2010 CENSUS PLANNING: DEMOGRAPHIC AND TECHNOLOGICAL CHANGE ESCALATING COMPLEXITY

Ruth Ann Killion, Sally M. Obenski, U.S. Census Bureau; Lisa E. Buckley, IBM Corporation
Sally M. Obenski, U.S. Census Bureau, 4301 Suitland Road, Rm 1002/2, Suitland, MD 20746

Key Words: Changing Demographics, technological change, privacy

Introduction[1]

Census 2000 was a good census. It may have even been a great census. Despite an array of risks confronting its successful implementation, the Census Bureau was able to execute what appears to be the best census ever conducted. So why are we concerned about our ability to successfully confront the challenges facing us in 2010? Although discussed in depth below, the simple answer is the rapidity of demographic and technological change.

The Census Bureau remains committed to learning from prior censuses and striving for continuous improvement. Decennial planners have always encountered and overcome numerous policy, operational, technical, and fiscal challenges. Just as in other decades, demographic and technological changes have prompted us to make bold decisions and take innovative approaches:

- While planning the 1940 Census, we determined that it was neither cost-effective nor necessary to ask all questions of every person. The census sample was developed as a major enhancement to the 1940 Census to increase efficiency and accuracy and decrease respondent burden.
- In planning for the 1970 Census, we determined that we could no longer effectively enumerate a growing and increasingly complex population through personal visits alone. We reengineered the 1970 design by moving to a self-response mailing strategy as the primary data collection mode--a strategy still in place and designed to improve data accuracy and reduce the number of personal visits.
- Declining response rates led to the use of a multi-faceted marketing approach in Census 2000 that included paid advertising, an expanded "Census in Schools Program," and broad-based community partnership programs.
- Concerns about coverage and cost led to a simplified short form questionnaire and the option of responding to Census 2000 in several ways, including the Internet.
- To more efficiently and effectively process about one and a half billion pages of enumeration information on schedule for Census 2000, the design, development, and operation of the data capture system were out-sourced to private firms for the first time.
- Measuring a differential undercount led to incorporating coverage measurement surveys into the 1980, 1990, and 2000 censuses.

During the Census 2000 planning cycle, we came under intense scrutiny and austere budgets. The Census 2000 planning environment included continuing resolutions and a lack of consensus between the Congress and the Administration on census design. This lack of consensus resulted first in a requirement to develop two different operational plans; then, per the January 1999 Supreme Court decision, a third plan. Nevertheless, decennial managers concluded that the successful implementation of Census 2000 indicated a program that was well-executed and under control. Census 2000 was an operational success, in large part, due to a large infusion of funds at the end of the decade.[2] At the same time, this success was at a high cost and cannot be replicated.

Just as in past censuses, Census 2010 faces a number of challenges. However, the rapid change in population size, diversity, and public attitudes coupled with ongoing technological change creates a potentially high-risk environment. It is the exponential progression of change on multiple fronts that makes successful 2010 census-taking a formidable challenge. Success requires an adequate understanding of the issues and the complex relationships among them supported by comprehensive,

[1]This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

[2]Bureau of the Census Memorandum, "A Qualitative Assessment of the Success of Census 2000," October 6, 2000.

integrated planning. The purpose of this paper is to describe the major challenges that the Census Bureau recognizes for the 2010 census. While numerous others exist, the challenges presented herein represent areas that have significant effects on policy, planning, operations, and resources. Opportunities to meet these challenges and our strategy for 2010 are presented in subsequent papers.[3]

Census 2010 Planning Is Confronted with Escalating Complexities That Require--but May Lack--Resources to Resolve

Decennial census planners will be facing a daunting task in preparing for the 2010 census that may very well be mission impossible if resources are not available early in the decade. The planning environment for 2010 is confronted with exponential demographic and socioeconomic change as the size of population groups grow in number and influence. Language barriers and irregular and complex household structures increase enumeration complexity that, in turn, increase the possibility of not being enumerated. However, measuring this undercount via traditional methods may no longer be adequate, leading to the need for both greater record linkage and alternative response modes. This response, in turn, increases data collection, processing complexity, and privacy and security burden for the Census Bureau. Overlaying these challenges is the need for adequate, early, and predictable funding throughout the decade to mount sufficient resources with the right skill sets to address and resolve 2010 issues. The following six topics represent our primary focus for 2010:

