The 2000 Census: A New Design for Count and Content Data

Robert D. Tortora, *U.S. Bureau of the Census*

The research program for the 2000 census is described and options for collecting count and content data are outlined. Count data are used to reapportion the United States House of Representatives and are the major data source for developing legislative district boundaries under the Voting Rights Act in each state. In all previous censuses in the United States these data have been collected by an enumeration of the entire population. However, since 1940 the Census Bureau has measured a differential undercount in minority/ethnic populations and for certain geographic areas, such as large cities. Content data, used by a wide array of government agencies, businesses, and institutions, are defined as the data historically collected from a sample on the census long form.

This paper is divided into two parts: (*a*) design features for the 2000 census count that are independent of content and (*b*) design features and options for data content.

The choice of design features not affecting content will be selected after a 1995 census test in three locations (Oakland, California; Paterson, New Jersey; and six rural parishes in northwest Louisiana) throughout the United States. Decisions affecting content design are on a different path. The content determination process ends in April 1997 when the Office of Management and Budget (OMB) and the Census Bureau recommend the final content to Congress. Federal data needs will be the major driver of content, but the Census Bureau will also obtain input from other government entities and other users of census data. In April 1998 the Census Bureau will recommend final wording of census questions to Congress. During that time period the Census Bureau will obtain input on the options being considered for content collection. In addition to use of a sample long form in 2000, the Census Bureau is conducting research on other ways to collect content. One option is to use multiple sample forms in 2000. Another option is to collect content through a continuous measurement survey—a survey conducted each month during the decade instead of the sample long form.

Census Design Features Not Affecting Content

The 2000 census research and development program began, in 1991, to consider options on how to fundamentally change census methodology. The overarching goals of the program were

Opinions expressed in this paper are solely those of the author.

to develop design features that (*a*) reduce the differential undercount, (*b*) reduce or contain costs, and (*c*) provide for an open process by informing stakeholders of the current status of the research and obtaining their input and advice. However, the research and development program began because of the differential undercount and costs. Thus, the program was designed to fundamentally change census methodology to attain these first two goals. Why are fundamental change and these goals so important?

There are at least five motivating factors for fundamental change: changing societal trends in the United States, cooperation by the U.S. population in responding by mail to the census, the cost of the census, the differential undercount, and the two-stage decision strategy (i.e., deciding after census counts are provided for reapportionment and redistricting whether to adjust the census to reduce the differential undercount). These five factors are interrelated.

Societal Trends

Our society is changing. Between 1980 and 1990 the United States population grew by about 10 percent. However, the Asian and Pacific Islander population grew by 108 percent; Hispanics (of any race) by 53 percent; American Indians, Eskimos, and Aleuts by 38 percent; and African Americans by 13 percent. Other changes are occurring as more households are composed of unrelated individuals, the number of single adult households grows, and more individuals stay in more than one household. All of these trends make conducting the census with the methodology used in 1970, 1980, and 1990 more difficult and unlikely to improve the picture for differential undercount or costs.

Cooperation

Since 1970 the mail response rate has been declining. The 1970 rate was 78 percent, the 1980 rate was 75 percent, and in 1990 the rate was 65 percent. At the same time the number of housing units grew from 69 million in 1970 to 88 million in 1980 and finally to 102 million in 1990. So, in 1990 (as in 1970 and 1980) the Census Bureau was faced with sending enumerators to each household not returning a questionnaire by mail up to six times to collect the data. This nonresponse follow-up was the single most costly component of data collection—$450 million out of the total $1.3 billion for data collection.

Cost

The overall cost of the 1990 census was $2.6 billion, approximately $25 per housing unit. In 1990 dollars, the 1970 and 1980 censuses cost about $11 and $20 per housing unit, respectively. Assuming 1990 methodology and using estimates for housing unit growth, inflation, and productivity increases, the 2000 census will cost more than $4.0 billion. Especially in these times of fiscal restraint, this is a dramatic increase in cost that is unacceptable to many, particularly with the persistent differential undercount.

Differential Undercount

Since the Census Bureau began measuring undercount in 1940 (using demographic analysis, which estimates the population by starting with the last census count, adding births and immigration, and subtracting deaths and emigration), there has been a differential undercount of the African American population at the U.S. level. Table 1 compares African American undercounts with non-African American undercounts from 1940 to 1990 using demographic analysis.

