2.). However, the extraordinary increase among certain groups such as "Hispanic" and "Latino" and the relative proportional decrease of certain groups such as "Nicaraguan" and "Spaniard" raises questions about the quality of the Hispanic data. This issue is pursued in more detail in the next section.

4. What do we know about data quality for the Hispanic question from Census 2000?

4.1 Is the change in the Hispanic population reasonable?

While Census 2000 data revealed that the Hispanic population had increased substantially between 1990 and 2000 (57.9 percent), the total population grew by only 13.2 percent.[14] This dramatic growth plus the disparity between the Census 2000 figure (35.3 million) and the demographic estimate for July 1, 2000 (32.2 million) prompted the Census Bureau to reexamine assumptions about international migration used in the development of population estimates.[15]

Shortly after the Census was completed, new international migration (legal and unauthorized) assumptions were developed for the Hispanic population based on a series of demographic research reports.[16] Although the number of Hispanics counted in Census 2000 was higher than expected, the growth of the Hispanic population was deemed reasonable after the results of the demographic review showed that the Census Bureau had originally underestimated Hispanic international migration during 1990s. Hispanic population adjustments were made to reflect these new findings (e.g., 35.6 million = July 1, 2000 estimate).[17]

4.2. Improved response rates

---

[14] Guzman (2001)

[15] Robinson (2001)

[16] Cresce, et al (2001). This paper is part of the Demographic Analysis of Population Estimates or DAPE series.

[17] Consistent with 2000 estimates base. See Table NA-EST2002-ASRO-02, "National Population Estimates - Characteristics," Population Division, U.S. Census Bureau, Washington, D.C. 20233

Sequencing of the Hispanic question before the race question may have finally resolved the problem of how to improve the response level of the Hispanic question without the use of a field content-edit followup.[18] Compared with results from the 1990 census, the Census 2000 total allocation rate fell from10.4 percent to 5.6 percent (Table 3). In addition, the Hispanic allocation rates declined in every state except Alaska, New Mexico, Vermont, and Wyoming. The total number of imputations for the Hispanic Origin question also dropped during the period from 25.5 million in 1990 to 16.8 million in 2000.

4.3. Improved imputation methodology

During Census 2000, in addition to improving the response rate and thereby reducing the number of required imputations, the Census Bureau improved the imputation process by:

• *Reducing the proportion of imputed "hot deck" cases.* In 1990, 75.6 percent of census data allocations were processed using the "hot deck "method. In Census 2000, only 41.2 percent of the imputations were processed in this manner.[19]

• *Introducing the use of Spanish and non-Spanish surnames into the hot deck procedure.* When the hot deck procedure was used, donors with Spanish surnames were used to impute values to Spanish surnamed cases with missing values (and vice versa for non-Hispanic cases). Approximately 31.4 percent of all imputations and 8.1 percent of all Hispanic imputation cases were affected by this enhancement of the hot deck procedure.

• *Combining the race and Hispanic imputation procedures.* In 1990, the race and Hispanic origin edit and imputation procedures were executed independently. This approach apparently contributed to the propagation of relatively rare race Hispanic combinations (for example, Black Mexicans), although in some part these esoteric combinations may also have been the result of misreporting. In Census 2000, the race and Hispanic origin edit specifications were integrated and the rules for 'within household' and 'hot deck procedures' restricted so that Hispanic donors and donees were matched before race was assigned.

---

[18] Cresce and Ramirez (2001)

[19] Census Bureau analysts consider the hot deck method the least reliable imputation method. Other methods impute values based on additional information about the individual in question (i.e. birthplace might be used to infer Hispanic Origin) or a member of his household (i.e. Hispanic origin of the household head might be used to assign Hispanic status to dependent children). Hot deck allocation involves the assignment of values from a set of stored values collected from other households. The phrase "hot deck" is used to describe this source because the deck is constantly refreshed by newly processed cases.

4.4. Good overall response consistency as measured by reinterview and by comparison with the 1990 version of the Hispanic question

Since 2000, the Census Bureau has conducted several studies evaluating the quality of Census 2000 data including the Census 2000 Content Reinterview Survey (CRS) and the Alternate Questionnaire Experiment (AQE).[20] Major findings from these studies are discussed below.[21]

Shortly after the last decennial census, the Census Bureau randomly selected a sample of 30,000 households that had received the long form in 2000. One person from each household in this sample was telephone interviewed by an experienced field representative. The primary goal of the survey was to evaluate the quality of the census data by comparing responses provided through a phone interview with those reported in the census questionnaire. Using data from this Content Reinterview Survey, analysts developed an index of inconsistency in reporting.

Results from the CRS indicate edited data for the Hispanic origin question displayed mixed results. The consistency of "Not Hispanic," "Mexican," "Puerto Rican," and "Cuban" responses fell in the good range (less than 20). "Other Hispanic" scored in the moderate range (20 to 50). "Multiple non-Hispanic," "Multiple Hispanic," and "Mixed Origin" scored in the poor range (over 50) (Table 4).[22]

The Alternate Questionnaire Experiment was designed to measure the total effect of the changes in the Census mail questionnaire from 1990 to 2000 by comparing two independent random samples of households. About 10,500 households received the1990-style short form while about 25,000 households received the census 2000 short form. The 1990-style form retained the same 1990 question wording, categories, order and format, but incorporated some recognizable elements of the Census 2000 design. Because this experiment was conducted by mail, the results of the study were generalizable only to the Census 2000 mailout-mailback universe.