1. Growing, Diverse Population Increases Enumeration Complexity.
2. Additional Methods for Measuring Undercount May be Required.
3. Increasing Number of Address and Enumeration Record Sources Complicate Data Collection and Processing.
4. Escalating Privacy and Confidentiality Concerns Can Further Complicate Data Collection.
5. Scarce Resources are available to Plan in a Dynamic Technological and Political Environment.
6. The Congress and Administration Must be Engaged and Reach Consensus Early on Key Design Components.

1. Growing, Diverse Population Increases Enumeration Complexity

As the demographic profile of the United States continues to grow and change, the Census Bureau is confronted with a wide and interdependent array of issues, increasing enumeration complexity. Nevertheless, the mandate of the decennial census is to enumerate every person resident in the United States, regardless of how hard-to-enumerate. The changing profile of emerging and growing population groups is prompting the Census Bureau to rethink the meaning of what has been called special populations and our response to groups and areas that may require special enumeration. For example, the Census Bureau must launch a response to increases both in numbers and usage of languages other than English that are spoken in some areas throughout the country and in irregular and complex household structures.

Changing demographics and increased public awareness and advocacy activities are quickly rendering the traditional definition of special populations inadequate. The successful Census 2000 awareness campaign has raised the expectations of various demographic groups in a number of types of non-traditional housing units. For example, the Census Bureau has initiated exploratory research to better understand the living situation of population groups who are residentially mobile. This population may require special enumeration techniques so we are interested in identifying and understanding patterns of residential mobility. The residential mobility of people who have two homes–one summer, one winter– and migrant workers is more predictable than that of persons whose residential mobility is dictated by factors other than weather and employment. The challenge for 2010 is to make our enumeration methods flexible to accommodate the life circumstances of a growing and complex population.

The Country Continues to Change

As seen in Table 1, with the growth in the size of the United States population over the past 60 years, the country has undergone dramatic demographic change. The primary changes have been the increase in age (e.g., baby boomers), the rapid growth of Hispanics (now equaling blacks), and the number and increasing size of Asian groups. With rapid immigration from Asia, Latin America, the Middle East, and Africa, there is an explosion of languages other than English that are spoken in some areas throughout the country. Additionally, the concept of racial identity has been in the forefront over

[3]Susan M. Miskura, "2010 Census Planning: Opportunities to Meet the Challenges," Draft, June 2001; Preston J. Waite, "2010 Census Planning: The Strategy," Draft, June 2001.

Table 1: U.S. Population Growth, Median Age, and Race Distribution: 1940 to 2000[4]

			Race						
Year	**Population**	**Median Age**	**White**	**Black**	**Amer Indian Alaskan Native**	**Asian Pacific**	**Hispanic**	**Some Other**	**Two or More**
2000	281,421, 906[5]	35.3	75.1	12.3	0.9	3.7[6]	12.5	5.5	2.4
1990	249,464, 396	32.8	83.9	12.3	0.8	3.0	9.0	NA	NA
1980	227,224,681	30.0	85.9	11.8	0.6	1.6	6.4	NA	NA
1970	205,052,174	28.0	87.6	11.1	NA	NA	NA	NA	NA
1960	180,671, 158	29.5	88.6	10.5	NA	NA	NA	NA	NA
1950	152, 271, 417	30.2	89.3	9.9	NA	NA	NA	NA	NA
1940	132,122, 446	29.0	89.8	9.8	NA	NA	NA	NA	NA

the last several years. There have been a number of efforts underway to simplify or improve race and ethnicity reporting. In the latter part of the last decade, a move was underway to include a multi race category on the census. Additionally, much research and analysis has been conducted on clarifying the difference in respondents' interpretation of race versus Hispanic Origin, including changing the order of the questions asking for this information on census questionnaires.