Table 1 shows a declining but nearly constant difference in undercount rates from 1940 through 1980. Note the upturn in 1990.

TABLE 1 Differential Undercount of African Americans and Non-African Americans, 1940–1990, Based on Demographic Analysis

CENSUS YEAR	AFRICAN AMERICAN UNDER-COUNT	CHANGE IN AFRICAN AMERICAN UNDER-COUNT	NON-AFRICAN AMERICAN UNDERCOUNT	CHANGE IN NON-AFRICAN AMERICAN UNDERCOUNT	DIFFERENCE[a]
1940	8.4	—	5.0	—	3.4
1950	7.5	−0.9	3.8	−1.2	3.7
1960	6.6	−0.9	2.7	−1.1	3.9
1970	6.5	−0.1	2.2	−0.5	4.3
1980	4.5	−2.0	0.8	−1.4	3.7
1990	5.7	1.2	1.3	0.5	4.4

[a] African American minus non-African American.

In 1990 an independent survey, conducted immediately after the census, was also used to measure the differential undercount. That survey obtained estimates of undercount for various population groups (Table 2) as well as for various geographic areas.

Table 2 indicates the difficulty in counting racial and ethnic populations as accurately as the non-Hispanic white and other population. Since many members of these populations live in large urban areas or are concentrated in a few states, these undercounts translate to undercounts in these types of geographic areas. Even some rural areas with large minority populations, say parts of the south and southwest, have a differential undercount.

Cost

The 1990 census methodology has been in use since 1970. As noted earlier, the cost of the census was $11 per housing unit, $20 per housing unit, and $25 per housing unit for 1970, 1980, and 1990, respectively, standardized to 1990 dollars. During those decades the largest increase in cost on a per housing unit basis comes from 1970 to 1980, when costs almost doubled. Using the data from demographic analysis (Table 1), the differential undercount of African Americans decreased by 2 percent (the largest decrease since measurement of this phenomenon began), and the difference between the African American and non-African American undercounts dropped to 3.7 percent. Between 1980 and 1990, a different picture emerges. Costs went up by 25 percent, from $20 per housing unit to $25 per housing unit, and differential undercounts increased for both African Americans and non-African Americans, to 5.7 percent and 1.3 percent, respectively. Would a $40 per housing unit census have resulted in a different story for differential undercount? Would Congress have funded a $4.0 billion census in 1990?

TABLE 2 Undercounts for 1990 by Population Subgroup Based on the Postenumeration Survey

POPULATION	UNDERCOUNT ESTIMATE	SAMPLING ERROR
Non-Hispanic White and other	0.7	0.2
American Indians	4.5	1.2
American Indians on reservations	12.2	4.7
Asian and Pacific Islander	2.3	1.4
African American	4.4	0.5
Hispanic	5.0	0.7
United States	1.6	0.2

Two-Stage Decision Strategy

In 1990 the Census Bureau had a two-stage decision strategy with respect to the final counts. The Census Bureau conducted the 1990 census and provided the counts to President Bush on December 26, 1990. The law requires that these data, used to reapportion the House of Representatives, be provided no later than December 31 of the census year. By law, the states were provided voting rights data (race and ethnicity of those at least 18 years old by block) no later than April 1, 1991. After the completion of census data collection (Stage 1), the 1990 Postenumeration Survey was started. The results from that survey and the evaluation of it were studied by the Census Bureau to determine whether the counts should be adjusted (Stage 2) to reduce the differential undercount as well as undercounts for all levels of geography. In July 1991 Secretary of Commerce Mosbacher decided not to adjust the 1990 census (*1*). (An August 1994 decision by the Second Appellate Court in New York ruled that an adjustment should have been made. It is possible that the 1990 census will continue its judicial journey into the last half of the 1990s.) Part of the rationale behind the decision not to adjust was the turmoil adjustment might cause, since 1990 census data had already been used for reapportionment and redistricting.

Research Program for 2000 Census

The five factors led the Census Bureau to conclude that methodology used since 1970 needed to be changed in a fundamental way, particularly to deal with differential undercount and costs. Recognition that traditional counting methods could no longer be used resulted in the start of research and development to redesign census-taking methodology in 1991. This research phase alone represents major change. In preceding decades, the bureau entered into a planning stage, knowing the final design goal, about the fourth year into a decade. Now, faced with introducing fundamental change, the Census Bureau developed a research philosophy with the following goals:

1. Consider a variety of census design alternatives or options with a wide range of design features (the components of a census method).
2. Use the time before 1995 to conduct many small-scale experiments to provide data on various design features.
3. Use the design options and research results to select the most promising design features as building blocks for a design to be tested in 1995.
4. Obtain continuous stakeholder input on the research and development program.
5. Conduct the 1995 test census to inform final design decisions by the end of calendar year 1995. ("Inform" is used instead of "select" because other nontechnical factors may affect the final choice and because some features may be found acceptable but will require refinement before 2000.)