According to the results of the AQE, changes to the Census 2000 questionnaire led to improved reporting of Hispanic origin as measured by item nonresponse.   For example, the overall item nonresponse to the question of Hispanic origin was 3.3 percent in the Census 2000-style questionnaire, compared with 14.5 percent in the 1990-style

---

[20] Singer and Ennis (2002); Martin (2002)

[21] Del Pinal (2003) provides a more in-depth summary.

[22] Del Pinal (2003). Table 3.2. "Response Variance Measures for Hispanic Origin (Edited Data)."

questionnaire. Nonresponse to the race question by Hispanics was also reduced by nearly 10 percentage points, from 30.5 percent in the 1990-style questionnaire to 20.8 percent in the 2000-style questionnaire.

4.5. Weaknesses in the Census 2000 Hispanic data

4.6. Less than expected growth for specific Hispanic groups; Substantial growth in reporting of "generic" Hispanic terms; Evidence that question wording and format led respondents to report more general responses instead of more specific responses

Despite many positive findings, the Alternative Questionnaire Experiment (discussed above) revealed the census mail questionnaire probably produced a few unwanted results. There was no difference between the two groups (those receiving the 1990-style questionnaire and those receiving the 2000-style questionnaire) in the percent of people reporting Hispanic (about 11.1 percent of each group surveyed). However, members of the group receiving the 2000-style questionnaires were less likely to report a specific Hispanic group (e.g., Mexican, Cuban, Puerto Rican) and more likely to report a general Hispanic term (e.g., Hispanic, Latino, Spanish) compared with the sample that received the 1990-style questionnaires. Specifically, the AQE found that 92 percent of the Hispanics who responded to the 1990-style form provided a specific Hispanic group identity compared with 80 percent of those who responded to a 2000-style form. Thus, the 2000-style forms produced about 10 percent more general responses than the 1990-style forms. AQE results suggest this difference is probably due to the combined effects of changes in the question wording (e.g., removal of the word "origin" which appeared on the 1990 form, and addition of the term "Latino" to the 2000 form) and/or the elimination of specific Hispanic origin examples from the Census 2000 questionnaire.

Findings from Census 2000 compared with those from the 1990 census show a similar pattern (Table 2). In Census 2000, the proportion of the Hispanic population providing a specific origin was 83.9 percent compared with 93.6 percent in the 1990 census.[23] In addition, the proportions of persons responding differed across groups. The Mexican origin population declined from 61.2 percent in 1990 to 59.3 in 2000; Puerto Rican origin declined from 12.1 percent to 9.7 percent; Cuban origin declined from 4.8 percent to 3.5 percent. On the other hand, general responses all experienced dramatic increases. For example, the percent "Latino" increased from less than 0.1 percent to 1.2 percent, "Hispanic" increased from 1.8 percent to 6.6 percent, "Spanish" increased slightly from

---

[23] N.B. These results differ from those in the preceding paragraph because they are actual 1990 and 2000 census results, whereas those in the preceding paragraph reflect the results from the Alternate Question Experiment.

2.0 percent to 2.2 percent and finally "Other Hispanic" increased from 2.6 percent to 5.8 percent.

Two independent external studies by Roberto Suro of the Pew Hispanic Center and John Logan of the Lewis Mumford Center raised additional concerns about the accuracy of the detailed Hispanic group data.[24] Although both analysts praised the Census Bureau for producing a good total count of Hispanics, Suro and Logan both provide additional evidence that the Census 2000 Hispanic question may have significantly underestimated the size of some Hispanic groups in the United States.

In 2002, Government Accounting Office auditors met with Census Bureau staff to discuss the decision that led to the selection of the version of the Hispanic question used in Census 2000 as well as the results from Census 2000. The GAO noted "the Bureau's lack of agency wide guidelines for its decisions on the level of quality needed to release data to the public" when the agency realized that question wording and format may have adversely affected reporting of detailed Hispanic groups.[25]

At the request of members of Congress and the Latino community, the Census Bureau further analyzed the Census 2000 Hispanic data in an effort to ascertain what kinds of detailed responses individuals might have provided in lieu of the more general responses they did provide. Using a 'what if' scenario, Census Bureau staff devised a simulation model that generated detailed Hispanic values based on information derived from responses to other census questions such as nativity and ancestry.[26] For example, if a respondent reported "Latino" in the Hispanic origin question and indicated he was born in Mexico, he was coded "Mexican" (a detailed response) in the simulation model.