Just as in all other census planning efforts, for Census 2000, the Census Bureau responded to new and vocal advisory and advocacy groups weighing in on the need to address racial and ethnic enumeration issues. For example, the Asian and Pacific Islander Advisory Committee was formed as this groups' numbers and presence increased. In response to this group's requests and concerns in the 1990 Census, the Census Bureau agreed to include specific Asian subgroups on the census form. By Census 2000, in response to this and other race and ethnicity advisory groups, the Census Bureau agreed to target neighborhoods with non-English speaking populations and include both an English and non-English questionnaire as part of the mailing strategy.

It appears that the number of groups requesting or requiring some type of special or customized outreach or enumeration is likely to grow. However, some of the promises made in 2000, such as the targeted dual mailing, proved impossible to keep due to operational infeasibility or extreme difficulty. Currently, the Race and Ethnic Advisory Committees and the Secretary's Decennial Census Advisory Committee are recommending a targeted mailing for the 2010 census. If such requests for the short form to be available in multiple languages and requests for inclusion of additional population groups on the short form are not adequately managed, the short form itself may become unmanageable and inefficient. The point is that if pressure for customization continues, mass enumeration may be threatened.

Impact of Change

While it is difficult to separate the effects associated with demographic characteristics and change, several effects appear inevitable. First, as Table 2 illustrates, the aging of the population will lead to increases in the number and type of institutional facilities, such as nursing homes and

[4]Source: Bureau of the Census, *U.S. Census of the Population:* 1940, vol. II, Part I and Vol. IV. Part 1; 1950 Vol II. Part 1; 1960, Vol. 1,. Part 1; 1970, vol 1, Part B' Current Population Reports, pp. 25-1095; "Resident Populations Estimates of the United States by Sex, Race, and Hispanic Origin: April 1, 1990, to July, 1999"; published 24 May 2000; Bureau of the Census, Census 2000.

[5]Source: Bureau of the Census, Census 2000. Other population specific data are based on Population Estimates Program, Population Division, Bureau of the Census.

[6]Native Hawaiian and Other Pacific Islanders accounted for 0.1 percent and Asians 3.6 percent.

assisted living quarters. The former example is a classic special place requiring a special enumeration for group quarter occupants. The latter example is generally considered as a housing unit but there is a growing trend toward a more hybrid arrangement--part residence and part institution--that will increase enumeration complexity. Further, as the population ages, the requirements for response to the disabled (e.g, visually, hearing impaired) will likely occur on a larger scale.

Table 2: U.S. Population Age Distribution 2000[7]

Age	Number	Percent (%)
under 5 years	19,175,798	6.8
5 to 9 years	20,549,505	7.3
10 to 14 years	20,528,072	7.3
15 to 19 years	20,219,890	7.2
20 to 24 years	18,964,001	6.7
25 to34 years	***39,891,724***	***14.2***
35 to 44 years	***45,148,527***	***16.0***
45 to 54 years	***37,677,952***	***13.4***
55 to 59 years	13,469,237	4.8
60 to 64 years	10,805,477	3.8
65 to 74 years	18,390,986	6.5
75 to 84 years	12,361,180	4.4
85 years and over	4,239,587	1.5

Second, the number and types of languages other than English continues to grow. Traditionally, the Census Bureau has identified linguistically isolated households as being a barrier to enumeration. As mentioned earlier, the initial plan to address this barrier was to deliver a non-English questionnaire to communities that had large numbers of linguistically isolated households. What we are beginning to see, however, is the emergence of linguistically isolated communities. English is neither the dominant language spoken in the home nor in the community. Colonias are one example of such linguistically isolated communities. Colonias are unincorporated generally low income residential subdivisions lacking basic infrastructure and services along the border between the U.S. and Mexico. A recent estimate of the population in Colonias totaled about 1.2 million persons, and the numbers have been increasing over the last 15 years. However, until Census 2000, these areas lacked accurate official census information needed to improve the living conditions of the Colonias. For Census 2000, the Census Bureau tabulated data on Colonias as a separate geographic entity rather than having the information intermingled with data from other geographic areas as in 1990.