As a result of this research philosophy, 14 census design alternatives were developed (*2*). The alternatives covered a broad spectrum of census designs—from an administrative records census to a sample census. The alternatives included high-tech options for data collection and capture, ways to increase the mail response rate, making questionnaires widely available, stopping data collection earlier than usual, using administrative records to improve coverage, using improved enumeration techniques in historically undercounted (small) areas, improving the list of addresses used for mailing questionnaires, and (most important) using statistical sampling and estimation to account for households that do not respond by mail and integrating statistical sampling and estimation into the census process to reduce differential undercount.

Table 3 gives the design features selected for testing in 1995. The features are categorized into five major groupings: new uses of sampling and estimation, new procedures to count the undercounted, new uses of technology, new avenues for greater involvement, and a new method for collecting long-form data (the other new method for collecting long-form data,

TABLE 3 Fundamental Changes and the 1995 Census Test (U.S. Bureau of the Census)

FUNDAMENTAL CHANGE FROM 1990	MAJOR GOALS: REDUCE DIFFERENTIAL UNDERCOUNT	MAJOR GOALS: REDUCE COST
New Procedures To Count the Undercounted		
Use an easy-to-fill-out questionnaire with multiple mail contacts to improve response		X
Use revised questions to ensure a complete listing of household members	X	
Mail Spanish-language questionnaires to areas with large concentrations of Spanish-speaking households	X	X
Make census questionnaires available at convenient locations for those who did not receive a questionnaire or feel that they were not counted	X	
Use special targeted methods to count historically undercounted populations and geographic areas	X	
Count persons with no usual residence by a method that uses the places where these individuals obtain services	X	
Use administrative records to identify persons missed in the census	X	
New Avenues for Greater Involvement		
Develop cooperative ventures with other federal agencies; state, local, American Indian tribal, and Alaska Native village governments; and private and nonprofit organizations to form partnerships in taking the census	X	X
Evaluate initial efforts to compile and maintain an address list and geographic files in cooperation with the U.S. Postal Service and state, local, American Indian tribal, and Alaska Native village governments	X	X
Use the U.S. Postal Service to identify vacant housing units or mistakes on the address list		X
New Uses of Sampling and Estimation		
Use sampling and estimation procedures to reduce the differential undercount and the cost of the census. This means using an integrated coverage measurement sampling and estimation technique. Use sampling and estimation techniques for housing units that do not return questionnaires by mail	X	X
New Uses of Technology		
Use "real-time" automated matching to improve census coverage	X	
Use new technologies to contact persons or to allow them to contact the Census Bureau		X
Develop a new data capture system using electronic imaging		X
New Method for Collecting Long-Form Data		
Experiment with collecting sample (long-form) data using multiple sample forms		X

continuous measurement, is in a research phase outside of the 2000 research program and will be discussed later). Table 3 indicates which of the two overarching goals of reducing the differential undercount and containing or reducing the cost each feature attempts to address.

Many of these features appear in the 1995 test census because of the successful implementation of the research philosophy. For example, the easy-to-fill-out questionnaire with multiple mail contacts is being used because of a series of tests (*3*) conducted between 1991 and 1993. The tests indicated that a significantly higher mail response rate can be obtained (at least in a noncensus environment) if one uses first-class postage with envelopes indicating that response

is required by law, a prenotice letter, a respondent-friendly questionnaire, a reminder/thank you card, and a replacement questionnaire for those not responding to the initial questionnaire.

Many other field and simulation experiments were conducted and include such studies as how people live and attach themselves to households, how people understand census concepts and words such as residency and "living or staying," effect on mail response rates of Spanish-language questionnaires in areas with large concentrations of households that speak Spanish, availability and quality of administrative record system, uncertainty introduced by sampling nonresponding households, and the feasibility of integrating into the census process a statistical method to reduce the differential undercount (*4*). Before discussing the content collection possibilities for 2000, it is worthwhile to consider the consequences of the latter two statistical techniques, namely sampling for nonresponse follow-up and integration of statistical sampling and estimation for reducing the differential undercount. There will now be *measured* uncertainty along with the final census results. The uncertainty comes in the form of sampling error from both the sampling of housing units that do not respond initially by mail and from the method, called Integrated Coverage Measurement (ICM), to reduce differential undercount.