When the criteria for refining the Hispanic detail were applied the numbers and proportions for many of the detailed groups increased. For example, the category Spaniard increased by about 69 percent (Census 2000 results = 112,999; Simulation model results = 190,656). In fact, all the detailed Hispanic group proportions increased at least 24 percent except Mexican (7 percent), Puerto Rican (4 percent), and Cuban (5 percent). The biggest numerical gainer was the Mexican category (Census 2000 results = 20.9 million; Simulation results = 22.3 million, or a 1.4 million gain). The Mexican cases accounted for nearly half (47 percent) of all the sample reassigned from a general to a specific response (1.4 million out of 3.1 million) in the simulation model.

---

[24] Suro (2002) compared Census 2000 results with Census 2000 Special Survey results. Logan (2002) compared Census 2000 results with information from the 1998-2000 Current Population Survey.

[25] GAO (2003)

[26] Cresce and Ramirez (2003).

Because long form information derived from the place of birth and ancestry question are only available for long form census cases, the simulation study findings may be of limited use in refining decennial census data.[27] However, this methodology opens up interesting possibilities with regard to the American Community Survey conducted annually.

Another group of Census Bureau analysts compared the Census 2000 Hispanic data results with data reported in the Census 2000 Supplemental Survey (C2SS).[28] These researchers found that both data sources provided similar totals for the Hispanic population. On the other hand, they noted Census 2000 produced lower detailed Hispanic group rates and higher general group rates. The report suggests:

> "the observed differences are due to the use of examples in the C2SS during telephone and personal visit interviewing. These aids were not provided during Census 2000 operations, although one could argue that the presence of the Hispanic origin checkbox groups act as examples. This reasoning does not explain why the Mexican percentage is also lower in Census 2000."[29]

Although the universe for these two data collection systems are somewhat different (Census 2000 represent the total population, i.e. the population in both households and group quarters; C2SS represents the household population alone), the results of this comparison add to the evidence that the Census 2000 question may have influenced respondents to report general Hispanic answers.

4.7. Evidence of slight decline in response consistency as measured by reinterview

Finally, even though response consistency was in the good range for the Hispanic question, response consistency declined between 1990 to 2000 not only for the Hispanic question overall, but also for the Mexican and Puerto Rican origin categories (Table 4.). These differences can be partly explained by: a) the Census 2000 CRS questionnaire used a somewhat different Hispanic question than that which appeared in the mail form, whereas the 1990 CRS used the same question as that used in the mail form, and b) the Census 2000 CRS used more Hispanic categories to derive the index of inconsistency than did the 1990 CRS. Thomas, et al note:

---

[27] While an estimated 5.7 million individuals provided a general Hispanic response in Census 2000, only 54 percent (3.1 million) of these people provided information in the long form nativity and ancestry questions.

[28] Bennett and Griffin (2002). C2SS is the 2000 variant of the American Community Survey.

[29] del Pinal (2003).

"the level of index is sensitive to the *number and detail* of categories in a classification system as well as to the *distribution of the population* over these categories." [30]

Nevertheless, these data add one more piece of evidence concerning the affect of the changes in the Census 2000 questionnaire with regard to reporting detailed and general Hispanic responses.

4.8. What is our overall assessment of data quality for the Census 2000 Hispanic question?

When viewed from the perspective that emerged following our evaluation of the 1990 census results, Census 2000 Hispanic data appear to be of a very high quality. Reversing the order of the race and Hispanic questions addressed the problem of nonresponse as shown by the relatively low allocation rates in 2000 (compared with 1990), as well as results from the Alternative Question Experiment. Refinements to the data editing process, particularly the new rule for imputing origin from people of the same race and imputing race from people with the same origin, dramatically reduced the artificial creation of relatively rare race/Hispanic origin combinations.

Despite the almost universal approval of the Hispanic population totals, it is clear that some of the changes introduced in Census 2000, such as the omission of the examples in the Hispanic question, probably encouraged respondents to provide general rather than detailed responses. This result casts a shadow on the quality of detailed Hispanic data. In the next section of this paper, will discuss efforts to address this problem as it relates to Census 2010 and the American Community Survey.

Given that the prime legislative mandate for Hispanic data is to provide an accurate count of the Hispanic population, and given the improvements the Census Bureau introduced in Census 2000 that largely addressed this directive, our overall assessment is that the Hispanic data quality is quite good.[31]

5. How are we addressing data quality issues for Census 2010 and how does this affect the American Community Survey?

When the Census Bureau conducted the National Content Survey (NCS) in 2003, the results of the evaluation studies discussed above played an important role in determining the issues the survey explored. The purpose of the NCS was to investigate and assess

---

[30] Thomas, Dingbaum, and Woltman (1993)

[31] This conclusion is supported in Schneider (2003) and del Pinal (2003).

proposed question changes, including those associated with race and ethnicity (Hispanic origin). The NCS contained experimental panels with differing versions of the Hispanic question. The various questions were composed of combinations of: a) examples for the "Other Spanish, Hispanic or Latino" category; b) the word "origin"; c) commas instead of slashes separating the words "Spanish," "Hispanic," and "Latino" in the question wording.