Introducing numerous languages other than English into every facet of the 2010 decennial census has profound policy and operational concerns. At a minimum, we will likely be asked to deliver more than what we promised in Census 2000--dual mailings, more forms in more languages readily available, and easy access to forms other than English via the Internet. However, having a census form in a non-English language addresses only part of the problem. There are populations that are also illiterate in their own language. For example, some Colonias residents were not able to deal with the Spanish language form provided in Census 2000 and needed additional assistance. Although it may seem that advances in telecommunications technology (e.g., Internet) may make communities less linguistically isolated, the argument could be made that such advances may actually contribute to linguistic isolation. For example, non-English language radio and TV stations make it possible to tailor broadcasting to the specific language needs of populations.

Third, while irregular and complex household arrangements pervade all segments of the population, growing numbers of new and existing population groups may be more likely to have such arrangements. In general, such households have one or more of the following features: (1) unrelated individuals, (2) mobile or ambiguous household members, (3) households formed for the sole purpose of sharing the rent and/of other living expenses, or (4) households that contain two or more "nuclear" families. Such households may not understand census rules of residence.[8] This can happen when household members have little or no knowledge of English or they are semi or entirely illiterate. Another

[7]Source: Bureau of the Census, Census 2000.

[8] Since the first decennial census in 1790, the concept of "usual residence" is the main principle in determining where people were to be counted. Usual residence has been defined as the place where a person lives and sleeps most of the time. This place is not necessarily the person's voting residence nor legal residence. All non-citizens living in the United States are included, regardless of their immigration status.

reason is the disjunction between how household is defined by the Census Bureau and how household is defined by respondents.

Such irregular and complex households require special enumeration techniques. For example, the Colonias along the southwestern border of the United States have many of the characteristics mentioned above. In Census 2000, the Census Bureau launched a special enumeration technique in which enumerators hand-delivered a form to the residents and enumerated them on the spot.

More work is required to fully understand the implications of the increasing population complexity to ensure improved coverage in 2010. Key policy, operational, and resource issues facing planners include:

> What is our strategy for dealing with such varied, large, and complex demographic and socioeconomic changes? How do we work with stakeholders but ensure that reasonable expectations can be met? How do we determine what situation requires an exception? What criteria do we use?
>
> There are many people asking whether it is appropriate for us to continue identifying the race of respondents. Others are pushing for single race identification while still others want more and more detail. What are the full implications of moving away from identifying race or moving toward a more detailed collection of race data?
>
> How do we translate the potentially large number of customization requests into something that is feasible for adequate research and testing prior to 2010?
>
> Will we have the resources to support needed special enumerations, including a substantially expanded language program? Will we have the right skill sets to implement them? Can we enumerate a growing number of population groups on schedule?

We are in the process of examining the expansion of the definition of what the Census Bureau has traditionally referred to as a special population to special types of enumeration. This expansion needs to encompass any group or area that requires some deviation from the mass enumeration approach and must address trends such as the growth of linguistically isolated communities. Addressing special needs is critically important to ensuring adequate coverage of all segments of the population. However, addressing special needs must be tempered by the requirement for standard procedures facilitating mass enumeration of the population.

2. Additional Methods for Measuring Undercount May be Required

While Census 2000 may have been the most accurate census ever conducted from a coverage perspective, managers and staff are in the midst of an extensive assessment of unexpected outcomes in our two key coverage measurement tools–demographic analysis and the coverage measurement survey. As illustrated in Table 3, from 1940 through 1990 the Census Bureau documented and measured a substantial differential undercount based on the statistical technique of Demographic Analysis.[9] The coverage measurement survey using dual system estimation was developed to more closely measure and potentially correct the undercount. Demographic Analysis benchmarks have been used to validate the coverage measurement survey estimates.

Table 3: Demographic Analysis of Percentage Net Undercount by Race: 1940 to 1990[10]

	1940	1950	1960	1970	1980	1990
Percent:						
Total	5.4	4.1	3.1	2.7	1.2	1.8
Black	8.4	7.5	6.6	6.5	4.5	5.7
Non-Black	5.0	3.8	2.7	2.2	0.8	1.3
Percent Difference:						
Black/ Non-Black	3.4	3.6	3.9	4.3	3.7	4.4

[9]Bureau of the Census, "Report to Congress–The Plan for Census 2000," originally issued July 1997, revised and reissued August 1997, pp. 2-6. Demographic Analysis is a statistical technique that measures trends and coverage by immigration, age, sex, emigration, and race using records and estimates of births, deaths, and Medicare enrollments to develop population estimates independent of the census.