In 1995 two sampling procedures for those not returning a mail questionnaire will be evaluated—a block sample and a unit or address sample. In the block sample all nonresponding households will be enumerated either by telephone (if a telephone number can be attached to an address) or by a personal visit. For the address sample all households not returning a mail questionnaire will have a chance of being selected for the nonresponse follow-up sample. Three questions need to be answered. First, how much sampling error is introduced into the census at various levels of geography such as block, tract, and site level for each procedure? Second, how much bias is introduced at these geographic levels for those units not sampled? And third, what is the cost of each procedure?

An ICM technique (sometimes referred to as Censusplus) being studied in 1995 calls for an independent listing of housing units in blocks randomly selected into the sample. (The Census Bureau will also compute estimates using the capture-recapture methodology of the 1990 Postenumeration Survey. These estimates will be compared with the Censusplus estimates. If the Censusplus results do not indicate a potential to reduce the differential undercount, a modified postenumeration survey will be used in 2000.) The independent listing will be compared with the Master Address File (MAF) to identify missed, duplicate, or incorrect addresses. The original MAF will be the frame for the census. In the randomly selected ICM blocks, some housing units will return a questionnaire and others will not. For the former an independent computer-assisted personal interview, using a refined measurement instrument, will be conducted to establish the household count on census day. For the nonresponding housing units a computer-assisted personal interview will be conducted, again using a refined measurement instrument, to establish household counts on census day. These interviews should result in a more accurate count (the 1995 test results will have to establish this fact) in each sampled block. Then a statistical estimation technique will be used to "adjust" for the undercount in the census. If the 1995 test is successful, it appears feasible to design an ICM survey that will produce direct estimates reducing the differential undercount for each racial and ethnic population, each state, and many major cities. For these populations and geographic areas the reduction in undercount will be larger than the error added by sampling. For smaller areas the Census Bureau will have to use an estimation technique to carry down adjustments. For these smaller areas it will be extremely difficult to decide on a case-by-case basis whether the census count was actually improved.

However, there will be measures of the uncertainty for these smaller areas. In contrast, in all past censuses except 1990, there were uncertainties in the accuracy of the counts for these areas as well as large geographic areas and population groups, but these uncertainties were not quantifiable. In 1990 the Census Bureau produced estimates of the differential undercount and the uncertainty (see Table 2) in these estimates using the Postenumeration Survey. But these data were not available until July 1991 (see preceding discussion of dual strategy).

From Table 3, the last category of change being tested in 1995 is a new method of collecting long-form data using multiple sample forms. Another option being considered is a continuous measurement survey to replace long-form content. The remainder of this paper compares three options for content—the long form, multiple sample forms, and continuous measurement.

Design Features Affecting Content

As mentioned earlier, OMB and the Census Bureau will report, for congressional approval, on the content needs for the 2000 census in April 1997. Specific question wording for the content will be recommended to Congress in April 1998. These two reports are the only dates cast in concrete with respect to content and the 2000 census. The methodology for collecting content data is still open. At least four possibilities exist: (*a*) use a 1990-like sample and (respondent-friendly) questionnaire design but with reduced content, (*b*) use a 1990-like sample and (respondent-friendly) questionnaire design with about the same content length as in 1990, (*c*) use multiple sample forms in place of the long form for content, and (*d*) use a new monthly continuous measurement survey to collect content over the decade. The remainder of this paper will compare these options in terms of the advantages and disadvantages of each, including a (simulated) comparison of sampling errors for the multiple sample forms and continuous measurement options. For any of these options the count portion of the census would include the fundamental changes described earlier.

Reduced Content, 1990-Like Design

Under this design the census would have one short form and one long form (it is always assumed that any form used by the Census Bureau will be respondent friendly and that the improved mail implementation strategy will be used) with the latter having fewer content items relative to the length of the 1990 long form. The short form would collect data that must be obtained for all persons (or housing units) for reapportionment, redistricting, and other statutory requirements. Currently this includes the following data: roster of persons in each housing unit, name, age, sex, relationship, race, Hispanic origin, tenure, telephone number, and coverage probes.