Preliminary results from the NCS indicate that both the inclusion of the word "origin" in the question wording and the inclusion of examples for the "Other Spanish, Hispanic, or Latino origin" category lead to a statistically higher percentage reporting a specific Hispanic group compared with the Census 2000 version of the question (the control panel) which had neither of these features. These test results raise the probability that the Census 2010 and ACS forms will include these key features with some minor modifications possible from future test results.

There are, however, a number of additional issues affecting the collection of data on the Hispanic population that will need to be addressed:

- *What examples should be included?* Generally, examples help respondents understand the intent of a question. However, specific examples potentially bias answers as some respondents may identify with an example instead of reporting actual group membership. Thus, inclusion of examples of groups with relatively small population totals (for example "Uruguayan") might result in a surge of reporting in this group. Testing panels to determine which examples and which order of presentation improve consistency and accuracy is difficult and time consuming and thus expensive.

- *What else can be done to increase the reporting of specific versus general Hispanic groups?* In addition to changes in question wording and inclusion of examples for the Hispanic write-in category, it may be possible, as shown in the study conducted by Cresce and Ramirez (2003), to use additional information such as ancestry and place of birth in the edit procedures to provide more specific Hispanic groups when a general response is given. This approach is possible with the American Community Survey which relies on an equivalent of the decennial long form (and thus includes nativity and ancestry information), but it will not be possible with the 2010 decennial census which relies on the short-form (and thus excludes the requisite additional questions). If the ACS relies on processing specifications that differ from those of the census, a problem of inconsistency arises.

- *What happens if there is an increase in the number of Hispanics who prefer to*

identify with the more general terms?   It is possible that the increase in the number of people reporting general instead of specific Hispanic responses may, in part, reflect a real trend toward identifying with more general terms. This is especially true for Hispanics who are one or more generations removed from an original immigrant ancestor, or whose ancestors lived in what became the territorial US after their ancestors had settled there (Spanish-Americans from New Mexico, for example).

Efforts beyond question wording and use of examples to encourage reporting of more specific Hispanic groups, raise an important philosophical issue. The Census Bureau has based its approach on counting the Hispanic population on the principle of self-identification. We do not fully understand all the reasons why people choose to respond the way they do. These responses reflect:

> "...the complexity underlying the reporting of ethnicity and highlight the problem of trying to simulate or 'second guess' the self-identification of respondents using other indicators of ethnicity. Trying to develop a composite measure of Hispanic ethnicity using a combination of responses from the Hispanic origin, place of birth, and ancestry questions undermines the principle of self-identification and can lead to endless discussion about who is 'Hispanic' and what is the size of the Hispanic population. In fact, the experience of using multiple indicators of Hispanic ethnicity in the 1970 census led the Census Bureau to decide that self-identification using a single question on Hispanic origin was the best method for counting this population."[32]

On the other hand, there is a legitimate concern among Hispanic groups to identify their group as fully as possible. We know that the Hispanic population is composed of diverse groups whose social, cultural, and economic characteristics may be quite different and for whom programs or policies need to be individually tailored .

These are difficult questions for which there are no easy solutions. Testing over the next two to three years will provide more information about changes in question wording and examples. Hopefully, the results of these tests, in combination with discussion of these results with key stakeholders, will help the Census Bureau develop an Hispanic question that will provide good data quality, not only for the total Hispanic population but also for the diverse groups that comprise the Hispanic population.

---

[32] Cresce and Ramirez (2003)

References

Bennett, C. and D. Griffin (2002). "Race and Hispanic Origin Data: A Comparison of Results from the Census 2000 Supplementary Survey and Census." *2002 Proceedings of the American Statistical Association, Section on Survey Research Methods*. American Statistical Association. Alexandria VA.

Costanzo, J., C. Davis, C. Irazi, D. Goodkind, R. Ramirez (2001). *Evaluating Components of International Migration: The Residual Foreign-Born*. U.S. Census Bureau, Population Division Working Paper No. 61. Washington DC.*

Cresce, A. (2002) "A Comparison of Editing Procedures for the Question on Hispanic Origin: 1990 Census and Census 2000." Paper presented at the Annual Meeting of the American Statistical Association, NYC.

_____ and R. Ramirez (2003). *Analysis of General Hispanic Responses in Census 2000*. U.S. Census Bureau: Population Division Working Paper No. 72. Washington DC.*

_____ R. Ramirez, and G. Spencer (2001). *Evaluating Components of International Migration: Quality of Foreign-Born and Hispanic Population Data*. U.S. Census Bureau: Population Division Working Paper No. 65. Washington DC.*

_____ S. Lapham, and S. Rolark. (1992). "Preliminary Evaluation of Data from the Race and Ethnic Origin Questions in the 1990 Census." Paper presented at the Annual Meeting of the American Statistical Association. Boston MA.

Deardorff, K. and L. Blumerman (2001). *Evaluating Components of International Migration: Estimates of the Foreign-Born Population by Migrant Status: 2000*. U.S. Census Bureau: Population Division Working Paper No. 58. Washington DC.*

del Pinal, J. (2003). *Race and Ethnicity in Census 2000*. U.S. Census Bureau: Census 2000 Testing, Experimentation, and Evaluation Program. Topic Report Series, No. 9. Washington DC.*

_____ (1996) "Treatment and Counting of Latinos in the Census," in Chabran Ed., *The Latino Encyclopedia, 1996*, Marshall Cavendish, NY.