[10]Bureau of the Census, "Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000", June 2000, p. 5.

The general trend was that Demographic Analysis estimates indicated a higher undercount than the coverage measurement survey estimates. One factor that confounded and complicated both independent population estimates in 2000 was the move to allow multiple responses to race. Demographic Analysis has been based on single race Black and non-Black populations. The Accuracy Coverage Evaluation (A.C.E.), was confronted with allocating varying race combinations into a manageable number of poststrata.

Demographic Analysis Estimates Were Inconsistent

We found that both the census and the A.C.E. were of very high quality, and the adjusted data are more accurate overall. However, concerns were identified, and we decided not to adjust for purposes of redistricting. Using demographic analysis as the baseline, we concluded that there was an apparent inconsistency between the A.C.E. and demographic analysis estimates. As illustrated in Table 4, the demographic analysis estimates of population are significantly lower than both Census 2000 and the A.C.E. estimates of population. The lower demographic analysis estimates resulted in net overcount percentages for both census and A.C.E. estimates.

One explanation that is being closely examined is that the number of undocumented immigrants is higher than we had estimated. However, this theory will not be able to be fully fleshed out until after the long form data on immigration become available. In addition to this primary concern, we also noted some concerns with the dual system estimates.

Dual System Estimates May Face New or Increased Error

While a discussion of the dual system estimation results is beyond the scope of this paper, dual system estimation results may be facing new or increased error due to an increasingly complex population. In other words, as demographic complexity increases, our assumptions about homogeneous behavior may need to change. The underlying assumption in dual system estimation is homogeneity or the synthetic assumption. Generally speaking, homogeneity or the synthetic assumption refers to the principle that people grouped in a given poststratum share the same coverage probability. Therefore, the results of the coverage measurement survey can be generalized from the large poststrata level down to small areas, including census blocks. These assumptions were the same ones we used in 1990. However, much demographic change has occurred since 1990 that could have affected these assumptions.

Why were the Demographic Analysis estimates so inconsistent? What are the full implications of the multiple response to race and the growth of Hispanics on Demographic Analysis, which focuses on administrative records denoting Black and non-Black at this time?

What are the full implications of changing demographics, including the increasingly multi-racial aspect of the population on dual system estimation?

In summary, we are currently analyzing all available data to better understand the underlying error structures in both Demographic Analysis and in the A.C.E. What is obvious is that continuing with a robust research and testing program for these important coverage measurement tools as well as examining innovative alternatives is imperative to a successful 2010 census.

3. Increasing Number of Address and Enumeration Record Sources Complicate Data Collection and Processing

One of the most daunting technical tasks facing the Census Bureau is determining methods to deal with the array of errors that emerge as a result of using multiple sources for address compilation and enumeration. Over the last decade, the Census Bureau has been taking advantage of existing sources of information as well as making it simple to respond to the census through providing multiple modes of response.

However, the use of multiple sources increases the probability that an entry has been collected more than once. For example, the Census 2000 address list was developed through the merging of multiple lists derived from multiple modes--the 1990 address file, the United States Postal Service Delivery Sequence Files, local governments' lists, and lists developed by address listers. We became concerned that the address list might contain a significant number of duplicates--housing units included more than once. To respond, decennial managers designed a complex, unplanned operation to try to identify and delete housing units that had been improperly duplicated in the address list. As a result, 2,411,743 addresses were identified as potential duplicates. Of these, 1,392,686 addresses were permanently removed from the address list and 1,019,057 were reinstated and included in the census results. Work is ongoing to fully understand how these duplicates remained in the census address list and to determine how to preclude this from

Table 4: Census 1990 and Census 2000 Estimates of Population and Percent Net Undercount[11]

	1990		2000	
	Number	**Percent (%)**	**Number**	**Percent (%)**
Demographic Analysis**	253,393,786	NA	279,598,121	NA
Total Census	248,709,873	(1.85)	281,421,906	(-.65)
Coverage Measurement Survey	252,712,820	(0.3)	284,683,782	(-1.8)

** Demographic Analysis represents the baseline; A minus sign (-) denotes a net overcount.

continuing to occur. This is important because we will continue to use the same sources as well as other administrative records to populate the address list.