Long-form questionnaires contain only data that are needed for only a sample of persons (or housing units). The reduction in content for the long form might take the form of the content in the 1995 test. Here, either a federal statute specifies that the decennial census must provide the data or federal agency statutes indicate that the census is the most appropriate source of the data. From the former the reduced long form might contain education, place of birth, citizenship, year of entry, language, income, number of rooms, year built, and farm residence. From the latter the reduced long form would also include marital status, disability, children ever born, veteran status, labor force status, place of work, journey to work, occupation, industry, class of worker, units in structure, rent, year moved in, number of bedrooms, plumbing, kitchen, telephones, number of vehicles, fuels, water, sewer, utilities, and ownership costs. On the basis of Census Bureau and OMB review of agency requirements, the following content items from 1990 would be dropped: ancestry, residence 5 years ago, work last year, year last worked, value of home, and condominium. (This content determination was based on a review of federal agency statutes that occurred primarily in late 1993 and early 1994. It does not mean that these items will have the same status in 2000.)

There are at least three potential advantages of a reduced-content long form. First, there is a potential for an improved mail response rate. The 1995 test will provide some insight into this potential, since three forms with varying lengths will be tested. Second, asking fewer questions reduces the burden placed on individual respondents in terms of time to complete the questionnaire. And third, some savings occur because of reduced data capture and processing costs. However, depending on the actual reduction in the number of questions, these savings might be small.

A reduced-content long form also has several disadvantages. Perhaps most important would be the loss of some data that many use. Second, because the 2000 census may incorporate sampling of housing units that do not return a questionnaire by mail, the sampling error for content data may increase somewhat. The remaining two disadvantages accrue to any census methodology that collects content only once a decade. Data collected once a decade become out-of-date. The median age of data from past censuses is 7 years. It is not uncommon for

important (funding) programs to use data at least this old or older. For example, in 1993 the federal government allocated education funds to states on the basis of 1980 census data with a 1979 reference year, using data 14 years old.

Finally, and particularly for small areas, sample data may reflect an unusual condition or temporary aberration, such as bad weather, that changes a person's behavior for the reference period. For example, unusual weather can affect data such as journey to work or mode of transportation. Such an anomaly is carried forward for all uses of the data over the decade.

Similar Amount of Content, 1990-Like Design

There would be one short form and one long form under this design. The main advantage of this scenario is that data users can expect to see about the same data content from 1990 to 2000. In addition, a respondent-friendly design should have a positive influence on initial mail response rates. This might only minimize the possibility of a further drop, or it may, when introduced with other components of the mail implementation strategy, bring about a small increase in response rate. Except that there is no loss of content, this design has the same disadvantages as the previous design.

Multiple Sample Forms

Under this design more than one sample form would be used to replace the traditional long form. Each form would have a different amount of content, although there would be certain questions common to all forms. The goal would be to develop several forms that, overall, produce the same amount of content as in 1990 without increasing respondent burden on any one form.

On the plus side, a multiple sample forms design can reduce individual respondent burden by asking fewer questions on many or all of the forms (when compared with the length of the traditional long form). To the extent that length of the questionnaire influences response, such a design may increase mail response rates.

There are at least three disadvantages of this design. First, multiple sample forms add operational complexities in properly fielding the design. The Census Bureau would have to develop a management control system that would ensure that the proper sample form was collected at the right address, particularly in nonresponse follow-up. Second, this design requires the development of new estimation and imputation techniques, since not all questions would be asked of each sample household. Third, sampling errors may generally increase under this design. Tables 4 and 5 present comparisons of relative sampling errors with 1990 long-form estimates for selected transportation variables. Finally, sampling error may increase for

TABLE 4 Illustrative Confidence Intervals Calculated for Selected Transportation Characteristics: 1990

CHARACTERISTIC	1990 ESTIMATE	CONFIDENCE INTERVALS (90 PERCENT)		
		CENSUS	MULTIPLE FORMS	CM
Percentage of workers using public transportation				
Forest Heights Town, Md.	12.9	+4.0	+5.0	+4.6
Census Tract 7044.01	20.4	+4.7	+5.9	+5.9
Mean travel time to work				
Forest Heights Town, Md.	33.0	+2.0	+2.5	+2.3
Census Tract 7044.01	30.6	+4.3	+5.4	+5.4
Persons per vehicle				
Forest Heights Town, Md.	1.27	+0.21	+0.26	+0.24
Census Tract 7044.01	1.04	+0.16	+0.20	+0.20