_____ (1994). "Social Science Principles: Forming Race-Ethnic Categories for Policy Analysis." Paper presented at the Workshop on Race and Ethnic Classification: An Assessment of the Federal Standard for Race and Ethnicity Classification, National

Research Council, Commission on Behavioral and Social Sciences and Education, Committee on National Statistics.

_____ and Audrey Singer (1997). *Generations of Diversity: Latinos in the United States* Population Reference Bureau. [Bulletin 52:3].

Fernandez E. and A. Cresce (1986). "Who Are the Other Spanish?" Paper presented at the Annual Meeting of the Population Association of America, San Francisco, CA;

Gibson, C. and K. Jung (2002). *Historical Census Statistics on Population Totals by Race, 1790 to 1990, and by Hispanic Origin, 1970 to 1990, For the United States, Regions, Divisions, and States*. U.S. Census Bureau: Population Division Working Paper No. 56. Washington DC.*

Guzman, B. (2001). *The Hispanic Population.* U.S. Census Bureau: Census 2000 Brief. Washington DC.* [C2KBR/01-3]

Hobbs, F. and N. Stoops. *Demographic Trends in the 20th Century.* U.S. Census Bureau: Census 2000 Report. Washington DC.* [CENSR-4]

Logan, J. (2002). Hispanic Populations and Their Residential Patterns in the Metropolis. Lewis Mumford Center for Comparative Urban and Regional Research, University at Albany. [www.mumford1.dyndns.org].

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*. U.S. Census Bureau: Alternative Questionnaire Experiment. Washington DC.*

McKenney, N. et al (1985). "The Quality of Race and Spanish Origin Information Reported in the 1980 Census." Paper presented at the Annual Meeting of the American Statistical Association.

_____ et al (1988). "Development of the Race and Ethnic Items for the 1990 Census." Paper presented at the Annual Meeting of the Population Association of America, New Orleans, Louisiana.

_____ and A. Cresce (1992). "Measurement of Ethnicity in the United States: Experiences of the U.S. Census Bureau." Paper presented at the Joint Canada-United States Conference on the Measurement of Ethnicity, Ottawa, Canada.

Robinson, J. Gregory (2001). *Accuracy and Coverage Evaluation: Demographic Analysis*

Results. U.S. Census Bureau: DSSD Census 2000 Procedures and Operations Memorandum Series B-4. Washington DC.*

Singer, P. and S. Ennis (2002). *Census 2000 Content Reinterview Survey: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview*. U.S. Census Bureau: Census 2000 Evaluation B.5. Washington DC.*

Schneider, Paula (2003). *Content and Data Quality in Census 2000*. U.S. Census Bureau: Census 2000 Testing, Experimentation, and Evaluation Program. Topic Series No. 12. Washington DC.*

Suro, R. (2002). *"Counting the 'Other Hispanics': How Many Columbians, Domincans, Ecuadorians, Guatemalans and Salvadorans are there in the United States?"* Pew Hispanic Center, Washington DC. [www.pewhispanic.org]

Thomas, K., T. Dingbaum, and H. Woltman (1993). *Content Reinterview Survey: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview*. U.S. Census Bureau: 1990 Census of Population and Housing: Evaluation and Research Reports. Washington, DC.*
[CPH-E-1].

U.S. Bureau of the Census (1997). *Findings on Questions on Race and Hispanic Origin Tested in the 1996 National Content Survey.* Population Division Working Paper No. 16. Washington DC.*

_____ (1996). *Results of the 1996 Race and Ethnic Targeted Test.* Population Division Working Paper No.18. Washington DC.*

U.S. General Accounting Office (2003). *Decennial Census: Methods for Collecting and Reporting Hispanic Subgroup Data Need Refinement.* Washington DC. [GAO-03-228]