In addition to duplicate housing units, we were concerned that the availability of multiple modes of response during Census 2000 could lead to a large number of duplicate persons. In addition to receiving a questionnaire in the mail, we offered several other opportunities for respondents to be enumerated. Three modes--the Internet, Telephone Questionnaire Assistance, and blank forms called Be Counted were employed. All three were very promising programs, especially the Internet but were kept small.

The Be Counted Program was initially designed to offer short forms in readily accessible locations throughout communities in a number of languages to increase coverage. The Internet response program was intended to provide respondents with the option of using the Internet to respond to the census. These were both very promising programs intended to improve coverage. But, we had grave concerns about the introduction of duplicates and our ability to process large numbers of responses in a timely, effective, and secure manner. Each Be Counted Form, for example, had to be matched and linked to any other forms that came from the designated housing unit. This means that the form must first be sent to geography staff to geocode it--adding time and complexity. If it was filled out in a language other than English, it required additional initial processing for translation or transcription into English. After all form processing was completed, the respondent(s) contained on the Be Counted Form were matched to all other respondents on forms that had the same housing unit identification. The management and processing of Be Counted Forms, even in small numbers, was resource-intensive and error-prone. As for the Internet, its use was in question almost up until the census itself. Given that we did not have the time to adequately test its use on a large scale or to ensure full protection of large numbers of responses, we did little to promote this option.

Another reason for concern about duplicate persons was the enumeration of people living in special places (e.g., nursing home). After gathering data on the special place, we made personal visits to enumerate its occupants. As expected, with no housing unit identifier and changing occupants, the potential for duplicates was high and unduplication efforts are difficult.

The complexity that emerges with the advent of using multiple sources for address listing and multiple modes of enumeration present challenges to 2010. The following summarize our concerns:

> How can we harness the power of both computer-based artificial intelligence and statistical methods to develop better models and algorithms for identifying and resolving duplicates? What are the differences among duplicates in the Master Address File compared with the census at large compared with special places and how should they be addressed?
> How do we minimize the processing complexity as numbers of potential entries increase, including mixing of duplicate responses in different languages?

We know that the desire for improved coverage and reduced cost will drive more and more merging of lists and sources to create our address list and multiple modes of enumeration. The use of administrative records, for example, is well-known in address building given the use

[11]Bureau of the Census, "Accuracy and Coverage Evaluation: Demographic Analysis Results, dated February 28, 2001, pp 24,39.

of the Delivery Sequence Files and local governments' lists. Other types of administrative records are being considered to augment and validate the list. Likewise, wide use of the Internet is being considered as a key data collection mode in 2010. However, numerous privacy and security concerns regarding increasing numbers of lists, sources, and modes must first be considered.

4. Escalating Privacy and Confidentiality Concerns Can Further Complicate Data Collection

In collecting and providing timely, relevant, and accurate data about the people and economy of the United States (i.e., the Census Bureau's mission), the Census Bureau is required by law to protect respondents' confidentiality and to use the collected data only for statistical purposes. Further, the Census Bureau is ethically obligated to respect respondents' privacy to the maximum extent possible. However, the increased availability of personal information in electronic data repositories coupled with computing and data linkage advances, in part, has heightened public sensitivity to privacy and confidentiality concerns. Thus, the natural tension that exists between the Census Bureau's mission and privacy and confidentiality concerns is also increasing. While this challenge exists across all Census Bureau census and survey activities, it is a significant challenge that needs to be resolved for a successful 2010 census design.

Declining public cooperation with the census and surveys has been attributed, in part, to increasing public concerns about privacy and confidentiality.[12] It is an issue that is as shrouded in perception as it is fact, and perception is a known key indicator of behavior. Census Bureau researchers found that privacy and confidentiality concerns became increasingly important barriers to public cooperation in the 1990 census and contributed to lower than anticipated mail response rates.[13]

Research concluded that the public's perception of the census as an invasion of privacy rises during the course of each census; however, there is evidence that the privacy reaction was greater during Census 2000 than it was during the 1990 census.[14] Specifically, this more recent study found the following:

> The experience of the census itself strongly influenced public concerns about privacy and confidentiality.
>
> The level of concern increased after mail out and was greater among long form recipients.
>
> There is evidence that people who heard about the widely publicized long form controversy became sensitized to the privacy issue.
>
> Data suggest that public reaction to Census 2000 differed from 1990, with several indications that Census 2000 engendered more sensitivity and a more diffuse privacy reaction than in 1990. For example, people increasingly came to believe that their answers could be used against them, a belief that was not apparent in 1990.