TABLE 5 Illustrative Confidence Intervals Calculated for Selected Commuter Flows for Persons Living in Washington, D.C.: 1990

		CONFIDENCE INTERVAL (90 PERCENT)		
PLACE OF WORK	1990 ESTIMATE	CENSUS	MULTIPLE FORMS	CM
Silver Spring, Md. (CDP)	3,134	+349	+436	+380
Hyattsville, Md. (CDP)	771	+174	+217	+189
Landover, Md. (CDP)	314	+111	+138	+121

some variables, not only because of the sampling for nonresponse follow-up, but also because some of the questions may be asked of fewer households.

Continuous Measurement

A continuous measurement (CM) design represents the most dramatic change for content collection. This design spreads content collection over the entire decade through a series of large monthly surveys. The current prototype design (*5*) envisions mailing questionnaires to about 250,000 addresses from the MAF each month. Each month a new sample of housing units, spread evenly across the country, will receive the questionnaire. Units that do not respond by mail, after several reminders, will be interviewed by telephone whenever the telephone number can be obtained.

Units not responding by mail or telephone will be sampled for a face-to-face interview at a rate of about one in three for most areas. A rate of about one in five will be used in remote areas. The total monthly interviewed sample size will be about 200,000 units, including vacant housing units; over a 5-year period the total number of interviewed housing units is about 12,000,000. In 1990, long-form data were obtained from 14,500,000 housing units.

The main objective of the CM design is to produce small-area (or small-domain) estimates that are better overall than those provided by the traditional long-form design. CM would provide estimates corresponding to any estimate that can be produced from the long form, including estimates for tracts, block groups, traffic analysis zones, school districts, and so forth, and small population subgroups consisting of about 0.1 percent of the population. The main differences between CM and the traditional long form are as follows:

1. The CM estimates will be an average over a 5-year period (3 years for 1999 to 2001 with a sample size of 400,000 per month).
2. The 5-year average will be updated annually.
3. The estimates from CM will have sampling errors typically about 25 percent larger than estimates from the long form.

Tables 4 and 5 compare simulated relative sampling errors from CM with sampling errors from the long form.

In addition, note that as the geographic level or the demographic subgroup size increases, CM will produce an abundance of estimates on a regular basis during the decade. Providing there is the need, one can imagine reliable estimates being produced quarterly, semiannually, or annually for states, large urban areas, congressional districts, and so forth. Table 6, extracted from Alexander (*5*), compares relative sampling errors for CM with 1990 long-form estimates for the estimate of the percentage of children 5 to 17 in poverty in Maryland for areas of different population size and for each of the eight congressional districts. The estimates from CM are based on 12- and 60-month accumulations.

The CM approach to content collection has several advantages. First, it simplifies decennial operations by allowing the entire effort to be directed at reducing the differential undercount and containing costs. Second, conducted over the decade, the CM model provides for updating

TABLE 6 Illustrative Comparison of Reliability Between Decennial Census and Continuous Measurement Data (Collection Systems for Areas in Maryland: Percentage of Children 5–17 in Poverty)

AREA	DECENNIAL CENSUS			ILF[a] CV[b] (PERCENT)	
	POPULATION SIZE	ESTIMATE	CV[b] (PERCENT)	12-MONTH[c]	60-MONTH[d]
Maryland total	4,781,468	10.5	1.1	3.2	1.45
Baltimore City	736,014	31.3	1.5	4.1	1.8
Anne Arundel County	427,239	5.3	5.6	15.7	7.0
Carroll County	123,372	3.6	10.0	32.3	14.7
St. Mary's County	75,974	9.4	9.2	N.A.	11.9
Gaithersburg	39,542	7.4	16.2	N.A.	21.2
Somerset County	23,440	15.6	14.0	N.A.	18.3
Kent County	17,842	12.9	14.9	N.A.	22.4
Hyattsville	13,864	5.6	25.8	N.A.	35.1
Harve de Grace	8,952	23.5	14.2	N.A.	19.6
Capitol Heights	3,633	7.2	46.5	N.A.	61.7
Cottage City Town	1,236	5.0	46.0	N.A.	103.8
Congressional District 1	597,684	10.2	3.2	9.1	4.1
Congressional District 2	597,683	6.3	4.2	11.8	5.3
Congressional District 3	597,680	11.9	3.0	8.4	3.7
Congressional District 4	597,690	8.0	3.6	10.4	4.6
Congressional District 5	597,681	4.7	4.7	13.9	6.2
Congressional District 6	597,688	8.3	3.5	10.2	4.6
Congressional District 7	597,680	30.2	1.6	4.7	2.1
Congressional District 8	597,682	4.1	5.2	14.8	6.6

NOTE: N.A.—Not applicable.