* Items available on the Census Bureau website: www.census.gov

Figure 1: Evolution of the Hispanic Question from the 1970 Census to the 2000 Census
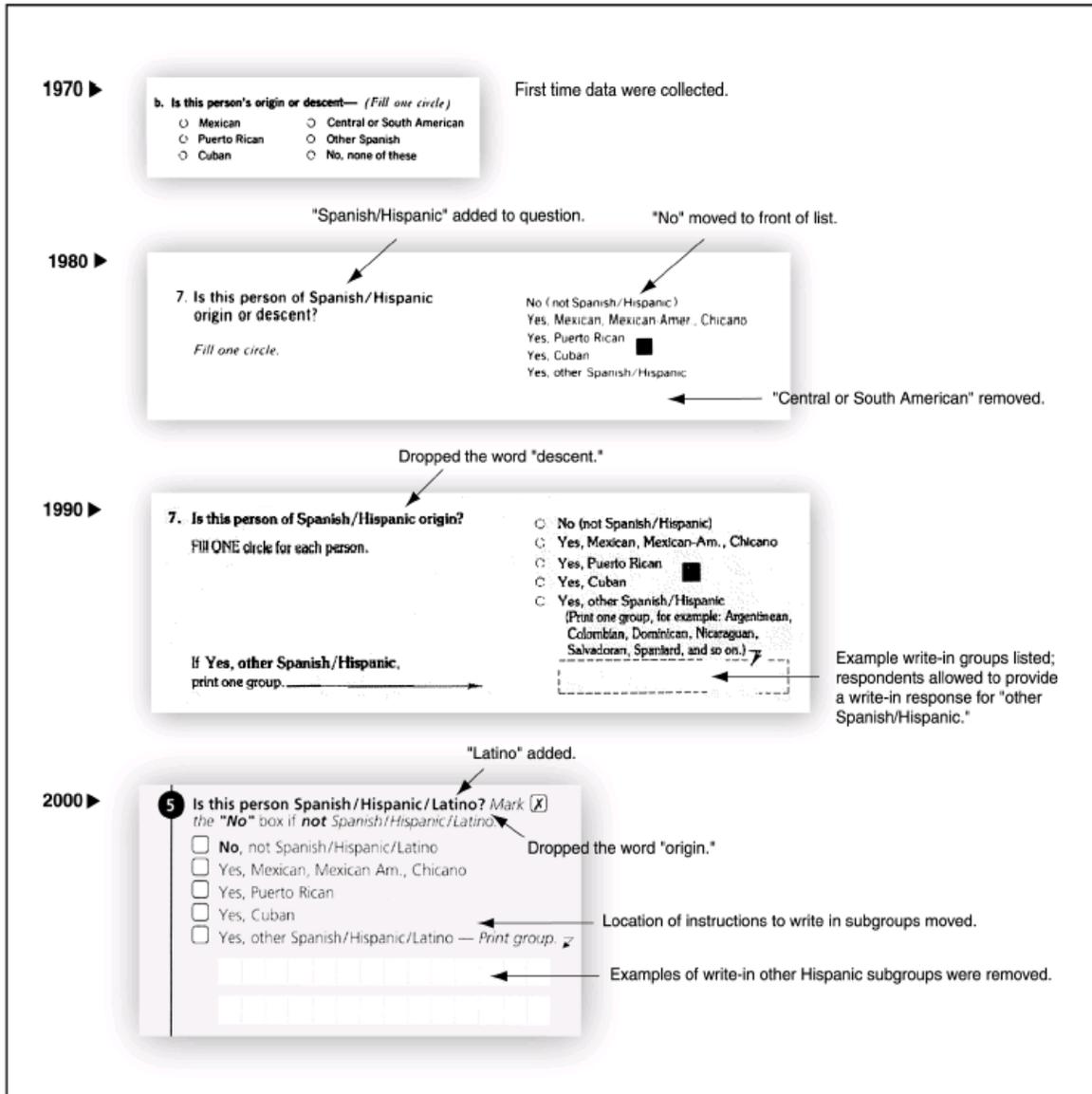
Source: U. S. Census Bureau.

NOTE: Figure 1 Reproduced from General Accounting Office Report GAO-03-228, "Methods for Collecting and Reporting Hispanic Subgroup Data Need Refinement," published February 992003.

Figure 2. Census 2000 Question on Place of Birth

12 Where was this person born?

○ In the United States — *Print name of state.*

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

Figure 3. Census 2000 Question on Ancestry

10 What is this person's ancestry or ethnic origin?

(For example: Italian, Jamaican, African Am., Cambodian, Cape Verdean, Norwegian, Dominican, French Canadian, Haitian, Korean, Lebanese, Polish, Nigerian, Mexican, Taiwanese, Ukrainian, and so on.)

Figure 4. RAETT Questions on Ancestry

Panel E. Combined Race, Hispanic Origin, and Ancestry Question With Mark One or More Boxes

Panel F. Combined Race, Hispanic Origin, and Ancestry Question With Mark One Box.

Table 1. Hispanic Population by Percent Change 1990 to 2000 for the United States, Regions, Divisions, and States