Privacy and confidentiality concerns are being further fueled by increased concerns and perceptions that better, more-sophisticated technology make it easier and easier to access personal information. The increased visibility and access to the Internet has exacerbated this issue. While there are legitimate security issues associated with electronic media, these issues may be resolved given advances in secure communications and encryption technologies. A more formidable obstacle may be public perception that respondent information collected and transmitted over the Internet, for example, will not be kept confidential. Finally, any widespread use of administrative records to augment the 2010 census may set off a firestorm of controversy over perceptions that dossier-like information is being used to profile individuals. As a result, the Census Bureau recognizes that significant policy issues are emerging:

> While we have some evidence of the short-term trends, are there long-term trends in public concern about privacy and confidentiality?
>
> How do we address privacy and confidentiality concerns since evidence suggests that raising the issues of privacy and confidentiality may backfire and increase concern rather than allay it?
>
> How do we resist possible public pressure from advocacy groups to increase the length of the short-form that would further fuel privacy concerns?
>
> How do we address the public's reaction to responding to the census, including use of the Internet and other response modes?
>
> How do we address the public's perception of the use of administrative records?
>
> How do we address the public's perception of providing long form data via the American Community Survey (ACS)?
>
> How do we measure/account for respondent

[12] Singer, Hippler, and Schwarz, 1992.

[13] Fay, Bates, and Moore, 1991; Kulka, et al, 1991, Singer, Mathiowetz, and Cooper, 1993).

[14] Martin, October 2000.

behavior related to privacy and confidentiality concerns?

Public attitudes are difficult both to measure and to predict. With respect to censuses and surveys, it is respondent behavior that matters most, and the more that the Census Bureau has studied privacy and confidentiality issues, the more we realize that we do not yet fully understand respondent behavior related to these issues. Without an appropriate understanding and approach toward addressing privacy and confidentiality issues, any major public outcry about a census design approach, such as the use of administrative records to augment data collection or massive use of the Internet, could result in last-minute design changes for 2010.

5. Scarce Resources are available To Plan in a Dynamic Technological and Political Environment

Empirical evidence, lessons learned, best practices, and common sense point to the fact that a systematic, building-block approach to designing the census beginning early in the decade is the optimum planning environment. Nevertheless, Census 2000 planners faced austere budgets, continuing resolutions, and scarce resources. As mentioned earlier, because the Congress and the Administration could not come to an agreement on the decennial design, the Census Bureau found itself with inadequate resources to staff even one design effort, yet were faced with planning for two designs--with and without sampling applications. Therefore, planning for Census 2000 faced an inherently unstable design.

Census 2000 resources were austere until 1998 which did not provide enough time to develop a strong staff knowledge-base, solid contingency plans, or alternative testing. The decennial planning staff was small and over-extended. Staff faced the prospect of implementing a complex Census 2000 design without the benefits of adequate resources to thoroughly research, test, and integrate the numerous components of the desired design. The design strategy included harnessing the power of technology to streamline and improve processes. However, the lack of in-house expertise and available funding for outsourcing coupled with an unstable design largely precluded this strategy from being enacted. Staff were concerned about making sure that individual, mission-critical functions were defined for either design contingency. However, a true, robust Census 2000 architecture design was never engineered. While some systems were well-planned and worked as intended, the overall architecture was, in large part, a collection of independent and inconsistent systems that had to be cobbled together.