[a] Calculations of reliability for intercensal long-form (ILF) estimates are based on (*a*) a sample size 64 percent of that needed to provide reliability comparable with that from the decennial census and (*b*) no oversampling of governmental units under 2,500 population.

[b] The coefficient of variation (CV) is a measure of sampling variability. The CV is the ratio of the standard error of a sample estimate to its expected value. There is no specific rule to determine whether a given CV is good or not. This determination is based on considerations such as use of the data, consequences of making the wrong decision, and so forth. In practice, a CV of 10 percent or less is often considered to be adequate, between 10 and 50 percent to be acceptable, and 50 percent or more to be undesirable.

[c] Estimates are based on weighted observations from 12 months of interviews.

[d] Estimates are based on weighted observations from 60 months of interviews.

of the MAF. Third, the annually updated moving averages provide for regular updates of the estimates, instead of updates once a decade. Fourth, it creates a permanent data collection staff knowledgeable about CM methods and procedures. Fifth, it creates long-term efficiency gains for other household surveys such as the Current Population Survey. When needed, it also provides an efficient sample (frame) for rare subpopulations. Sixth, it allows for more uniform treatment of seasonal effects, seasonal resort areas, and aberrations due to adverse weather or other causes. Finally, it allows more flexibility in adjusting sample sizes when necessary, correcting for errors in estimates identified by independent local sources, conducting ongoing experiments to evaluate and improve the quality of the design, using variable reference periods to reduce recall error, and responding to new data needs as they arise.

On the other side of the coin, CM has disadvantages. First, over the entire decade CM will cost more than collecting traditional long-form data (precise cost estimates are not available at this time). Second, with the added sample size comes added total respondent burden. Third, data users will have to adjust to using moving 5-year averages. Fourth, it will have less complete coverage of housing units than the long-form survey. Fifth, it is possible to increase within-household undercoverage, as is the case with all current household surveys. Sixth, there is an inability to control tract-level estimates to short-form counts. Finally, income measurement inaccuracies may occur if the current recall period of income last year is retained over an annual cycle.

Conclusions

The 2000 census design program has been conducted with the aim of addressing the two major problems of continuing differential undercount and rising costs in past decennial censuses. To attain both goals, the design program has looked at fundamentally different ways of conducting the census, recognizing that past methods will not be able to overcome these problems. The two most important method changes are integrated coverage measurement and sampling for nonresponse follow-up. The former is designed to reduce differential undercoverage—the latter to reduce or contain cost.

Fundamental change includes not only new methods of improving housing unit and person coverage and reducing costs, but also different ways to collect content. In addition to the use of 1990-like options for the 2000 census, this paper discussed two other approaches—multiple sample forms and continuous measurement. The results from the 1995 census test will provide data important in making the final design decisions at the end of calendar year 1995.

Acknowledgments

The author would like to acknowledge the assistance of Art Cresce, Alfredo Navarro, and Charles Alexander in the preparation of this paper.

References

1. Mosbacher, R. Decision of the Secretary of Commerce on Whether a Statistical Adjustment of the 1990 Census of Population and Housing Should Be Made for Coverage Deficiencies Resulting in an Overcount or Undercount of the Population. *Federal Register*, 56, 1991, pp. 33,582–33,642.
2. The 2KS Memorandum Series. Design Alternative Recommendations: 1–14. Bureau of the Census, U.S. Department of Commerce, 1993.
3. Dillman, D., J. Clark, and J. Treat. Influence of 13 Design Factors on Completion Rates to Decennial Census Questionnaires. Presented at Census Bureau Annual Research Conference, Washington, D.C., 1994.
4. *Final Report of the Technical Committee of the 2000 Census Task Force*. Bureau of the Census, U.S. Department of Commerce (forthcoming).
5. Alexander, C. *A Prototype Continuous Measurement System for the U.S. Census of Population and Housing*. CM-17, Continuous Measurement Research Series, Bureau of the Census, U.S. Department of Commerce, 1994.