Regions, Divisions, and States	Hispanic Population		Percent Change 1990 to 2000
	1990 Census	Census 2000	
Total	**22,354,059**	**35,305,818**	**57.9**
Northeast	**3,754,389**	**5,254,087**	**39.9**
New England	568,240	875,225	54.0
Maine	6,829	9,360	37.1
New Hampshire	11,333	20,489	80.8
Vermont	3,661	5,504	50.3
Massachusetts	287,549	428,729	49.1
Rhode Island	45,752	90,820	98.5
Connecticut	213,116	320,323	50.3
Middle Atlantic	3,186,149	4,378,862	37.4
New York	2,214,026	2,867,583	29.5
New Jersey	739,861	1,117,191	51.0
Pennsylvania	232,262	394,088	69.7
Midwest	**1,726,509**	**3,124,532**	**81.0**
East North Central	1,437,720	2,478,719	72.4
Ohio	139,696	217,123	55.4
Indiana	98,788	214,536	117.2
Illinois	904,446	1,530,262	69.2
Michigan	201,596	323,877	60.7
Wisconsin	93,194	192,921	107.0
West North Central	288,789	645,813	123.6
Minnesota	53,884	143,382	166.1
Iowa	32,647	82,473	152.6
Missouri	61,702	118,592	92.2
North Dakota	4,665	7,786	66.9
South Dakota	5,252	10,903	107.6
Nebraska	36,969	94,425	155.4
Kansas	93,670	188,252	101.0
South	**6,767,021**	**11,586,696**	**71.2**
South Atlantic	2,132,751	4,243,946	99.0
Delaware	15,820	37,277	135.6
Maryland	125,102	227,916	82.2
District of Columbia	32,710	44,953	37.4
Virginia	160,288	329,540	105.6
West Virginia	8,489	12,279	44.6
North Carolina	76,726	378,963	393.9
South Carolina	30,551	95,076	211.2
Georgia	108,922	435,227	299.6
Florida	1,574,143	2,682,715	70.4
East South Central	95,285	299,176	214.0
Kentucky	21,984	59,939	172.6
Tennessee	32,741	123,838	278.2
Alabama	24,629	75,830	207.9

Regions, Divisions, and States	Hispanic Population		Percent Change
	1990 Census	Census 2000	1990 to 2000
Mississippi	15,931	39,569	148.4
West South Central	4,538,985	7,043,574	55.2
Arkansas	19,876	86,866	337.0
Louisiana	93,044	107,738	15.8
Oklahoma	86,160	179,304	108.1
Texas	4,339,905	6,669,666	53.7
West	**10,106,140**	**15,340,503**	**51.8**
Mountain	1,991,732	3,543,573	77.9
Montana	12,174	18,081	48.5
Idaho	52,927	101,690	92.1
Wyoming	25,751	31,669	23.0
Colorado	424,302	735,601	73.4
New Mexico	579,224	765,386	32.1
Arizona	688,338	1,295,617	88.2
Utah	84,597	201,559	138.3
Nevada	124,419	393,970	216.6
Pacific	8,114,408	11,796,930	45.4
Washington	214,570	441,509	105.8
Oregon	112,707	275,314	144.3
California	7,687,938	10,966,556	42.6
Alaska	17,803	25,852	45.2
Hawaii	81,390	87,699	7.8
Puerto Rico	---	3,808,610	n/a

Source: US Census Bureau, Census 2000 and Census 1990

For Census 2000, housing unit population and group quarters population are included.

For Census 1990, housing unit population only.

1990 totals do not include Puerto Rico because race and Hispanic Origin questions were first added to the PR census form in Census 2000.

Internet Release date: July 27, 2004

Table 2. Hispanic Origin Population by Detailed Group: 1990 and 2000

Hispanic Population by Origin Response	1990 Census [2/]		Census 2000 [2/]		Change 1990 to 2000	
	Number	Percent	Number	Percent	Number	Percent
Total	21,900,089	100.0	35,238,481	100.0	13,338,392	60.9
Mexican	13,393,208	61.2	20,900,102	59.3	7,506,894	56.1
Puerto Rican	2,651,815	12.1	3,403,510	9.7	751,695	28.3
Cuban	1,053,197	4.8	1,249,820	3.5	196,623	18.7
Dominican	520,151	2.4	799,768	2.3	279,617	53.8
Central American	1,323,830	6.0	1,811,676	5.1	487,846	36.9
Costa Rican	57,223	0.3	72,175	0.2	14,952	26.1
Guatemalan	268,779	1.2	407,127	1.2	138,348	51.5
Honduran	131,066	0.6	237,431	0.7	106,365	81.2
Nicaraguan	202,658	0.9	194,493	0.6	-8,165	-4.0
Panamanian	92,013	0.4	98,475	0.3	6,462	7.0
Salvadoran	565,081	2.6	708,741	2.0	143,660	25.4
Other Central America	7,010	0.0	93,234	0.3	86,224	1230.0
South American	1,035,602	4.7	1,419,979	4.0	384,377	37.1
Argentinean	100,921	0.5	107,275	0.3	6,354	6.3
Bolivian	38,073	0.2	45,188	0.1	7,115	18.7
Chilean	68,799	0.3	73,951	0.2	5,152	7.5
Colombian	378,726	1.7	496,748	1.4	118,022	31.2
Ecuadorian	191,198	0.9	273,013	0.8	81,815	42.8
Paraguayan	6,662	0.0	8,929	0.0	2,267	34.0
Peruvian	175,035	0.8	247,601	0.7	72,566	41.5
Uruguayan	21,996	0.1	20,242	0.1	-1,754	-8.0
Venezuelan	47,997	0.2	96,091	0.3	48,094	100.2
Other South American	6,195	0.0	50,941	0.1	44,746	722.3
Spaniard	519,136	2.4	112,999	0.3	-406,137	-78.2
General Hispanic	1,403,150	6.4	5,540,627	15.7	4,137,477	294.9
Hispanic	390,945	1.8	2,316,515	6.6	1,925,570	492.5
Latino	1,577	0.0	411,559	1.2	409,982	25997.6
Spanish	444,896	2.0	765,879	2.2	320,983	72.1
Other Hispanic response [1/]	565,732	2.6	2,046,674	5.8	1,480,942	261.8