In spite of this monumental effort by staff, substantial change was again imposed after the January 1998 Supreme Court decision. The Census Bureau found itself facing the development of yet a third design--one with no sampling in the enumeration phase but including a coverage measurement survey. This third design was never tested prior to the census since the dress rehearsal included the prior two options. For 2010, notwithstanding the possible existence of optimum budgets and plans, the importance and magnitude of the decennial census will require planners to be poised for such externally imposed design changes. Resource-related issues we face include:

> How do we deal with the inherent complexity of long-lead time architectural development and implementation given the scarcity of such in-house knowledge, skills, and abilities?
>
> What strategies should be used to address rapidly changing technology in order to strike a balance between availability but not obsolescence? How do we approach planning for such sophisticated development and testing techniques (e.g., spiral, rapid prototyping)?
>
> How do we strike the balance between growing integration issues due to the growing dependence on technology while still planning for contingencies?

Given adequate time and resources, we have tremendous confidence in the ingenuity and adaptability of our staff to meet these design and technical challenges. Our concern is that we are even more dependent on technology for infrastructure and for facilitating collection and processing methods for 2010 than in 2000. Such dependence requires substantial early planning that will not be possible without adequate and steady funding over the decade.

6. The Congress and Administration Must be Engaged and Reach Consensus Early on Key Design Components

The lack of consensus between the Administration and the Congress coupled with a lack of early funding to support required research and testing were major risk factors for Census 2000. As demographic and technological complexity escalates, our ability to successfully execute a census that has not been well-planned and tested diminishes. Because of the long lead time of a decennial census, it is difficult to get policymakers to focus on early design work and the need for agreement on major design methods. It is also difficult when the administration and majority party

change during the cycle.

Since the planning cycle for a decennial census is 13 years long while the funding cycle historically spikes toward the last few years in the decade, this problem is exacerbated. Historically, the federal budget has not supported the critical years coming off a decennial census where planners can take full advantage of the lessons learned from the just-completed census. The total funding for Census 2000 was more than adequate. The problem was the distribution of funds. During the critical planning years up to final preparation for the Census 2000 dress rehearsal in 1998, decennial planners received only about three percent of the total life cycle appropriation.

The lack of consensus and engagement between the Congress and the Administration and the lack of adequate funding to support Census 2000 planning led to a census implementation fraught with risk. The design was not finalized until after the dress rehearsal, resulting in some software being developed during production activities, escalating the risk of processing errors. Further, as mentioned previously, although the Internet was a promising technology, its inclusion was not resolved until very late in the decade. Therefore, it was not well-advertised for fear that Internet responses could not be fully managed and protected.

We were able to manage much of the risk because of the enormous infusion of funds during execution. The full cost of Census 2000 will be about $6.5 billion dollars. Our concern is how we approach the Administration and the Congress about full funding for all three strategic program components--the (MAF/TIGER)[15] modernization, full implementation of the ACS, and early planning for 2010 census design given the following:

> How do we get the Congress and the Administration to focus on something that is nine years away?
> How do we make the case for full funding given the reality of net present value and the long lead time in decennial development? Any costs incurred now will be difficult to recover in savings later in the decade.
> How do we make the case for funding and consensus now given that all through the latter part of the 1990s we were deemed high risk by the General Accounting Office and the Inspector General but managed to conduct a really good--perhaps great--census in 2000?
> As more and more lists or sources are used, how do we determine when more errors are being introduced than removed and the structures of those errors?

Conclusions

The fortuitous set of circumstances underlying the success and the challenges of exponential demographic and technological change preclude repeating Census 2000 in 2010. The challenges facing 2010 decennial planners are significant and growing in complexity. The Census Bureau will be confronted with requests and demands for more tailored, customized approaches to deal with real and perceived enumeration challenges among population groups. An aging population may require a different array of collection methods. Demographic analysis and coverage measurement surveys may have to be rethought, given changing demographic and socioeconomic characteristics. Using multiple sources for address listing and enumeration responses concurrently resolve and create issues, not the least of which is growing privacy concerns.

Finally, without adequate, early, and predictable funding resources to support 2010 planning, the appropriate set of knowledge, skills and abilities required to develop the necessary the 2010 design, procedures, and infrastructure to resolve these challenges may not exist. Given the complexity of change, the Census Bureau is facing a large and complex task. However, responding to change is an integral part of decennial census planning. Therefore, just as in other decades, we are addressing the challenges by taking advantage of multiple opportunities and by adopting a bold and innovative strategy for reengineering 2010.

[15]Master Address File/Topologically Integrated Geographic Encoding and Referencing.