Table 3. Allocation Rates for the Hispanic Question for the United States, Regions, and States: 1990 and 2000.[1]

Geographic Area	1990 Census [3]	Census 2000 [4]	Difference in Allocation Rates: 1990 and 2000
United States [2]	**10.4**	**5.6**	**4.8**
Region			
Northeast	11.6	5.8	5.8
Midwest	10.6	4.7	5.9
South	11.5	6.0	5.5
West	7.2	5.8	1.4
State			
Alabama	12.3	6.9	5.4
Alaska	4.6	5.5	-0.9
Arizona	7.0	6.3	0.8
Arkansas	6.6	5.5	1.0
California	7.2	6.1	1.1
Colorado	7.4	5.2	2.2
Connecticut	12.2	4.8	7.3
Delaware	9.7	7.0	2.7
District of Columbia	18.3	11.0	7.3
Florida	10.7	5.7	5.0
Georgia	13.7	6.9	6.8
Hawaii	8.2	7.7	0.5
Idaho	4.2	4.3	0.0
Illinois	11.9	5.8	6.1
Indiana	11.1	5.2	5.9
Iowa	10.1	3.5	6.6
Kansas	8.0	4.1	4.0
Kentucky	13.7	4.9	8.7
Louisiana	14.4	6.4	8.0
Maine	7.4	4.1	3.4
Maryland	12.2	6.6	5.6
Massachusetts	11.8	5.2	6.6
Michigan	11.2	4.9	6.3
Minnesota	9.0	4.0	5.0
Mississippi	15.2	7.0	8.2
Missouri	11.5	4.4	7.1
Montana	4.8	4.7	0.2
Nebraska	8.1	3.5	4.6
Nevada	9.0	6.5	2.6
New Hampshire	8.6	4.8	3.7
New Jersey	11.6	5.6	6.0
New Mexico	6.0	6.4	-0.3
New York	11.4	7.1	4.3
North Carolina	13.9	5.7	8.2
North Dakota	6.1	3.6	2.5
Ohio	10.9	4.3	6.6
Oklahoma	9.5	4.8	4.8
Oregon	8.2	4.5	3.6
Pennsylvania	12.3	4.9	7.5

Geographic Area	1990 Census [3/]	Census 2000 [4/]	Difference in Allocation Rates: 1990 and 2000
Rhode Island	14.7	6.2	8.5
South Carolina	13.7	6.6	7.1
South Dakota	7.4	4.1	3.3
Tennessee	12.3	5.4	6.9
Texas	7.9	5.9	2.0
Utah	7.3	4.2	3.1
Vermont	4.3	4.9	-0.6
Virginia	12.0	5.5	6.5
Washington	7.8	5.4	2.4
West Virginia	12.4	4.6	7.8
Wisconsin	9.7	4.6	5.1
Wyoming	5.0	5.3	-0.3

Source: U.S. Bureau of the Census, 2000 Census and Census 1990.

1/ Total Allocation rates do not include pre-edit procedures such as obtaining Hispanic origin from muliple ethnic origin
or from the race question.

2/ Total does not include Puerto Rico.

3/ For Census 1990, housing unit population only.

4/ For Census 2000, housing unit population and group quarters population are included.

Internet Release date: July 27, 2004

Table 4. Indexes of Inconsistency for the Hispanic Question and for Selected Origin Groups: Census Content Reinterview Survey Results, 1990 and 2000 [1]

Hispanic Question	1990 Content Reinterview Survey Index		2000 Content Reinterview Survey	
	Estimate	90-percent confidence Interval	Estimate	90-percent confidence Interval
Total	**12.2**	**11.2 to 13.2**	**17.2**	**16.1 to 18.4**
Not Hispanic	903	8.4 to 10.2	10.1	9.2 to 11.0
Mexican	8.5	7.5 to 9.7	13.4	12.2 to 14.8
Puerto Rican	8.6	6.4 to 11.5	14.2	11.5 to 17.6
Cuban	13.6	9.4 to 19.5	13.7	9.3 to 20.1
Other Hispanic	34.1	30.8 to 37.7	33.8	30.7 to 37.3

Source: Singer, Phylis, and Sharon R. Ennis (2002). Census 2000 Content Reinterview Survey: Accuracy of the Data for Selected Population and Housing Characteristics as Measured by Reinterview. Table 36. US Census Bureau: Census 2000 Evaluation B.5. Washington DC.

1/ The Census 2000 Content Reinterview Survey (CRS) used a test-retest methodology in which a sample of households from Census 2000 long form respondents were contacted a second time and re-asked most of the long form questions. The intent was to measure the simple response variance. The measure used to summarize this response variance is the index of inconsistency. The higher the index value, the more problematic is the interpretation of the data from the census item. Historically, an index value of less than 20 has been viewed as a low or a good level of response variance; an index between 20 and 50 as a moderate variance; and an index over 50 as a high variance.

Internet Release date: July 27, 2004